Symbolic Dynamics for Geometrically Finite Groups

by David Fried*

In loving memory of Julius Fried

Abstract. *We use a horoball packing of hyperbolic n-space to encode geodesics. Given a geometrically finite group of isometries that is not convex cocompact, we obtain a symbolic dynamics for the recurrent part of the corresponding geodesic flow. Often this is based on a Markov partition with a finite number of transversals.*

We will develop a symbolic dynamics for geodesics on any finite volume noncompact quotient H^n/Γ of hyperbolic n-space by a discrete group of isometries. When $n = 2$ and $\Gamma = PSl(2, \mathbf{Z})$, this relates the geodesics on the modular surface to least-remainder continued fractions, much as simple continued fractions arise in Artin's description of these geodesics [Ar]. When $n = 3$ and Γ is the Bianchi group $PSl(2, O_k)$ of a quadratic imaginary field k, this gives a reduction theory for complex binary quadratic forms [F3]. More generally, say that Γ is geometrically finite but not convex cocompact and let $L(\Gamma)$ be the limit set of Γ. We find a symbolic dynamics for those geodesics on H^n/Γ that are covered by lines in H^n asymptotic to $L(\Gamma)$, in particular for closed geodesics.

Our interest in such codings comes from their use in the study of dynamical zeta functions [R, M, F1, F2]. Our methods, however, are suggested by Hurwitz's work on least-remainder continued fractions and reduction of real binary quadratic forms [H] as well as Bowen's theory for compact hyperbolic flows [B4].

Consider the "test case" where Γ is a discrete group of isometries of H^2, H^2/Γ has finite area and just one cusp, and the stabilizer of ∞ is the group $\mathbf{Z}$ of integer translations. The *Ford region* F consists of all points in H^2 whose height is at least as great as any point in their Γ-orbit. F is bounded by a family of smooth arcs C, each of which admits an edge-pairing g to another arc, determined up to composition with a translation. When we factor F by the action of $\mathbf{Z}$ and identify the boundary arcs by their edge-pairings, we get the orbifold H^2/Γ. We will use these same edge-pairings and translations to describe the geodesic flow $\overline{\phi_t}$ over this orbifold. Each C lies above a certain interval I in $\mathbf{R}$. We transform each point of I, except $g^{-1}(\infty)$, by g to give a point that is well-defined up to the action of $\mathbf{Z}$. One so obtains a piecewise smooth map from $\mathbf{R}$ to the circle $\mathbf{R}/\mathbf{Z}$, independent of the choices of edge-pairings, with the caveats that the map is undefined at points $g^{-1}(\infty)$ and ambiguously defined at the endpoints of I. This map induces an expanding map E from the circle to itself. Much as in [BS], $\overline{\phi_t}$ can be described using the orbits of E (see §8). In particular, periodic orbits of E correspond to closed geodesics on H^2/Γ.

When $\Gamma = PSl(2, \mathbf{Z})$, there is only one boundary arc C up to translation. Taking C on the unit circle, the projection of C to $\mathbf{R}$ is $[-1/2, 1/2]$ and we may take $g(w) = -1/w$ to be our edge-pairing. Thus

(0.1) $$E(x + \mathbf{Z}) = -1/x + \mathbf{Z},\ \ 0 < |x| \leq 1/2.$$

As shown in §2, E arises naturally when discussing least-remainder continued fractions.

* Mathematics Department, Boston University, 111 Cummington St., Boston MA 02215
df@bu.edu

§1 and §2 examine horoball packings of H^n and relate the Ford disc packing of H^2 to least-remainder continued fractions. We use fan tilings of H^n that, at least for finite volume quotients H^n/Γ, are dual to those studied by Epstein and Penner [EP]. The boundaries of the fans have long been used to study the cohomology of discrete groups and there are analogous tilings for other locally symmetric spaces, as in [As], [Me], [Vo], [L], and [Sa]. §3 and §4 introduce shadow families and aimed sequences and prove the two theorems that were announced in [F3]. A shadow family generalizes the family of level sets of the nearest integer relation. An aimed sequence generalizes the sequence of convergents to a least-remainder continued fraction. Eichler employed similar sequences to study orbits of Fuchsian groups on H^2, using Dirichlet domains instead of fans and aiming at a point of H^2 rather than a boundary point [E].

§5, §7, and §8 construct transversals and flowboxes and find systems of expanding maps that describe geodesics. We begin to discuss symmetry groups in §8. §6 discusses special features of horoball packings of H^2. §9 and §10 define chain-finite shadow families and show that they yield Markov partitions and a theory of symbolic dynamics, much like the known theory for compact quotients [B4].

§11 concerns a geometrically finite group Γ that is not convex cocompact. The recurrent part $\overline{\phi_t}$ of the geodesic flow over H^n/Γ is described using a piecewise smooth expanding map on a compact space with a natural invariant measure. §12 discusses an iteration scheme for computing the implicit factors in our Markov partitions.

§13 and §14 give criteria for shadow families to be chain-finite. §14 produces Markov partitions for all finitely generated Fuchsian groups. The criterion of §13 is used on various examples in §15−18, treated in order of increasing dimension. §19 poses some open questions.

Appendix 1 proves a property of geometric measure used in §11. Appendix 2 gives instances where the ridge shadow families of invariant horoball packings fail to be chain-finite. Appendix 3 treats marked sequences, which have an advantage over aimed sequences when there are finite order symmetries.

We are grateful to Paul Gunnells for many useful suggestions and to Genya Zaytman for the quaternion description of the conformal group of S^4. We are indebted to Rufus Bowen and Dennis Sullivan, whose ideas are used freely in this paper, and to Julius Fried for his lifelong encouragement.

Table of Contents

Section 1. Horoball packings and measures

Let Q be an infinite subset of the boundary ∂H^n of hyperbolic n-space and let L be its closure. We suppose given a horoball $B(q)$ based at q for every $q \in Q$ so that no two horoballs overlap. We call such an infinite family a *horoball packing.* Note that the horoballs in our packing do not have to touch one another. Although we are not yet assuming any invariance for our packing, the examples of interest are preserved by a discrete group of isometries. With this in mind, we call points of Q *cusp points* and points of L *limit points.*

We also suppose there is a positive constant D such that every point of every line $\overline{az}$ joining two limit points is distance at most D from $\cup\ B(q),\ q \in Q$. In this case we say our horoball packing is D*-dense.* We can rephrase this in terms of the unit tangent bundle SH^n, that is the sphere bundle of unit tangent vectors to H^n. Let $\mathcal{R} \subset SH^n$ consist of all unit vectors tangent to the rays $\overrightarrow{az}$ with a and z limit points. As L is closed in ∂H^n, $\mathcal{R}$ is closed in SH^n. Let $\pi : \mathcal{R} \to H^n$ be the basepoint projection, so $\pi(\mathcal{R})$ is closed in H^n. Then our packing is D-dense if and only if $\pi(\mathcal{R})$ lies in the D-neighborhood of $\cup\ B(q),\ q \in Q$.

In this paper, the interior Int S, the closure Cl S, and the boundary Bd S of a subset S of one of the spaces $L,\ L \times L$, or $\mathcal{R}$ are taken relative to that space.

The *geodesic flow* over H^n is the flow on SH^n whose flowlines are the lifts of the unit-speed geodesics on H^n. $\mathcal{R}$ is invariant by the geodesic flow and we let $\phi_t : \mathcal{R} \to \mathcal{R},\ t \in \mathbf{R}$, be the restriction of the geodesic flow to $\mathcal{R}$. Our goal is to analyze the flow ϕ_t in terms of how geodesics pass through or near the horoballs in the given packing, see §8 and §10.

Some of our uniqueness results will depend on L having no seams, where a *seam* of L is a subsphere $S \subset \partial H^n$ such that Int $S \cap L$ is nonempty but $L \not\subset S$. For instance, a 0-sphere is a seam of L if and only if it contains an isolated point of L. So if L has no seams then L has no isolated points and the converse holds for $n = 2$. Our examples below will have no seams by the following lemma.

Lemma 1.1. *The limit set $L = L(\Gamma)$ of a nonelementary discrete group Γ of isometries of H^n has no seams.*

We know that L is infinite and compact and has no isolated points ([Ra], Thm. 12.1.8). Suppose S is a subsphere of ∂H^n and Int $S \cap L$ is nonempty. We may choose S so that no proper subsphere of S has these properties. As L has no isolated points, S has dimension $d > 0$. For each $\gamma \in \Gamma$ we let $S_\gamma =$ Int $\gamma(S) \cap L$. As the action of Γ on L is minimal ([Ra], Thm. 12.1.4), the S_γ form an open cover of L. By our choice of S, the sets S_γ are disjoint or equal. As L is compact, there are only finitely many distinct S_γ. Thus a finite index normal subgroup $\Gamma' \subset \Gamma$ stabilizes each of these sets. But L is also the limit set of Γ' ([Ra], Thm. 12.1.16) and so Γ' acts minimally on L and there is only one S_γ. Hence S contains L and S is not a seam, proving the lemma. We note that this proof also shows that any real analytic set that does not contain L must meet L in a set with empty interior, c.f. Cor. A1.6.

We need to know the distance d in the upper halfspace model of H^n from a horoball B to a point w not in B. Suppose first that B is based at $p = \infty$ and so consists of all points in the upper halfspace whose height is at least some positive constant h_0. Suppose w has height $v < h_0$. Then $d = \log(h_0/v)$, by a trivial calculation. Suppose, however, that B is based at a finite point p, so $B \cup \{p\}$ is a Euclidean ball of some diameter $h = h(B)$. If v is the height of w then d is determined by

(1.2) $|w-p|^2 = hve^d$.

To prove this, let B' be the horoball based at p with $w \in \partial B'$. Then $h(B') = he^d$ and $h(B')/|w-p| = |w-p|/v$ by similar triangles, so the formula follows.

The following is a sort of ergodicity result that guarantees that many rays pass through our horoballs.

Proposition 1.3. *Let $\mathcal{E}$ consist of all distinct pairs $(z,a) \in L \times L$ such that some ray in $\overline{az}$ asymptotic to a does not meet the interior of any $B(q)$, $q \in Q$. Let $\mathcal{E}_z$, $z \in L$, consist of all $a \in L$ such that $(z,a) \in \mathcal{E}$. Then $\mathcal{E}$ (respectively $\mathcal{E}_z$) is meager in $L \times L$ (respectively L). The Lebesgue measure of $\mathcal{E} \subset \partial H^n \times \partial H^n$ is zero, as is the Lebesgue measure of $\mathcal{E}_z \subset \partial H^n$.*

To prove this, we fix a hyperplane $H \subset H^n$ and nonempty compact subsets Z, A of $L - \partial H$ that are separated by ∂H. We let H^+ be the open halfspace of H^n on the A side of H and we let $\Sigma \subset Z \times A$ consist of all pairs (z,a) such that $\overline{az} \cap H^+$ is disjoint from $\cup$ int $B(q)$, $q \in Q$. It is easily seen that $\mathcal{E}$ is covered by a countable family of such sets Σ for suitable choices of H, Z, and A. Fix $z \in Z$ and let Σ_z consist of all $a \in A$ such that $(z,a) \in \Sigma$. Then Q is disjoint from Σ_z, so the latter has no interior. Thus Σ has no interior. Moreover it is clear that Σ is closed and so Σ_z is closed, too.

The last statement in the proposition will hold if Σ_z has Lebesgue measure zero. By the Lebesgue density theorem, it suffices to fix an $a \in \Sigma_z$ and to prove that a is not a Lebesgue density point. Consider an upper halfspace model with $a = \infty$ so that H is a hemisphere and H^+ is its exterior. If $\overline{az}$ meets some $B(q)$ then this intersection lies outside H^+ and so $h(B(q))$ is bounded. Thus there is a horoball B based at a such that $B(q)$ is disjoint from B if $B(q)$ meets $\overline{az}$.

Now we pass to an upper halfspace model of H^n with $z = \infty$ and $a = 0$. Let w_m, $m \geq 0$, be a monotone divergent sequence in $\overline{0\infty}$ asymptotic to 0 and choose $p_m \in Q$ so that w_m is at distance $d_m \leq D$ from $B(p_m)$. For m large, $B(p_m)$ meets B and hence is disjoint from $\overline{0\infty}$. This implies $|p_m| \geq h_m/2$, where $h_m = h(B(p_m))$. We let v_m be the height of w_m, so (1.2) gives $v_m^2 + |p_m|^2 \leq h_m v_m e^D$. These two inequalities imply that all the ratios of h_m, v_m, and $|p_m|$ are bounded. Since $v_m \to 0$, we see that $p_m \to 0$ and $h_m \to 0$. Thus for m large, $B(p_m) \subset H^+$. Fix such an m.

Let $\pi_z : H^n \to \partial H^n$ denote the projection from $z \in \partial H^n$ along hyperbolic lines. Then $\pi_z(B(p_m))$ is a ball in ∂H^n of diameter h_m and center p_m that is disjoint from Σ_z. As this ball is contained in the ball of diameter $h_m + 2|p_m|$ centered at 0 and as the ratio of the diameters of these balls is bounded, the ratio of their volumes is also bounded. As $m \to \infty$, we see that 0 is not a Lebesgue density point of Σ_z and the proposition is proved.

For each $\rho > 0$ and each cusp point q let $B_\rho(q)$ be the horoball based at q consisting of those points of $B(q)$ of distance at least ρ from $\partial B(q)$. We let $\mathcal{O}_\rho$ be the union of the interiors of the horoballs $B_\rho(q)$. Now Proposition 1 applies to the horoball family $B_\rho(q)$ since this family is $(D+\rho)$-dense. Let Ω be the open subset of $L \times L$ consisting of distinct pairs (z,a). We say that (z,a) is *deep* if for each $\rho > 0$ the subset of the line $\overline{az}$ that belongs to $\mathcal{O}_\rho$ is unbounded in both directions. By applying Proposition 1 for a sequence of ρ tending to infinity, we find

(A) almost all pairs $(z,a) \in \Omega$ are deep.

At this point, "almost all" refers to Lebesgue measure. But (A) is trivial when L itself has Lebesgue measure zero. We suppose μ is a nonnegative Borel measure on ∂H^n with support in L such that (A) holds for the measure $\mu \times \mu$ on Ω. For instance the counting measure on Q satisfies (A). Lebesgue measure is a suitable μ when $L = \partial H^n$. For other natural examples, arising from geometrically finite groups, we refer to Thm. 11.2. As only null sets and sets of full measure matter to us, it is the measure class of μ that is significant.

Section 2. Lengths and shadows

We may use $B(p)$ to make the punctured boundary $\partial H^n - \{p\}$ a Euclidean space. The restriction of the hyperbolic metric to $\partial B(p)$ is flat, as is easily seen in an upper halfspace model with $p = \infty$. We may transfer this metric to a flat metric on $\partial H^n - \{p\}$ by the diffeomorphism $\pi_p : \partial B(p) \to \partial H^n - \{p\}$. We denote the length of a tangent vector v in this metric by $|v|_p$ and the distance function for this metric by $d_p(a, z)$. We also extend this distance function to the cases $a = p \neq z$ and $a \neq p = z$ by setting $d_p(a, z) = \infty$. Choose an upper halfspace model so that $B(p)$ is the *standard horoball* B_n consisting of all points of height at least one. Then $p = \infty$, $|v|_p$ is the Euclidean length, and $d_p(a, z) = |a - z|$ is the Euclidean distance function.

When p and q are distinct cusp points, their flat metrics are conformally equivalent. More precisely, the conformal factor at $z \in \partial H^n - \{p, q\}$ depends only on the orthogonal projection z' of z to the line $\overline{pq}$. We parametrize this line at unit speed by $t \in \mathbf{R}$ so that one approaches q as t approaches $+\infty$ and so that $t = 0$ corresponds to the midpoint of the shortest path from $B(p)$ to $B(q)$. Then for tangent vectors at z one has

Lemma 2.1. $|v|_p = e^{-2t(z')}|v|_q$.

To see this, let δ be half the distance from $B(p)$ to $B(q)$. Choose an upper halfspace model with $q = \infty$, $p = 0$, and e^t the height function on $\overline{0\infty}$. We may assume v has unit Euclidean length. Then $|v|_q = e^{-\delta}$ and $|v|_p = e^{-\delta}|z|^{-2}$. As $t(z') = \log |z|$, the lemma follows.

Now suppose $z \neq p$ lies closer to $B(q)$ than to $B(p)$ in the sense that some $w \in \overline{pz}$ lies at the same distance d from $B(p)$ as from $B(q)$. We now give another expression for our conformal factor, namely

$$e^{-2t(z')} = 1 - e^{2\delta - 2d}. \tag{2.2}$$

Inversion in the unit sphere interchanges p and q and fixes w so it interchanges $z = \pi_p(w)$ with $\pi_q(w)$. Using the coordinates of the preceding paragraph, $1 = |w|^2 = (e^{\delta - d})^2 + |\pi_q(w)|^2 = e^{2\delta - 2d} + 1/|z|^2$ and the formula follows. (2.2) shows that the conformal factor in Lemma 2.1 is less than one, expressing the fact that vectors at z look larger when projected to $\partial B(q)$ than when projected to $\partial B(p)$ since z is closer to $B(q)$ than to $B(p)$.

We now consider a codimension one sphere $\Sigma \subset \partial H^n$. For each cusp point q we let $\kappa(\Sigma, q)$ be the curvature (inverse radius) of the sphere or hyperplane $\Sigma - \{q\}$ in the Euclidean space $\partial H^n - \{q\}$ with the metric d_q. In the context of Lemma 2.1, we can compare the curvatures $\kappa(\Sigma, q)$ and $\kappa(\Sigma, p)$ as follows:

(2.3) if $z \in \Sigma$ with $t = t(z') > 0$ and $\kappa(\Sigma, p) \leq \kappa(\Sigma, q)$ then $\kappa(\Sigma, q) \leq e^{\delta}/\sinh(t)$.

This curvature estimate is needed to prove Thm. 13.1. It is a simplified version of Lemma 5 of [F3]. To check it, we again use the coordinate system in the proof of Lemma 2.1. We may assume that Σ does not pass through ∞. We choose a point z_0 on Σ nearest 0 and let z_1 be the point on Σ furthest from z_0, so $|z_0| \leq |z| = e^t$ and $1 < |z| \leq |z_1|$. We have $\kappa(\Sigma, q) = 2/d_q(z_1, z_0) = 2e^\delta/|z_1 - z_0|$ and $\kappa(\Sigma, p) = 2/d_p(z_1, z_0) = 2e^\delta|z_1||z_0|/|z_1 - z_0|$. Thus the curvature assumption implies that $|z_1|^{-1} \geq |z_0|$ and hence $|z_1 - z_0| \geq |z_1| - |z_0| \geq |z_1| - |z_1|^{-1} \geq |z| - |z|^{-1} = 2\ \sinh(t)$. (2.3) follows at once.

Now suppose z, w, d, and δ are as above with $z \neq q$. We consider an upper halfspace model of hyperbolic n-space in which z corresponds to ∞. Let $h(p) = h(B(p))$ and $h(q) = h(B(q))$. We find

Proposition 2.4. *$e^{2\delta}h(p)h(q) = |p-q|^2 = e^{2d}\ h(p)(h(q) - h(p))$, so $|p - q| = e^\delta\sqrt{h(p)h(q)}$, with $0 \leq \delta \leq d \leq D$. We have $|v|_p/h(p) = |v|_q/h(q)$ for all tangent vectors v to ∂H^n at ∞.*

We may increase the size of both horoballs by δ to reduce to the case of tangent horoballs. The first equation, with $\delta = 0$, then follows from the Pythagorean theorem applied to the right triangle whose base has constant height and whose hypotenuse joins the centers of our 2 horoballs. Next we apply (1.2) to w and $B(q)$. The height of w is $v = e^d h(p)$, since w lies over p at distance d from $B(p)$. Also $h = h(q)$ and $|w - q|^2 = v^2 + |p - q|^2$. We put these values into (1.2) and simplify to obtain the second equation. Finally we compare these two equations to get $h(p)/h(q) = 1 - e^{2\delta - 2d}$. As this equals $e^{-2t(z')}$, Lemma 2.1 gives the last equation in the proposition.

We now use the horoball packing $B(q)$, $q \in Q$, to divide hyperbolic space into polyhedra akin to Dirichlet domains. We let $F(q)$ consist of all points whose distance to $B(q)$ is no more than their distance to $B(p)$ for any $p \in Q$. Then $F(q)$ is a convex hyperbolic polyhedron, possibly infinite-sided, that contains $B(q)$. The sets $F(q)$, $q \in Q$, cover H^n without overlap. We call $F(q)$ a *fan* (c.f. [C]) and the family $F(q)$, $q \in Q$, the *fan tiling* associated to our horoball packing.

Lemma 2.5. *When $K \subset \pi(\mathcal{R})$ is a bounded set, the number of fans that meet K is bounded in terms of n and the diameter of K. The fan tiling is a locally finite closed cover of H^n.*

It suffices to bound the number of $B(p)$ that meet the D-neighborhood of K. K lies in a ball $B \subset H^n$ of radius $r = D + \mathrm{diam}(K)$, so it is enough to bound the number of $B(p)$ that meet B in terms of n and r. Let $B'(p) \subset B(p)$ be the subhoroball tangent to B and choose a disc model of H^n with origin at the center of B. We project each $B'(p)$ from the origin to ∂H^n to get a spherical ball whose volume depends only on n and r. Since the sets $B'(p)$ do not overlap, neither do their projections. Thus the number of p's is bounded, as desired. The proof of local finiteness is similar.

Suppose p and q are distinct cusp points. The intersection of $F(p)$ and $F(q)$ is either empty or a convex hyperbolic polyhedron of dimension at most $n-1$. When the intersection is nonempty and of dimension $n-1$, we say that p and q are *adjacent* and we define the *ridge* $R(q,p)$ to be this intersection of fans. This name is suggested by regarding the distance to $\cup\ B(q)$, $q \in Q$, as a height function on H^n.

We define the *shadow* of the ridge $R(q,p)$ as lit from p to be $S(q,p) = L \cap \pi_p(R(q,p)) \subset \partial H^n$. Note that the set of points at distance at most D from each of $B(p)$ and $B(q)$ is a compact set that meets each line $\overline{pz}$ for $z \in S(q,p)$. Hence $S(q,p)$ is compact. We let A denote the set of all adjacent pairs of cusp points (q,p) such that Int $S(q,p)$ is nonempty (as holds automatically when $L = \partial H^n$.)

Figure 1 shows some horoballs, fans, and ridges in the Poincare disc model of the hyperbolic plane.

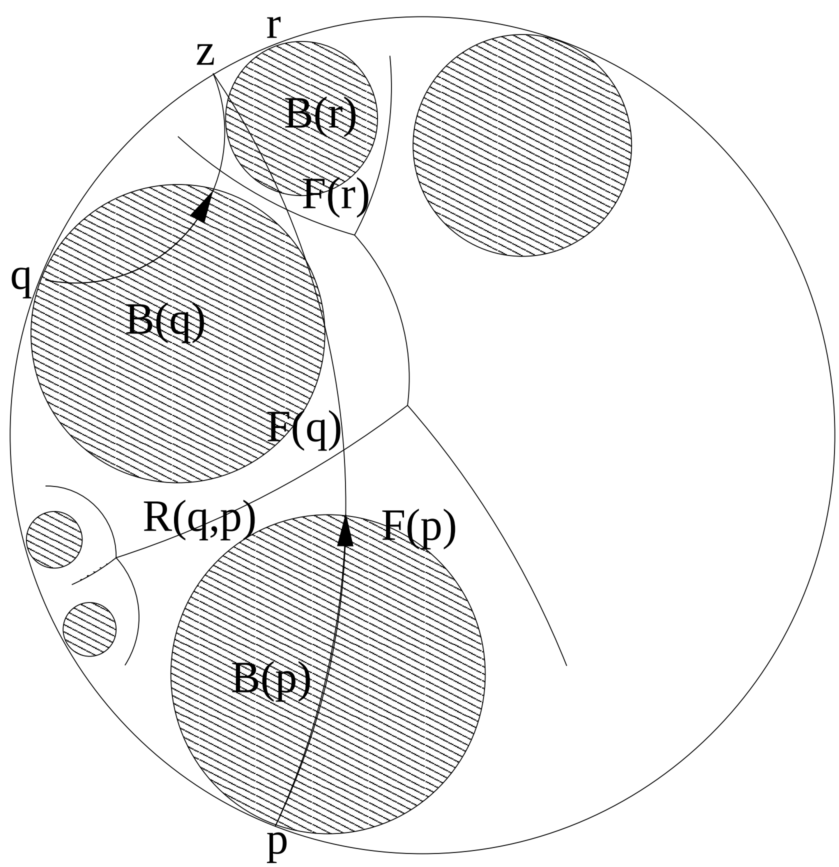

Figure 1. A point z in S(q,p) and S(r,q)

For $c = 1 - e^{-2D}$ and for all $z \in S(q,p)$, $z \neq q$, one has $|v|_p \leq c|v|_q$ for all tangent vectors v to ∂H^n at z. This follows from Lemma 2.1 and (2.2) since $1 - e^{2\delta - 2d} \leq 1 - e^{-2D}$. As $c < 1$, this length comparison expresses the fact that $S(q,p)$ is uniformly closer to q than to p. For the same value of c, we have $h(B(p)) \leq ch(B(q))$ in any upper halfspace model with $z = \infty$. This follows from the last equation in Prop. 2.4. This factor c will eventually explain the expanding properties of the mappings we use to describe geodesic flows.

We note that each set $\pi_p(R(q,p))$ is a Euclidean polyhedron, possibly infinite-sided, and that for a fixed p these polyhedra do not overlap. They form a locally finite tiling of $\partial H^n - \{p\}$ as q varies over all the cusp points adjacent to p. The corresponding shadows $S(q,p)$ form a locally finite cover of $L - \{p\}$ by compact sets. We say that q is a *z-successor* to p if $z \in S(q,p)$, $(q,p) \in A$.

Let p, q, and q' be distinct cusp points with q and q' adjacent to p. The ridges $R(q,p)$ and $R(q',p)$ lie on hyperplanes in H^n perpendicular to $\overline{qp}$ and $\overline{q'p}$, respectively. Thus $R(q,p) \cap$

$R(q',p)$ lies in a hyperplane whose boundary sphere contains p but does not contain q or q'. It follows that $S(q,p) \cap S(q',p)$ lies in a proper subsphere of ∂H^n that does not contain L. Therefore if L has no seams and $p \in Q$ is fixed, the shadows $S(q,p)$ have disjoint interiors.

We say that $z \in L$ is *generic* if it belongs to no shadow boundary Bd $S(q,p),\ (q,p) \in A$. The nongeneric points form a set of the first category in L. If L has no seams and $p \in Q$ is fixed then every generic $z \in L,\ z \neq p$, lies in a unique shadow $S(q,p),\ (q,p) \in A$, and is in the interior of this shadow.

We say that a family of sets is *essentially disjoint* if any two sets in the family intersect in a null set. The condition

(B) for each $p \in Q$, the family $S(q,p),\ (q,p) \in A$, is essentially disjoint

holds if every subsphere of ∂H^n that does not contain L is null. The latter is a measured version of the no seams property. (B) trivially holds if $L = \partial H^n$ and μ is Lebesgue measure. (B) also holds for the geometric measures of §11 attached to geometrically finite groups. (B) implies that almost every point of L is generic.

We now consider our motivating example, the *Ford disc* packing of the upper halfplane. Let the horoball $B(\infty) = B_2$ consist of all points with $\Im(w) \geq 1$. For each rational number $q = a/b$ in lowest terms, let $r = 1/(2b^2)$ and let $B(q)$ be the horoball given by $|w - (q + ir)| \leq r, w \neq q$. These Ford discs are permuted by $z+1,\ -\overline{z}$, and $-1/z$. These are generators of the discrete group $\Gamma = PGl(2, \mathbf{Z})$, so this group acts by permutations on the Ford discs. This permutation action is transitive. As the $B(q),\ q \neq \infty$, do not overlap $B(\infty)$, these discs are nonoverlapping. So we have a horoball packing of H^2 with $Q = \mathbf{Q} \cup \{\infty\},\ L = \partial H^2 = \mathbf{R} \cup \{\infty\}$. The complement of the union of the Ford discs consists of bounded open sets, which are also transitively permuted by Γ. Thus this horoball packing is D-dense for some D, and it is not hard to check that the smallest such D is $\log(2/\sqrt{3})$, which is the distance from $B(\infty)$ to $\exp(\pi i/3)$. We find $c = 1 - e^{-2D} = 1/4$.

In this case the cusp points adjacent to ∞ are the integers. An integer q is an x-successor of ∞ when $|x - q| \leq 1/2$, so the shadows $S(q, \infty)$ are just the level sets of the nearest integer relation. To go further, we recall the *least-remainder continued fractions* developed by Hurwitz [H], also called *continued fractions to the nearest integer.*

Suppose $x = x_0$ is irrational and let $a_0 = \langle x_0 \rangle$ be the integer nearest to x_0. We write $x_0 = a_0 - 1/x_1$ for some irrational number x_1. We write $x_1 = a_1 - 1/x_2$ with $a_1 = \langle x_1 \rangle$, and so on. The resulting formal expansion

$$x_0 = a_0 - 1/(a_1 - 1/(a_2 - ...)) =< a_0, a_1, a_2, ... > \tag{2.6}$$

has the following features. The a_i are integers, $|a_i| \geq 2$ for $i \geq 1$, and if $a_i = \pm 2$ for some $i \geq 1$ then $a_i a_{i+1} < 0$. Conversely, if integers $a_i,\ i \geq 0$, obey these conditions then Hurwitz showed that $< a_0, a_1, a_2, ... >$ is the least-remainder continued fraction of an irrational number x.

When we successively truncate the least-remainder expansion of x, we get a sequence p_m of *convergents* to x with $p_0 = \infty,\ p_1 = a_0,\ p_2 = a_0 - 1/a_1,\ p_3 = a_0 - 1/(a_1 - 1/a_2)$, etc. The convergents, as their name suggests, do converge to x [H]. Convergents are related to Ford discs as follows:

(2.7) for each $m \geq 0$, p_{m+1} is an x-successor to p_m.

(2.7) is the geometric basis of the theory of least-remainder continued fractions. Note that the least remainders $r_i = x_i - a_i$ lie in $[-1/2, 1/2]$ and are related by the expanding map E of (0.1), that is $r_{i+1} + \mathbf{Z} = -1/r_i + \mathbf{Z}$.

Section 3. Aimed sequences

While the shadows of §2 (which we now call *ridge shadows*) suffice for many applications, including the reduction theory of [F3], it is important to work more broadly (see Thm. 14.1, (19.1), and Prop. A2.3). Accordingly, we axiomatize the properties of ridge shadows proven in §2. Let a *shadow family* for the horoball packing $B(q),\ q \in Q$, be a family $S(q,p)$ of subsets of L, indexed by a set A of ordered pairs of distinct cusp points, obeying these conditions:

(3.1) For each $p \in Q$, $L - \{p\} = \cup S(q,p)$, where q varies over all cusp points with $(q,p) \in A$. Moreover this is a locally finite cover by nonempty compact sets with disjoint interiors.

(3.2) For some constant c with $0 < c < 1$ and all $(q,p) \in A$, $h(B(p)) \leq ch(B(q))$ in any halfspace model with $\infty \in S(q,p),\ q \neq \infty$.

(3.2*) For some constant c with $0 < c < 1$ and all $(q,p) \in A$, $|v|_p \leq c|v|_q$ for each vector v tangent to ∂H^n at a point of $S(q,p) - \{q\}$.

(3.3) For some constant $\Delta > 0$ and all $(q,p) \in A$, the distance between the horoballs $B(p)$ and $B(q)$ is no more than 2Δ.

For example, the ridge shadows of §2 form a shadow family with $\Delta = D$ and $c = 1 - e^{-2D}$, provided L has no seams. Note that (3.2) and (3.2*) are equivalent, by Prop. 2.4. Both mean that $S(q,p)$ lies uniformly closer to $B(q)$ than to $B(p)$. Although (3.3) does not imply that the horoball packing is Δ-dense, the horoball packing is D-dense for some $D > \Delta$ as shown in Prop. 7.2. If $z \in S(q,p)$ we say, as before, that q is a z-successor to p.

Given a shadow family, the next theorem describes each pair $(z,a) \in \Omega$ using a sequence of cusp points indexed by an interval of integers I. I is bounded below if $a \in Q$ and is bounded above if $z \in Q$ but otherwise $I = \mathbf{Z}$. For this, we need some definitions. If $f : I \to X$, with X some topological space, the *upper limit* $\lim^+ f(i)$ is $f(\sup\ I)$ when I is bounded above and $\lim_{i\to+\infty}\ f(i)$ when I is unbounded above and the limit exists. We define the *lower limit* $\lim^- f(i)$ in a similar fashion. Note that when l is an integer and $I' = I + l$, the shifted sequence $p'_i = p_{i-l},\ i \in I'$, has the same limits as $p_i,\ i \in I$.

Given $z \in L$, a sequence $p_i \in Q,\ i \in I$, is *aimed at* z if I contains at least two elements, $z \in S(p_{i+1}, p_i)$ for all $i \in I,\ i < \sup(I)$, and $\lim^+ p_i = z$. The latter condition is often redundant, as in Thm 3.4 below, but not always. For instance, if $q \notin S(q,p)$ (see §6) and $z \in S(q,p)$ then the sequence $p_0 = p,\ p_1 = q$ is not aimed at z.

A *chain* is a nonempty set of the form $S(p_l, ..., p_k) = S(p_l, p_{l-1}) \cap ... \cap S(p_{k+1}, p_k) \subset \partial H^n$, $l > k$, with $(p_{i+1}, p_i) \in A$ for $k \leq i < l$. When $p_i,\ i \in I$, is aimed at z then each finite interval in I with at least two elements defines a chain containing z. For $z \in \partial H^n - \{p\}$, $B(z,p)$ denotes the horoball based at z that touches $B(p)$ at one point.

We now associate aimed sequences to geodesics with endpoints in L.

Theorem 3.4. *Let $S(q,p),\ (q,p) \in A$, be a shadow family for the packing $B(q),\ q \in Q$.*
Monotonicity: *When $p_i,\ i \in I$, is aimed at z, the $|\log c|$-neighborhood of $B(z,p_i)$ is contained in $B(z,p_{i+1})$ for $i \in I,\ i+1 < \sup(I)$.*
Convergence: *If $p_i,\ i \in I$, is aimed at z then the p_i are distinct, $\lim^- p_i = a$ exists, and*

$a \neq z$. I is bounded above if and only if $z \in Q$ and I is bounded below if and only if $a \in Q$. If I is not bounded above and $z \in \cap\, S(p_{i+1}, p_i)$, $i \in I$, then p_i, $i \in I$, is aimed at z.
Continuation: *If $z \in S(p_l, ..., p_k)$, $k < l$, then $p_k, ..., p_l$ may be extended to a sequence p_i, $i \in I$, aimed at z with $\inf(I) = k$.*
Existence: *Given distinct limit points z, a there is an aimed sequence with these limits.*
Uniqueness: *This aimed sequence is unique up to a shift except for (z, a) in a meager subset of Ω. This meager set is null for any measure μ satisfying (A) and (B).*

So each geodesic $\overrightarrow{az}$ with endpoints in L is described by an aimed sequence with limits (z, a) and this sequence is uniquely determined, up to a shift of its indices, for most geodesics. Next we compare such sequences for a fixed z.

Theorem 3.5. *Let μ be a measure satisfying (A) and (B). For almost all z, when two sequences p_i, $i \in I$, and q_j, $j \in J$, are aimed at z there is an $l \in \mathbf{Z}$ and $j_0 \in J$ so that $q_j = p_{j+l}$ for all $j \geq j_0$. The set of exceptions is meager in L.*

Taking a ridge shadow family in Theorems 3.4 and 3.5, with $L = \partial H^n$ and Lebesgue measure for μ, proves Theorems 6 and 7 in [F3]. In this case it is clear that L has no seams and (B) holds. (A) was shown in §1.

To begin the proof of Thm. 3.4, suppose $i \in I$, $i + 1 < \sup(I)$. Choose an upper halfspace model with $z = \infty$. Monotonicity now follows from (3.2).

Now consider Convergence. The p_i are distinct by Monotonicity. Suppose $z = \infty \in S(p_{i+1}, p_i)$ for each $i \in I$, $i < \sup(I)$, and let $h_i = h(B(p_i))$. For $i > j$, (3.2) gives $h_i \geq c^{j-i} h_j$. When $z \in Q$, the h_i are bounded and so I must be bounded above. Since $B(p_i)$ and $B(p_j)$ are disjoint, $|p_i - p_j| \geq \sqrt{h_i h_j} \geq c^{(j-i)/2} h_j$ so $d_{p_j}(z, p_i) \leq c^{(i-j)/2}$. When I is not bounded above, we fix j and find $\lim^+ p_i = z$, so p_i is aimed at z. Indeed the p_i converge exponentially as i increases, with error $O((\sqrt{c})^i)$.

(3.3) and Prop. 2.4 show $|p_m - p_{m-1}| \leq e^{\Delta} \sqrt{h_m h_{m-1}}$ for $p_m \neq \infty$. Taking i, $j \in I$ with $j < i < \sup(I)$ this gives $|p_i - p_j| \leq \Sigma\ |p_m - p_{m-1}| \leq e^{\Delta}\ \Sigma\ \sqrt{h_m h_{m-1}}$, where m varies from $j + 1$ to i. Since $h_m \leq c^{i-m} h_i$, this gives $|p_i - p_j| \leq b h_i$ where $b = \sqrt{c} e^{\Delta}/(1 - c)$ is a positive constant. When I is not bounded below, this shows that the sequence p_j, $j \to -\infty$, is Cauchy, so $\lim^- p_j = a$ exists. Indeed

(3.6) $|p_i - a| \leq b h_i$

so the convergence is exponential as i decreases, with error $O(c^{-i})$. Note that $a \neq \infty = z$. All parts of Convergence have been shown except the assertion for $a \in Q$.

For example, consider the Ford disc packing of H^2 with the ridge shadow family, where $c = 1/4$ and $\Delta = 0$. Suppose p_i, $i \in \mathbf{Z}$, is an aimed sequence with limits (z, a) and with $p_0 = \infty$. As in §2, p_1, $p_2, ...$ is the sequence of convergents in a least-remainder continued fraction expansion of z. Here $\sqrt{c} = .5$, so $|p_i - z| = O((.5)^i)$, $i > 0$. By [H, pp. 93-97], however, p_i approximates z like the inverse square of its denominator and these denominators grow by a factor of at least $(1 + \sqrt{5})/2$, so the error is $O(r^i)$ for $r = (3 - \sqrt{5})/2 < .4$. Thus the rate of convergence that follows from the D-density of the Ford disc packing is not optimal. On the other hand, p_{-1}, $p_{-2}, ...$ is the sequence of convergents in an "H-dual" continued fraction expansion of a, introduced by Hurwitz [H]. Then (3.6) shows $|p_{-i} - a| = O((.25)^i)$, $i > 0$, for a finite. This is consistent with [H, pp. 95-97], where it is shown that the rational number

p_{-i} approximates a like the inverse square of its denominator and that the denominators at least double at each step.

Continuation is proven by induction on l for a fixed k. When $p_l \neq z$, (3.1) implies $z \in S(q, p_l)$ for some q, so one may continue the aimed sequence with $p_{l+1} = q$. Continuing in this way, either $p_m = z$ for some $m \geq l$, in which case $m = \sup(I)$, or p_m is defined for all $m \geq k$ and $\lim^+ p_i = z$ by Convergence. Either way, Continuation is proven.

We now relate the limit pair of an aimed sequence to certain pairs of cusp points.

Lemma 3.7. *Suppose p_i, $i \in I$, is an aimed sequence with limit pair (z, a) and $m \in I$, $m < \sup(I)$. Suppose $q \in Q$, $q \neq p_m$, $q \neq z$, and $B(z,q) \subset B(z, p_m)$. Then*

$$\frac{|z-q||p_m - a|}{|z-a||q-p_m|} \leq b \text{ and } \frac{1}{\alpha} \leq \frac{|z-p_m||q-a|}{|z-a||q-p_m|} \leq b+1$$

for a positive constant α depending only on c and Δ.

For example, Monotonicity permits one to take $q = p_{m+1}$ here in order to bound the relative position of the points z, p_{m+1}, p_m, and a. Whereas certain coincidences are possible between these points, namely one can have $p_m = a$ or $p_{m+1} = z$ or both, the other coincidences are not, that is $p_m \neq \; p_{m+1} \neq a \neq z \neq p_m$. The lemma gives a quantitative form of these forbidden coincidences.

Choose an upper halfspace model with $z = \infty$ and $q = 0$, and assume $h(B(q)) = 1$. Let $h_i = h(B(p_i))$, $i \in I$, $i < \sup(I)$. By assumption we have $h_m \leq 1$. Then Prop. 2.4 gives $|p_m| \geq \sqrt{h_m}$ and (3.6) gives $|p_m - a| \leq b h_m$. Thus $|p_m - a|/|p_m| \leq b$ and so $|a|/|p_m| \leq b + 1$. This proves the upper bounds in the lemma.

For the lower bound, the case $p_m = a$ is trivial so suppose $m > \inf(I)$. By Prop. 2.4, $|p_i| \geq \sqrt{h_i}$ for $i \in I, i < m$. (3.3) and Prop. 2.4 give $|p_i - p_{i+1}| \leq e^{\Delta}\sqrt{h_i h_{i+1}}$. So

$$\log\frac{|p_{i+1}|}{|p_i|} \leq \frac{|p_{i+1} - p_i|}{|p_i|} \leq e^{\Delta}\sqrt{h_{i+1}} \leq e^{\Delta}\sqrt{c^{m-i-1}}$$

where the first inequality uses $\log x \leq x - 1$. Now sum over all $i \in I$, $i < m$, and simplify to get $|p_m|/|a| \leq \alpha$, where $\alpha = \exp[e^{\Delta}/(1-\sqrt{c})]$. This proves the lemma.

When $a \in Q$ and $a \neq p_m$, the lower bound in Lemma 3.7 shows $B(z,a) \not\subset B(z, p_m)$. Thus the sequence h_m is bounded below and, by (3.2), I is bounded below. This finishes the proof of Convergence.

Section 4. Existence and Uniqueness

Now consider Existence. This follows from Continuation when $a \in Q$. So fix $a \in L - Q$, $a \neq z$, and write $a = \lim a_j$, $j = 1, 2, 3, \ldots$, where the a_j are distinct cusp points with $a_j \neq z$. For each j, Continuation gives a sequence p_{jk}, $k \in I_j$, aimed at z with I_j bounded below and $\lim^- p_{jk} = a_j$. We will construct a sequence p_i aimed at z with $\lim^- p_i = a$ by a diagonal process. We first treat the special case when $0 \in I_j$ for all j and $p_{j0} = p_0$ is independent of j.

Choose an upper halfspace model with $z = \infty$ and $a = 0 \notin Q$ and let $h_{jk} = h(B(p_{jk}))$ for $p_{jk} \neq \infty$. Assume for now that ∞ is not a cusp point, so I_j is not bounded above. Let w_j

be the point where $\overline{p_0 p_{j1}}$ meets $\partial B(p_{j1})$. Since $\infty \in S(p_{j1}, p_0)$, (3.2) implies that the height of w_j is at least $h(B(p_0))/2$. (3.3) implies that w_j is in the $2D$-neighborhood of $B(p_0)$. Thus w_j lies in a compact set in H^n. By Lemma 2.5, this compact set meets only a finite number of horoballs in our packing. It follows that there are infinitely many j with the same value of p_{j1}, which we denote p_1. As $0 \notin Q$, $p_0 \neq 0$. For j large enough, $|a_j| < |p_0|/3$ and so

$$2|p_0|/3 < |p_0| - |a_j| \leq |p_0 - a_j| \leq |p_0 - p_{j,-1}| + |p_{j,-1} - a_j| \leq |p_{j,-1} - p_0| + bh_{j,-1}.$$

Thus either $|p_{j,-1} - p_0| > |p_0|/3$ or $h_{j,-1} > |p_0|/(3b)$. Of the horoballs $B(q)$ of distance at most 2Δ from $B(p_0)$, all but finitely many have q very near to p_0 and very small diameter. Thus $p_{j,-1}$ can only take finitely many values. So within the set of j for which $p_{j1} = p_1$, there is an infinite set with $p_{j,-1}$ taking the same value, say p_{-1}. Continue in this way, taking $k = 0, 1, -1, 2, -2, ...$, to construct the desired aimed sequence.

When ∞ is a cusp point, one can suppose $p_{j0} = \infty$ for all j. Repeating the argument just given for negative k gives a suitable aimed sequence. Thus Existence is proved in the special case.

Returning to the general case, one may suppose $z = \infty$ and $a = 0$ are not cusp points. By deleting a finite number of j and shifting the index sets I_j as needed, one may suppose that $h_{j,-1} < 1 \leq h_{j0}$ for all j. By (3.6), $|a_j - p_{j,-1}| < b$ so $p_{j,-1}$ is bounded. By (3.3), $B(p_{j0})$ passes through the horoball based at $p_{j,-1}$ of diameter $e^{2\Delta}$, so $B(p_{j0})$ meets a bounded Euclidean ball centered at 0. But there are only finitely many $B(q)$ of diameter at least one that meet such a ball. So only finitely many cusp points can occur as p_{j0}. We now pass to a subsequence of the j's to reduce to the special case. This concludes the proof of Existence.

The proof of Uniqueness uses an additional lemma.

Lemma 4.1. *Suppose p_i, $i \in I$, is an aimed sequence with limit pair (z, a) and q is a cusp point distinct from all the p_i. For some positive ρ, depending only on c and Δ, $B_\rho(q) \cap \overline{az}$ is empty. Equivalently, if $q \in Q$ and $(z, a) \in \Omega$ with $d_q(z, a) \geq 2e^\rho$ then every aimed sequence with limit pair (z, a) contains q. We may take ρ with $2e^\rho = \alpha\beta$ where $\beta = 1 + e^\Delta$.*

By Convergence, $z \neq q$. We may assume $\overline{az}$ meets $B(q)$. Using an upper halfspace model with $z = \infty$, one finds an $m \in I$, $m < \sup(I)$, such that $B(z, q) \subset B(z, p_m)$ but either $p_{m+1} = z$ or $B(z, q)$ overlaps $B(p_{m+1})$. Using an upper halfspace model with $B(q) = B_n$, where $h(B(z, q)) = 1 \geq h(B(p_{m+1}))$, one finds that $d_q(z, p_{m+1}) < 1$. By Lemma 3.7, $d_q(z, a) \leq \alpha d_q(z, p_m) < \alpha(1 + d_q(p_{m+1}, p_m))$. By Prop. 2.4, $d_q(p_{m+1}, p_m) \leq e^\Delta$, so $d_q(z, a) < \alpha\beta$ and the lemma follows.

Now consider Uniqueness. As in §2, we say z is generic if it does not lie in the boundary of a shadow. For a generic z, (3.1) implies that each $p_i \neq z$ has a unique z-successor and so p_i determines p_j for $j > i$. But Lemma 4.1 provides certain cusp points that must occur in aimed sequences. In particular, if z is generic and (z, a) is deep then (z, a) is represented by a unique aimed sequence, up to a shift. Such pairs form a Borel set. By (A) and (B), this set has full measure. Proposition 1.3 shows that the set of pairs (z, a) that are not deep is meager in Ω. As the set of nongeneric points is meager in L, we see that the aimed sequence is unique, up to a shift, outside a meager null set of Ω. This proves Uniqueness and also Theorem 3.4.

The proof of Theorem 3.5 is similar to that of Uniqueness. The set of $z \in L$ such that z is generic and (z, a) is deep for some $a \in L,\ a \neq z$, has full measure by (A) and (B). We fix such a pair (z, a) and choose an upper halfspace model with $z = \infty,\ a = 0$. If $a' \in L,\ a' \neq \infty$, then one can find a cusp point q such that $\overline{a'\infty}$ and $\overline{0\infty}$ meet $B_\rho(q)$. If $\infty \in Q$ we simply take $q = \infty$. Otherwise there are $q \in Q$ such that $\overline{0\infty}$ meets $B_{1+\rho}(q)$ with $h(B(q))$ arbitrarily large, so one may choose such a q with $|q| > |a'|$. Then $\overline{a'\infty}$ penetrates $B(q)$ to a distance

$$\log \frac{h(B(q))}{2|a'-q|} = \log \frac{h(B(q))}{2|q|} + \log \frac{|q|}{|a'-q|} > 1 + \rho - \log 2 > \rho.$$

So Lemma 4.1 shows that q is in any aimed sequence with limits (∞, a') or with limits (∞, a). As ∞ is generic, such aimed sequences will have the same tail from q on, and the theorem is proved.

Henceforth we will assume that μ satisfies (A) and (B) so that Uniqueness and Thm. 3.5 hold.

Section 5. Some sets in orbit space

A system of flowboxes for the geodesic flow ϕ_t on $\mathcal{R}$ will be constructed in §7. As preparation, we analyze certain subsets of the orbit space of ϕ_t. We identify this space of flowlines with the space Ω of ordered pairs of distinct points in L by mapping the flowline $\{\phi_t(v) : t \in \mathbf{R}\}$ to $(z, a) \in \Omega$ when $\pi(\phi_t(v))$ is asymptotic to z as $t \to +\infty$ and to a as $t \to -\infty$.

For each cusp point q, consider all aimed sequences $p_i,\ i \in I$, such that $0 \in I$ and $p_0 = q$, and form the set $K(q) \subset \Omega$ of all limit pairs $(z, a) = (\lim^+ p_i, \lim^- p_i)$. One can roughly describe $K(q)$ as follows.

Proposition 5.1. *$K(q)$ is compact, Bd $K(q)$ is null, and*

$$\{(z, a) : d_q(z, a) \geq \alpha\beta\} \subset K(q) \subset \{(z, a) : d_q(z, a) \geq 1/b\}.$$

For $(r, q) \in A$, let $K(r, q)$ consist of all limit pairs (z, a) of aimed sequences with $\{0, 1\} \subset I,\ p_0 = q$, and $p_1 = r$. Note $z \in S(r, q)$. Define the *slice* $K_z(r, q)$ for each fixed $z \in S(r, q)$ as the set of all $a \in L$ such that $(z, a) \in K(r, q)$. The sets $K_z(r, q)$ and $K(r, q)$ admit the following description.

Proposition 5.2. *$K(r, q)$ is compact, Bd $K(r, q)$ is null, and for each $z \in S(r, q)$,*

$$\{a : d_q(z, a) \geq \alpha\beta\} \subset K_z(r, q) \subset \{a : d_q(z, a) \geq 1/b\}.$$

There is a set of cusp points in Int $K_z(r, q)$ that is dense in this slice. For a fixed q, the sets $K(r, q)$ form a locally finite cover of $\{(z, a) \in K(q) : z \neq q\}$ by essentially disjoint compact sets with nonempty disjoint interiors. For a fixed q, the sets $K(q, p)$ form a locally finite cover of $\{(z, a) \in K(q) : a \neq q\}$ by essentially disjoint compact sets with nonempty disjoint interiors.

Consider the first inclusion in Prop. 5.2. If $q \neq z$ and $d_q(z, a) \geq \alpha\beta$ then we take an aimed sequence $p_i,\ i \in I$, with limit pair (z, a) and we apply Lemma 4.1 to conclude that

$q = p_j$ for some $j \in I$. Shifting indices, we may suppose that $p_0 = q$. Using Continuation we can build a sequence $p'_0 = q,\ p'_1 = r, \ldots$ aimed at z. We splice these two sequences to build a sequence aimed at z with terms $p_i,\ i \leq 0$, and $p'_i,\ i \geq 0$. This sequence has limits (z, a) so $(z, a) \in K(r, q)$, as desired. To prove the first inclusion in Prop. 5.1, one reasons similarly when $a \neq q \neq z$. When $(z, a) \in \Omega$ with $z = q$ or $a = q$, one instead uses Existence and Convergence to show $(z, a) \in K(q)$.

The second inclusion in either proposition follows from (3.6). Using these inclusions, we see that compactness will hold if $K(q)$ and $K(r, q)$ are closed in Ω. But this follows from the proof of Existence, which shows that each p_i can take on only finitely many values.

The local finiteness of the family $K(r, q)$ follows from (3.1). (3.1) and our inclusions show that the interiors are nonempty and disjoint. (B) shows that this family is essentially disjoint. Note that if $(z, a) \in K(q)$ with $z \neq q$ then any aimed sequence with limits (z, a) and $p_0 = q$ must have a term p_1 and so $(z, a) \in K(p_1, q)$. This shows that the $K(r, q)$ cover $\{(z, a) \in K(q) : z \neq q\}$.

Similar reasoning shows that the $K(q, p)$ cover $\{(z, a) \in K(q) : a \neq q\}$. Each set $K(q, p) \cap K(q, p'),\ p' \neq p$, is null and meager by Uniqueness. By the Baire category theorem, Int $K(q, p)$ and Int $K(q, p')$ are disjoint. Say $(z_0, a_0) \in K(q),\ a_0 \neq q$, and $(z, a) \in K(q, p)$ is near (z_0, a_0). By the second inclusion in Prop. 5.2, $1/b \leq d_p(z, a)$. Using an upper halfspace model with $B(q) = B_n$, one sees $d_p(z, a) = h(B(p))|z - a|/(|z - p||a - p|)$ and $h(B(p)) \leq 1$. By the lower bound in Lemma 3.7, $|z - a|/|z - p| \leq \alpha$. Altogether, these inequalities give $|a - p| \leq b\alpha$. Since a is near a_0, p lies in a bounded subset of $L - \{q\}$. Since the 2Δ-neighborhood of $B(q)$ must meet $B(p)$, there are only finitely many possibilities for p. This shows that the family $K(q, p)$ is locally finite.

Say $p_i,\ i \in I$, is aimed at z with $p_0 = q$ and $p_1 = r$. We choose $m \in I,\ m \leq 0$, and let $p = p_m \in K_z(r, q)$. Points such as p are surely dense in this slice. To show that p is in the interior of this slice, it suffices to prove $a \in K_z(r, q)$ when $d_p(z, a) > \alpha\beta$. By Lemma 4.1, $(z, a) \in K(p)$. So there is an aimed sequence $p'_i,\ i \in I'$, with limits (z, a) such that $m \in I'$ and $p'_m = p$. Let p''_i be the spliced sequence with $p''_i = p_i,\ i \geq m$, and $p''_i = p'_i,\ i \leq m$. Then p''_i is an aimed sequence with limits (z, a), $p''_0 = q$, and $p''_1 = r$. So $a \in K_z(r, q)$, as desired.

The above propositions are proved, except for the assertions that the topological boundaries are null. These are proven in §7 using Uniqueness and flowboxes for the geodesic flow.

In many cases of interest, the slices $K_z(q, p)$ take only finitely many values as z varies over $S(q, p)$, with consequences that we study in §9 and §10. Consider for instance the Ford disc packing and let $\Gamma = PSl(2, \mathbf{Z})$. The action of Γ on the set A of adjacent pairs of cusp points is simply transitive so all the $K(q, p)$ are equivalent under Γ to $K(\infty, 0)$. As shown in §12 or [F3] (but see [H], Fig. 4)

$$K(\infty, 0) = [2, \infty] \times [r - 1, r] \cup\ [\infty, -2] \times [-r, 1 - r] \tag{5.3}$$

where $r = (3 - \sqrt{5})/2$. Thus $K_z(\infty, 0)$ is $[r - 1, r]$ for $z \geq 2$, $[-r, 1 - r]$ for $z \leq -2$, or $[r - 1, 1 - r]$ for $z = \infty$. These slices are intervals as required by Lemma 6.6.

The translates of $K(\infty, 0)$ by integers n give nonoverlapping rectangles on which $|z - a| \geq 2 - r,\ a \neq \infty$. Their union contains all finite pairs (z, a) with $|z - a| \geq 3 - r$. These bounds are sharp. They are consistent with the inclusions of Prop. 5.1 since for $c = 1/4$ and $\Delta = 0$ one has $1/b = 3/2 < 2 - r$ and $\alpha\beta = 2e^2 > 3 - r$.

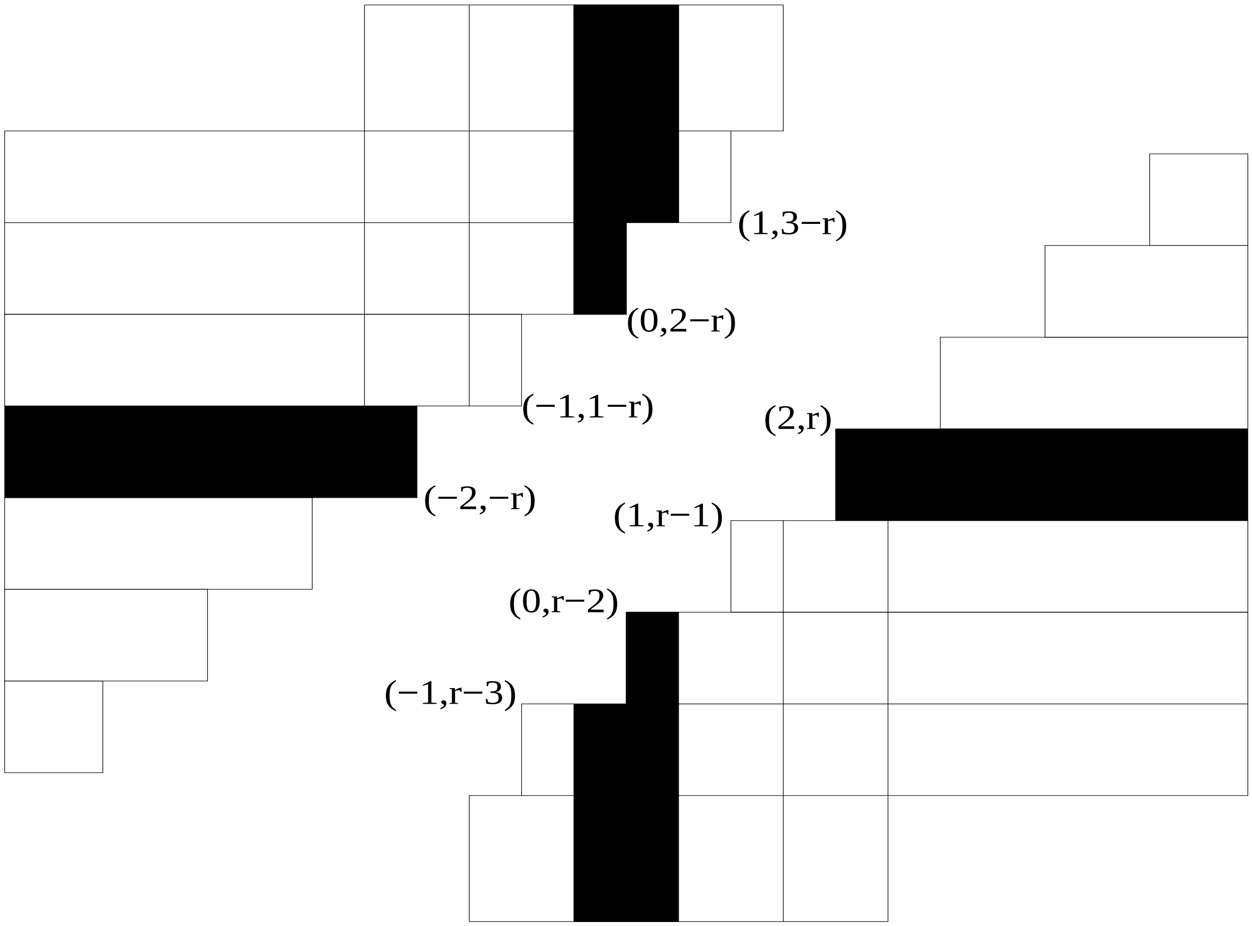

Figure 2. Some sets K(q,p) for the Ford disc packing

If we apply $-1/z$ to $K(\infty, 0)$ we obtain the two rectangles $\pm([0, 1/2] \times [3 - r, r - 2])$, where we have used the identities $r(3 - r) = 1 = (1 - r)(2 - r)$. Here the second factor is an interval containing ∞. Taking the translates by integers of these two rectangles gives the same union as before, except that the points with $z = \infty$ are deleted and replaced by points with $a = \infty$. This is consistent with the covering assertions of Prop. 5.2. The sets $K(\infty, 0)$ and $K(0, \infty)$ are illustrated in Figure 2, together with several of their integer translates. Note that the portion $K^>$ of $K(\infty) \cap \mathbf{R}^2$ with $a > z$ is the region above the graph of a monotone function and the portion $K^<$ of $K(\infty) \cap \mathbf{R}^2$ with $a < z$ is the region below the graph of a monotone function. This will be explained in the next section. For a reason why these monotone functions are step functions we refer to Cor. 9.1. We note that similar pictures arise in the work of Adler and Flatto on symbolic dynamics for a compact hyperbolic surface ([Ad], Fig. 24).

The following lemma relates chains and slices. Fix $(q, p) \in A$. For $z \in S(q, p)$ we let $S(z)$ be the intersection of all chains $S(q, p, ...)$ containing z.

Lemma 5.4. *If $z' \in S(z)$ then $K_z(q,p) \subset K_{z'}(q,p)$ and $S(z') \subset S(z)$.*

The second inclusion is obvious. To see the first, we consider an aimed sequence $p_i,\ i \in I$, with $(p_1, p_0) = (q,p)$ and with limit pair $(z,a) \in K(q,p)$. Let $q_j,\ j \geq 0$, be an aimed sequence with $(q_1, q_0) = (q,p)$ and $\lim^+ q_j = z' \in S(z)$. Then z and z' lie in each chain $S(p_1, p_0, ..., p_m),\ m \leq 0$, and so the spliced sequence p'_i with $p'_i = p_i,\ i \leq 1,\ i \in I$ and $p'_i = q_i,\ i \geq 0,\ i \in J$, is aimed at z'. The limit pair of this spliced sequence is (z', a), so the lemma is proven.

Section 6. The picture for $n = 2$

In this section we will examine the behavior of aimed sequences, chains, and the sets $K(q)$ for a ridge shadow family, when $n = 2$ and L has no seams (that is, no isolated points). One of our goals is to explain why the regions in Figure 2 are bounded by the graphs of monotone functions. The reasoning is based on three topological properties of ridge shadow families and our results will apply to families with these properties.

We fix a horoball packing $B(q) \subset H^2,\ q \in Q$, such that L has no seams, and construct the fans $F(q),\ q \in Q$, and the ridge shadows $S(q,p),\ (q,p) \in A$, as in §2. We say that a set of limit points in an *interval in L* if it is the intersection of L with a closed interval in ∂H^2. We will need these properties of the ridge shadows.

(6.1) Each $S(r,q)$ is an interval in L and for a fixed q these shadows occur in same order on $\partial H^2 - \{q\}$ as the cusp points r.

(6.2) Pairs (q,p) and (q',p') in A are unlinked in ∂H^2 whenever the cusp points q, p, q', and p' are distinct.

(6.3) If $(q,p) \in A,\ (r,q) \in A$, and $S(q,p)$ meets $S(r,q)$ then $S(q,p) \cup S(r,q) \neq L$.

(6.1) is clear since $S(r,q) = L \cap \pi_q(R(r,q))$. (6.2) holds since the fans $F(p)$ and $F(q)$ share an edge and cannot separate $F(p')$ from $F(q')$. To check (6.3) we assume $S(q,p) \cup S(r,q) = L$. Then $q \in S(q,p)$ and $p \in S(r,q)$ and so $\overline{pq}$ meets $R(q,p) \cap R(r,q)$. So unless $p = r$, these two ridges lie on different sides of $\overline{pq}$. Either way, $S(q,p)$ and $S(r,q)$ are disjoint and so (6.3) holds.

For a horoball packing $B(q) \subset H^2,\ q \in Q$, a shadow family that satisfies (6.1), (6.2), and (6.3) will be called *standard*. For instance, certain perturbations of a ridge shadow family are standard, as shown in §14. We now fix a standard shadow family and extend (6.2). Given an aimed sequence $p_i,\ i \in I$, we say that a subsequence defined by a finite nonempty interval in I is an *aimed segment*. If this segment contains N points then its complement in ∂H^2 is the union of N component intervals.

(6.2*) If an aimed sequence lies in the complement of an aimed segment then it lies in one component of this complement.

The case $N = 1$ is trivial. The case $N = 2$ holds by applying (6.2) to each successive pair of terms in the aimed sequence. The cases with $N > 2$ follow from the two-point case.

We now show that the p_i converge monotonically on each side of z.

Lemma 6.4. *For any three indices $j < k < l$ in I, p_l lies in the component of $\partial H^2 - \{p_j, p_k\}$ that contains z.*

For motivation, consider the Ford disc packing and suppose $j = 0,\ p_0 = \infty$. This lemma says that the convergents in a least-remainder continued fraction expansion of z that are greater than z are decreasing whereas the convergents that are less than z are increasing. It may be that all the convergents lie on one side of z. By comparison, the convergents in a simple continued fraction alternate from one side of z to the other and are monotone on each side of z.

The lemma is trivial if $l = \sup(I)$ since then $z = p_l$. Otherwise, use (6.2*) with the aimed segment $p_j, ..., p_k$ and the aimed sequence $p_i, i \in I, i \geq l$, and the lemma follows.

Next we prove that every chain is an interval in L. Fix a family of compact intervals $\mathcal{I}(q,p),\ (q,p) \in A$, such that $S(q,p) = L \cap \mathcal{I}(q,p)$. When $S(p_{m+1}, ..., p_k)$ is a chain, let $\mathcal{I}(p_{m+1}, ..., p_k) = \mathcal{I}(p_{m+1}, p_m) \cap ... \cap \mathcal{I}(p_{k+1}, p_k)$, so $S(p_{m+1}, ..., p_k) = \mathcal{I}(p_{m+1}, ..., p_k) \cap L$.

Lemma 6.5. *$\mathcal{I}(p_{m+1}, ..., p_k)$ is an interval.*

We prove this by induction on $m-k$. The case $m = k$ is trivial. If $m > k$ and $\mathcal{I}(p_m, ..., p_k)$ is an interval then $\mathcal{I}(p_{m+1}, ..., p_k) = \mathcal{I}(p_{m+1}, p_m) \cap \mathcal{I}(p_m, ..., p_k)$ is an interval as well since $\mathcal{I}(p_{m+1}, p_m) \cup \mathcal{I}(p_m, ..., p_k) \neq \partial H^2$. In fact this union lies in $\mathcal{I}(p_{m+1}, p_m) \cup \mathcal{I}(p_m, p_{m-1})$ which, by (6.3), does not contain L.

Lemma 6.5 shows that chains are intervals in L. We now treat slices.

Lemma 6.6. *For each $z \in S(q,p)$, the slice $K_z(q,p)$ is an interval in L containing p in its interior.*

Prop. 5.2 shows that p is in the interior of this slice. Say $a \in K_z(q,p)$ with $a \neq p$, and $p_i,\ i \in I$, is an aimed sequence with limit pair (z,a) such that $p_0 = p,\ p_1 = q$. Suppose r is a cusp point and the boundary points z, a, r, p are distinct and in cyclic order on the circle ∂H^2. For some $h \in I, h < 0$, the points z, p_h, r, p_0 are also distinct and in cyclic order. Then by (6.2*), any sequence $p'_j,\ j \in J$, aimed at z that starts at r must include some $p_i,\ h \leq i \leq 0$. But if $p'_j = p_i$ one may shift indices in J so that $j = i$. Consider the spliced sequence q_k with $q_k = p'_k,\ k \leq i$, and $q_k = p_k,\ k \geq i$. Here the index set is $K = \{k :\ k \leq i,\ k \in J\} \cup \{k :\ k \geq i,\ k \in I\}$. Clearly q_k is aimed at z and $\lim^- q_k = \lim^- p'_j = r$. Also $q_0 = p$ and $q_1 = q$, so $r \in K_z(q,p)$. Taking limits of such cusp points r, this slice must contain all points of L in an interval from a to p. The lemma follows.

For each cusp point q and each limit point $a \neq q$, let the *level* $K^a(q)$ be the set of all z for which $(z,a) \in K(q)$.

Lemma 6.7. *Each level $K^a(q)$ is an interval in L containing q in its interior.*

Prop. 5.1 implies that q is in the interior of this level. Choose $z \in K^a(q),\ z \neq q$. Suppose $z' \in L$ is such that q and z separate a from z' in ∂H^2. We will prove that $(z', a) \in K(q)$.

Assume for now that a is a cusp point and choose a sequence $p_0 = a, ..., p_k = q, ...$ aimed at z. Pass to an upper halfplane model with $a = \infty$ and $q < z' < z$. Consider a sequence $p'_j,\ j \in J$, aimed at z' with $p'_0 = a$. We wish to show that $p'_j = p_i$ for some $j > 0$ and some $i \leq k$. If not, then (6.2*) gives a component C in the complement of the aimed segment $p_0, ..., p_k$ that contains all the $p'_j,\ j > 0$. Since $z' < z$, (6.1) implies that $p'_1 \leq p_1$. As $p_k < z'$ and z' is in the closure of C, we must have $p_k < p'_1 < p_1$. Thus $k > 1$ and $C \subset (p_k, p_1)$. Now one may apply (6.2*) to the aimed segment $p_1, ..., p_k$ and the aimed sequence $p'_j,\ j \geq 0$.

We find that the aimed sequence lies in (p_k, p_1), contradicting the fact that $p'_0 = \infty$. Thus $p'_j = p_i$ for some $j > 0$ and some $i \leq k$. Since the p'_j are distinct, $i > 0$. If $i = k$ we have $(z', a) \in K(q)$. If $i < k$ we let $a^* = p_i < q$ and $k^* = k - i < k$. By induction on k, we find $p'_l = q$ for some $l \geq j$ so $(z', a) \in K(q)$.

Now take any $a \in L - Q$, and choose an aimed sequence p_i with limits (z, a) and $p_0 = q$. When $i < 0$ and q and z separate p_i from z', as holds for sufficiently negative i, the previous paragraph shows that $(z', p_i) \in K(q)$. Since $K(q)$ is closed, $(z', a) \in K(q)$. Thus $K^a(q)$ contains $L \cap I_z$ where I_z is the interval in $\partial H^2 - \{a\}$ with endpoints z and q. Taking the union over all z proves the lemma.

Note that Lemma 6.7 follows from Lemma 6.5 when each chain $S(q, ...)$ contains q. This may fail to hold, however, even for ridge shadow families, as we will see shortly.

The monotone boundary arcs in Figure 2 are explained by the following theorem.

Theorem 6.8. *Consider a horoball packing $B(p)$, $p \in Q$, of H^2 and a standard shadow family. Fix a cusp point q and choose an upper halfplane model with $q = \infty$ and $B(q) = B_2$. For finite points $(z, a) \in K(q)$ and $(z', a') \in \Omega$ with $z' \leq z < a \leq a'$ or $z' \geq z > a \geq a'$ we have $(z', a') \in K(q)$. The set $K^>(q)$ of finite points of $K(q)$ with $a > z$ consists of all points on or above the graph of a function $\sigma^>(x)$, where $\sigma^> : L - \{q\} \to L - \{q\}$ is monotone and left-continuous. Likewise the set $K^<(q)$ of finite points of $K(q)$ with $a < z$ consists of all points on or below the graph of a function $\sigma^<(x)$, where $\sigma^< : L - \{q\} \to L - \{q\}$ is monotone and right-continuous. The differences $\sigma^>(x) - x$ and $x - \sigma^<(x)$ lie in $[1/b, \alpha\beta]$.*

The first assertion follows from Lemmas 6.6 and 6.7. As $K(q)$ is compact by Prop. 5.1, the sets $K^>(q)$ and $K^<(q)$ are as described. The last statement now follows from Prop. 5.1.

We now construct a one-parameter family of horoball packings of H^2 that includes the Ford disc packing as well as packings that differ from it in some important ways. Consider distinct points p, $q \in \partial H^2$ and a point $w \in H^2$ with $\angle qwp = \pi/3$. Let R_p be the reflection that fixes p and w and define R_q in a similar fashion. Let ι be the rotation of order two that interchanges p and q. The group

$$\Gamma = < R_p, R_q, \iota > \tag{6.9}$$

is discrete and the triangle with vertices p, q, and w is a fundamental domain for Γ. Γ is the free product of $D_3 = < R_p, R_q >$ and $C_2 = < \iota >$. Let w' be the projection of w to $\overline{pq}$ and d the distance from w' to $\iota(w')$. Then Γ is determined up to conjugacy by the nonnegative parameter d.

Since H^2/Γ has just one cusp, there is a unique maximal Γ-invariant horoball family. Consider the Poincare disc model with center w. For $d = 0$, the horoballs nearest to w are 6 equal discs, each tangent to its neighbors on either side. For $0 < d \leq \log 3$, three of these discs grow and three shrink while they remain tangent and keep their D_3-symmetry. For $d = \log 3$, the large discs are tangent to one another and one has the Ford disc packing. For $d > \log 3$, the large discs stay the same but lose contact with the small discs. For each d, x interchanges the small disc based at q with its large neighbor based at p.

Since H^2/Γ has only one cusp, any two invariant horoball packings of H^2 give the same fans and ridge shadows. We will now calculate these, up to the action of Γ, for $d > 0$. One may suppose that $\iota(w')$ is between w' and q. Let ξ be the midpoint of $\overline{w'\iota(w')}$, so that $\iota(\xi) = \xi$. There is a point $\xi' \in \overline{wq}$ so that $\angle \xi' \xi q = \pi/2$. Then $R(p,q) = \overline{\xi' \iota(\xi')}$ and $R(p,p') = \overline{w\xi'}$, where $p' = R_q(p)$. As any pair in A is Γ-equivalent to (p,q) or (p,p'), this determines all the ridges.

Consider the upper halfplane model with $p = \infty$, $q = 0$, and $w = \exp(2\pi i/3) - 1$, as depicted in Figure 7 of §15. In these coordinates, $R_p(z) = -3 - \overline{z}, R_q(z) = -\overline{z}/(1+\overline{z})$, and $\iota(z) = -\mu/z$ for some $\mu > 0$. Thus $w' = \sqrt{3}i$ and $\iota(w') = (\mu/\sqrt{3})i$ so that $\mu < 3$ and $d = \log(3/\mu)$. Also $p' = -1$, $\xi = \sqrt{\mu}i$, and ξ' satisfies $|z|^2 = \mu$, $|z+1|^2 = 1$ so $\Re\xi' = -\mu/2$. Thus $S(0,\infty) = [-\mu/2, \mu/2]$, $S(-1,\infty) = [-3/2, -\mu/2]$, and every shadow is Γ-equivalent to one of these two.

For $2 < \mu < 3$ one has $-1 \notin S(-1,\infty)$, as in Figure 7. This possibility was mentioned before Thm. 3.4. The only aimed sequence with $(z,a) = (-1,\infty)$ is ∞, 0, -1 even though -1 is adjacent to ∞. Applying R_q, we find $\infty \notin S(\infty,-1)$ and so $K(\infty,-1)$ has finite width, unlike the sets $K(\infty,p)$ in Figure 2. Moreover, contrary to what one expects from continued fraction theories, one can have an aimed sequence such as $p_0 = \infty$, $p_1 = -1$, $p_2 = -3/2$ such that the truncated sequence p_0, p_1 is not aimed.

§15 gives special values of μ for which $K(\infty)$ is bounded by two staircase curves, much like Figure 2.

Section 7. Transversals and flowboxes

We will used aimed sequences to describe how flowlines of the geodesic flow ϕ_t on $\mathcal{R}$ cross a certain family of closed transversals $T(p)$, $p \in Q$.

Fix $p \in Q$. For $(z,a) \in \Omega$ with $z \neq p$ we let $v(z,p,a)$ be the unit tangent vector to $\overrightarrow{az}$ at the point where this directed line enters the horoball $B(z,p)$. The set of such $v(z,p,a)$ is closed in $\mathcal{R}$. We define $T(p) \subset SH^n$ as the set of all $v(z,p,a)$ such that there is an aimed sequence p_i, $i \in I$, with $0 \in I$, $0 < \sup(I)$, $p_0 = p$, and limit pair (z,a). Thus $(z,a) \in K(p)$ and $z \neq p$. Indeed, under the projection from $\mathcal{R}$ to Ω, $T(p)$ is homeomorphic to $K(p) - \{(z,a) : z = p\}$. It follows from Prop. 5.1 that this set is closed in $\Omega - \{(z,a) : z = p\}$ and so $T(p)$ is a closed transversal to ϕ_t.

By Existence, every flowline in $\mathcal{R}$ meets at least one $T(q)$. Indeed a flowline that meets only one such closed transversal consists of the unit tangent vectors to a ray $\overrightarrow{qr}$, where $(r,q) \in A$.

We now use this family of transversals to define a family of flowboxes $\Phi(q,p) \subset \mathcal{R}$, $(q,p) \in A$. If p_i, $i \in I$, is an aimed sequence with limit pair (z,a) then Monotonicity implies that the vectors $v(z,p_i,a) \in T(p_i)$, $i \in I$, $i < \sup(I)$, occur along $\overrightarrow{az}$ in the order of their indices. We fix $(q,p) \in A$ and define $\Phi(q,p)$ as the union over $(z,a) \in K(q,p)$ of certain closed intervals $I(z,q,p,a)$ in the flowline of the vector $v(z,p,a)$. Here $I(z,q,p,a)$ starts at $v(z,p,a) \in T(p)$ and it either ends at $v(z,q,a) \in T(q)$, provided $z \neq q$, or is unbounded if $z = q$. We will show

Proposition 7.1. *If $z \neq q$ then the return time τ for $v(z,p,a) \in T(p) \cap \Phi(q,p)$ (that is, the change in t over $I(z,q,p,a)$) is $\tau(z,q,p) = 2t(z')$. $\Phi(q,p)$ is closed in $\mathcal{R}$.*

The formula for τ is easily checked in the coordinates used to prove Lemma 2.1, where $p = 0$, $h(B(0)) = e^{-\delta}$, $q = \infty$, and $\partial B(\infty)$ has height e^{δ}. Then $h(B(z,\infty)) = e^{\delta}$ and, as in the proof of Prop. 2.4, $h(B(z,p)) = |z|^2 e^{\delta}$. This gives $\tau = \log|z|^2 = 2t(z')$. The flowbox is closed because $K(q,p)$ is closed, $T(p)$ is closed, τ is continuous for $q \neq z$, and τ approaches infinity as z approaches q.

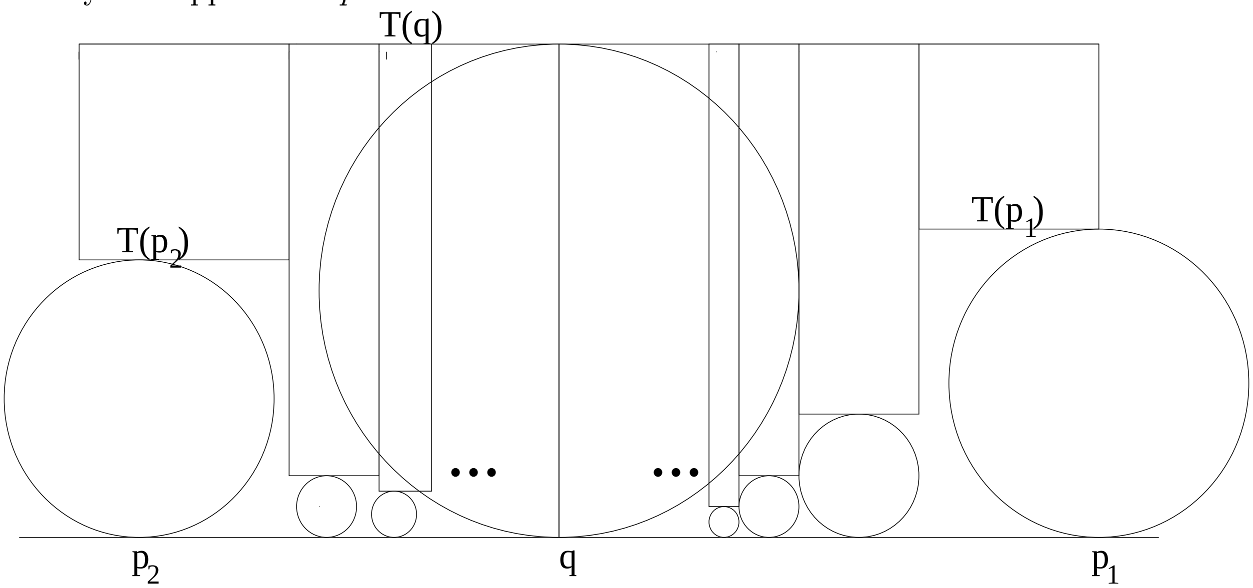

Figure 3. Sections of flowboxes with z fixed, projected to hyperbolic space

We must mention another natural family of transversals. Let $T'(p),\ p \in Q$, consist of those $v \in \mathcal{R}$ that are both tangent to $\overrightarrow{az}$ for some $(z,a) \in K(p)$ and tangent to the boundary of some horoball based at p. Suppose a flowline $\phi_t(v)$ has limits $(z,a) \in \Omega_L$ and that $p_i,\ i \in \mathbf{Z}$, is an aimed sequence with these limits. Then there are $t_i \in \mathbf{R},\ i \in \mathbf{Z}$, such that $\phi_{t_i}(v) \in T(p_i)$. In any upper halfspace model of H^n, one has

$$t_{i+1} - t_i = \log\left(\frac{|p_{i+1}-a||p_i-z|}{|p_i-a||p_{i+1}-z|}\right).$$

This is easily seen in a model with $z = \infty$ and $a = 0$. When $t_{i+1} - t_i$ is always positive, one can build a flowbox family $\Phi'(q,p),\ (q,p) \in A$, bounded by the transversals $T'(p),\ p \in Q$. For the Ford disc packing and the ridge shadow family, for instance, one can reduce to the case $p_{i+1} = \infty$ and $p_i = 0$ and use (5.3) to see that $t_{i+1} - t_i \geq \log(1+\sqrt{5})$. The flowboxes used in [F2] are similar to these $\Phi'(q,p)$'s.

Next we fix a cusp point q and consider the various flowboxes $\Phi(q,p)$. Define the *unstable set* $W^u(q) \subset \mathcal{R}$ to consist of all the unit tangent vectors to the rays $\overrightarrow{qz},\ z \in L,\ z \neq q$. We define the *flowbox cluster* $\Phi(q) = (W^u(q) \cap \pi^{-1}B(q)) \cup (\cup\, \Phi(q,p)) \subset \mathcal{R}$, where p varies over all cusp points with $(q,p) \in A$. In Figure 3 we show the projection to H^n of the vertical vectors in a flowbox cluster $\Phi(q)$, that is the section of this set corresponding to $z = \infty$. The section of $\Phi(q,p)$ with $z = \infty$ projects to a rectangle whose upper and lower boundaries correspond respectively to the transversals $T(q)$ and $T(p)$. The horizontal factor is the slice $K_z(q,p)$ and the vertical factor has hyperbolic length $\tau(z,q,p)$.

We now show that the flowbox cluster $\Phi(q)$ approximates $\pi^{-1}B(q)$, in the sense that their Hausdorff distance is bounded.

Proposition 7.2. *For some $D = D(c,\Delta) > 0$, $\pi(\Phi(q))$ lies in the D-neighborhood of $B(q)$*

and $\pi^{-1}B_D(q) \subset \Phi(q)$. The $\Phi(q)$ form a closed cover of $\mathcal{R}$ and the packing $B(q)$, $q \in Q$, is D-dense.

Let $D = \log(c + e^{2\Delta}/(1-c)^2) > 2\Delta$. Choose $v \in I(z,q,p,a)$ and let d be the distance from $\pi(v)$ to $B(q)$. We will show $d \le D$. Choose an upper halfspace model with $z = \infty$ and let 2δ be the distance from $B(p)$ to $B(q)$, h the height of $\pi(v)$, $h(p)$ the diameter of $B(p)$, and $\sigma = bh(p) + \sqrt{h(p)}e^{\delta}$. Since $d \le 2\delta < D$ when $z = q$, we may suppose $z \ne q$ and that the model is chosen so $q = 0$ and $B(q)$ has diameter 1. By (1.2), $e^d = h + |a|^2/h$. Since $|a| \le |a-p| + |p-q| \le \sigma$ by (3.6) and Prop. 2.4, we get $e^d \le h + \sigma^2/h$. Since $\sigma^2 > h(p)$ and $h(p) \le h \le 1$, we find $e^d \le h(p) + \sigma^2/h(p)$. As $\delta \le \Delta$, $b = e^{\Delta}\sqrt{c}/(1-c)$, and $h(p) \le c$ by (3.2), we get $e^d \le c + (b\sqrt{c} + e^{\Delta})^2 = e^D$. So $d \le D$ and the first assertion is proved.

We now show $\Phi(q)$ is closed. Since $\Phi(q)$ is defined as a countable union of certain closed sets, it suffices to take $v_m \in \Phi(q, p_m)$ for a sequence of distinct p_m such that v_m converges to some $v \in \mathcal{R}$ and to show that $v \in W^u(q) \cap \pi^{-1}B(q)$. Say that $v_m \in I(z_m, q, p_m, a_m)$ and that v is tangent to $\overrightarrow{az}$. Then $(z,a) = \lim(z_m, a_m)$. By the local finiteness in Prop. 5.2, $a = q$. Thus $v \in W^u(q)$, $z \ne q$, and $z_m \ne q_m$ for m large. For such m there is a $t_m \ge 0$ with $v'_m = \phi_{t_m}(v_m) \in T(q)$. Since $T(q)$ is a closed transversal to the flow ϕ_t, the limits $v' = \lim v'_m$ and $t = \lim t_m$ exist, with $v' \in T(q)$ and $t \ge 0$. This shows that $v' = v(z,q,q)$ and $v = \phi_{-t}(v')$, $t \ge 0$, so that $\pi(v) \in B(q)$. Thus $\Phi(q)$ is closed.

Let $\mathcal{U} = \cup\ \Phi(q,p)$, $(q,p) \in A$. If $v \in \mathcal{R} - \mathcal{U}$ then $v \in W^u(q) \cap \pi^{-1}(\text{int } B(q))$ for some $q \in Q$. It follows that the packing $B(q)$ is D-dense and that the $\Phi(q)$ cover $\mathcal{R}$. If $\pi(v) \in B_\rho(q)$, $\rho > D$, then v does not lie in $\Phi(q')$ for $q' \in Q$, $q' \ne q$, so v must lie in $\Phi(q)$. As $\Phi(q)$ is closed and contains each set $\pi^{-1}(B_\rho(q))$, $\rho > D$, it contains $\pi^{-1}(B_D(q))$ and the proposition is proved.

Let $\mathcal{R}_0 = \text{Int } \mathcal{U}$. From our description of $\mathcal{R} - \mathcal{U}$, we see that $\mathcal{R} - \mathcal{R}_0$ is the union of the disjoint closed sets $W^u(q) \cap \pi^{-1}B(q)$, $q \in Q$.

We now prove an important bound on our flowbox family.

Theorem 7.3. *A vector $v \in \mathcal{R}$ lies in none of the flowboxes $\Phi(q,p)$, $(q,p) \in A$, if and only if $v \in W^u(r) \cap \pi^{-1}(\text{Int } B(r))$ for some $r \in Q$. If $v \in W^u(r) \cap \pi^{-1}(B(r))$ for some $r \in Q$ then $v \in \Phi(q,p)$ for some $(q,p) \in A$ if and only if $p = r$ and $v = v(z,p,p)$ for some $z \in S(q,p)$. If $v_0 \in \mathcal{R}_0$ then there is a neighborhood of v_0 that meets only a bounded number of flowboxes, with a bound depending only on n, c, and Δ. The flowbox family $\Phi(q,p)$, $(q,p) \in A$, is point finite.*

The first statement has already been noted. For the second, suppose $v \in W^u(r) \cap \pi^{-1}(B(r))$ and $v \in \Phi(q,p)$. Choose z so v is tangent to $\overrightarrow{rz}$ and choose an upper halfplane model with $z = \infty$. Then $(z,r) \in K(q,p)$ so $z \in S(q,p)$ and there is an aimed sequence with limit pair (z,r) with $p_1 = q$, $p_0 = p$, and $p_i = r$ for some $i \le 0$. Let h be the height of $\pi(v)$. Then $h \ge h(B(p))$, since v lies beyond $v(z,p,r)$, and $h \le h(B(r))$, since $\pi(v) \in B(r)$. $c^{-i}h(B(p)) \ge h(B(r)) \ge h(B(p))$, we must have $i = 0$ and $p = r$. Furthermore we must have $v = v(z,p,p)$. The converse is simpler and will be omitted.

Next we fix $v_0 \in \mathcal{R}_0$ and $v \in \Phi(q,p) \cap \mathcal{R}_0$ with $\pi(v)$ at distance at most 1 from $\pi(v_0)$. Then $\pi(v)$ is distance at most D from $B(q)$ by Prop. 7.2, so $B(q)$ meets the $(1+D)$-neighborhood of $\pi(v_0)$. Now Lemma 2.5 shows that q lies in a finite set, bounded in terms of n and D. We now fix q and bound the number of p for which $v \in \Phi(q,p)$.

Let $(z, a) \in \Omega$ be the limit pair of the flowline $\phi_t(v)$. Let p' be the intersection point of $\partial B(p)$ and $\overline{zp}$. By (3.6), the distance from p' to $\pi(v(z,p,a)) \in \overline{za}$ is at most b. But every point of $\overline{za}$ lies at distance less than 1 from $\overline{zq} \cup \overline{qa}$ (this is the Rips property of hyperbolic space). Therefore p' has distance at most $b+1$ from $\overline{zq} \cup \overline{qa}$. The distance from p' to $B(q)$ is bounded by D, as in the proof of Prop. 7.2, and p' is not in the interior of $B(q)$. Thus p' belongs to one of two sets of bounded diameter. To see this, choose an upper halfspace model with $q = \infty$ and $B(q) = B_n$ and note that the cone of points of distance at most $b+1$ from $\overline{qa}$ (respectively from $\overline{zq}$) meets the slab of points of height between 1 and e^{-D} in a compact set, determined up to isometry by n, b, and D, hence by n, c, and Δ. Now we apply Lemma 2.5 again to obtain a uniform bound on the number of p. This gives a uniform bound on the number of flowboxes that meet a neighborhood of v_0.

It only remains to show that each v lies in only finitely many flowboxes when $v \in \mathcal{U} - \mathcal{R}_0$. This follows from the local finiteness of the shadow family $S(q,p)$ for fixed p. Thus the theorem is proved.

Proposition 7.4. *With respect to the measure class on $\mathcal{R}$ associated to the measure $\mu \times \mu$ on the orbit space Ω of ϕ_t, Bd $\Phi(q,p)$ is null. The flowboxes $\Phi(q,p)$ are essentially disjoint and their interiors are disjoint.*

Note that $K(q,p)$ is the orbit space of the flowbox $\Phi(q,p)$. By Uniqueness, a subset of $K(q,p)$ of full measure consists of pairs (z,a) that are represented by a unique aimed sequence, up to shift, which may be taken with $q = p_1$, $p = p_0$. Then for $0 < t < \tau(z,q,p)$, $v_0 = \phi_t(v(z,p,a))$ lies outside all the flowboxes except $\Phi(q,p)$. This shows that the flowboxes are essentially disjoint. In a similar way, Uniqueness implies that the intersections of two flowboxes in our family is nowhere dense in $\mathcal{R}$, hence the interiors of the flowboxes are disjoint. If $v_0 \in \mathcal{R}_0$ then the flowbox cover is locally finite at v_0 and so $v_0 \in$ Int $\Phi(q,p)$. That is $\mathcal{R}_0 - \cup\, \Phi(q',p')$, $(q',p') \in A - \{(q,p)\}$, is an open subset of $\Phi(q,p)$ that contains v_0. If v_0 is not in $\mathcal{R}_0$ then it lies in one of countably many null sets $W^u(q') \cap \pi^{-1}B(q')$, $q' \in Q$. This shows that Bd $\Phi(q,p)$ is null and proves the proposition.

It also follows from Prop. 7.4 that Bd $K(q,p)$ is null, finishing the proof of Prop. 5.2. The interior of $K(p)$ contains the union of the interiors of the sets $K(q,p)$, which has full measure in $K(p)$. Thus Bd $K(p)$ is null, which finishes the proof of Prop. 5.1.

The results of this section shed some light on Thm. 3.4. By Uniqueness, some null set in Ω contains all pairs that are not uniquely described by aimed sequences, up to shift. In fact, such a null set is $\cup$Bd $K(q,p), (q,p) \in A$. By Lemma 4.1, certain cusp points q must occur in aimed sequences with certain limit pairs (z,a). In fact, this holds if some v tangent to $\overrightarrow{az}$ lies outside $\Phi(q')$ for all $q' \neq q$.

We now examine how many aimed sequences can describe a particular geodesic $\overrightarrow{az}$ with endpoints in L. Let us say this geodesic is *forked* if the corresponding flowline meets two transversals at once, that is if some $v \in T(q) \cap T(q')$, $q \neq q'$, is tangent to $\overrightarrow{az}$. In an aimed sequence for an unforked geodesic, each pair of consecutive terms determines the entire aimed sequence. Thus the bound in Thm. 7.3 gives the following.

Corollary 7.5. *If a geodesic is not forked then only a bounded number of aimed sequences, up to shift, describe it.*

A forked geodesic, however, may be described by an infinite number of aimed sequences, up to shift, at least if $n \geq 3$. We will give natural examples of this in §16.

Section 8. Correspondences and symmetry

We now relate aimed sequences to correspondences and their inverse limits.

If Y and Z are topological spaces, a *correspondence* from Y to Z is given by a continuous mapping $f : W \to Y \times Z$, where W is some topological space. If the first component $f_1 : W \to Y$ is a homeomorphism to its image, we obtain a mapping $f_2 \circ f_1^{-1} : f_1(W) \to Z$, so correspondences from Y to Z generalize mappings from subsets of Y to Z. The composition of correspondences $g : V \to X \times Y$ and $f : W \to Y \times Z$ is $h = f \circ g : U \to X \times Z$, where U is the pullback of f_1 and g_2 and h is the natural mapping. Specifically, $(v, w) \in U \subset V \times W$ if $g_2(v) = f_1(w)$, in which case $h_1(v, w) = g_1(v)$ and $h_2(v, w) = f_2(w)$. This generalizes the composition of mappings.

When $Y = Z$ one has a *self-correspondence* on Y that one may iterate. Define a $\mathbf{Z}-$*orbit* of f to be a sequence $w_m \in W$, $m \in \mathbf{Z}$, with $f_2(w_m) = f_1(w_{m+1})$ for all $m \in \mathbf{Z}$. Let the *inverse limit* of f be the space $\hat{f}$ of all $\mathbf{Z}-$orbits of f, topologized as a subspace of the Cartesian product $W^{\mathbf{Z}}$. This space carries a natural *shift* homeomorphism σ with $\sigma(w_m) = (w'_m)$ where $w'_m = w_{m+1}$. When f is understood, we will denote the correspondence by W and its inverse limit by $\hat{W}$. If $Y = Z$ and f_1 is a homeomorphism to its image then $\hat{W}$ can be identified with a space of sequences in Y. Namely $y_m = f_1(w_m)$ satisfies $f_2(f_1)^{-1} y_m = y_{m+1}$ for all integers m, and any such sequence (y_m) corresponds to a point of $\hat{W}$.

Given a horoball packing and shadow family, consider the space of all aimed sequences indexed by $\mathbf{Z}$, or $\mathbf{Z}$-*aimed sequences*, with its natural shift mapping. We will identify this space and shift with the inverse limit of a certain natural correspondence. In fact, various correspondences work, for instance the inverse of the Gauss correspondence of [F3], so we select one that seems the simplest. For $(q, p) \in A$, let $\mathcal{A}(q, p) = S(q, p) - \{q\}$ and consider the correspondence $\mathcal{A}(q, p) \to (L - \{p\}) \times (L - \{q\})$ given by the diagonal mapping $z \to (z, z)$. This diagonal correspondence is an expanding mapping provided we use the Euclidean distance functions d_p on $L - \{p\} \subset \partial H^n - \{p\}$ and d_q on $L - \{q\} \subset \partial H^n - \{q\}$. This holds because $|v|_p \leq c|v|_q$ holds for tangent vectors v on a conformal disc in ∂H^n and this disc contains $\mathcal{A}(q, p)$ by (3.2*).

In order to build a self-correspondence, we take disjoint unions. Let $\mathcal{L} = \coprod L - \{p\}$, $p \in Q$, be the disjoint union space of the punctured limit sets, so a point of $\mathcal{L}$ is an ordered pair $(p, z) \in Q \times L$, $z \neq p$. Let $\mathcal{A}$ be the space $\coprod \mathcal{A}(q, p)$, $(q, p) \in A$, so that a point of $\mathcal{A}$ is a triple (q, p, z) with $(q, p) \in A$ and $z \in \mathcal{A}(q, p)$. The union of the diagonal correspondences defined above gives a self-correspondence $\mathcal{A} \to \mathcal{L} \times \mathcal{L}$, $(q, p, z) \to ((p, z), (q, z))$, that we call the *aimed correspondence*. When p_i, $i \in \mathbf{Z}$, is aimed at z, the sequence $l_i = (p_i,\ p_{i-1}, z) \in \mathcal{A}(p_i, p_{i-1})$, $i \in \mathbf{Z}$, is a $\mathbf{Z}-$orbit of $\mathcal{A}$. Conversely, any $\mathbf{Z}-$orbit of the correspondence $\mathcal{A}$ defines a $\mathbf{Z}$-aimed sequence p_i, so the space of all $\mathbf{Z}$-aimed sequences is identified with the inverse limit $\widehat{\mathcal{A}}$ of the aimed correspondence by a homeomorphism that respects the shifts.

Consider a flow ψ_t on a locally compact metrizable space S that preserves a Borel measure class. An *orbit-equivalence* from a self-correspondence W to ψ_t is a continuous mapping $\Pi : \widehat{W} \to S$ that sends orbits of the shift σ to flowlines of ψ_t such that the inverse image of a flowline consists of a single orbit of σ, with the exception of a union of flowlines in S that is meager and null. The next proposition will show that the aimed correspondence is orbit-equivalent to ϕ_t when Q is null.

Define $V : \widehat{\mathcal{A}} \to \mathcal{R}$ by $V(l_i,\ i \in \mathbf{Z}) = v(z, p_0, a)$, where $l_i = (p_i, p_{i-1}, z) \in \mathcal{A}(p_i, p_{i-1})$ and $p_i,\ i \in \mathbf{Z}$, is an aimed sequence with limits (z, a). We give $\mathcal{R}$ the measure class of Prop. 7.4. Let $\mathcal{R}^*$ be the set of all $v \in \mathcal{R}$ such that $\phi_t(v)$ corresponds to a pair $(z, a) \in \Omega$ that is represented by only one $\mathbf{Z}$-aimed sequence, up to a shift. The following properties of V then follow from Uniqueness.

Proposition 8.1. *V maps orbits of $\sigma : \widehat{\mathcal{A}} \to \widehat{\mathcal{A}}$ to flowlines of ϕ_t. There is a set $\mathcal{R}^* \subset \mathcal{R}$, invariant by ϕ_t, such that for all $v \in \mathcal{R}^*$, the inverse image $V^{-1}\{\phi_t(v) : t \in \mathbf{R}\}$ consists of a single orbit of σ. $\mathcal{R} - \mathcal{R}^*$ is meager in $\mathcal{R}$. When Q is null, $\mathcal{R}^*$ has full measure in $\mathcal{R}$ and so $\mathcal{A}$ is orbit-equivalent to ϕ_t.*

We define a *symmetry group* of our horoball packing and shadow family to be a discrete group Γ of hyperbolic isometries that stabilizes Q and A such that for each $\gamma \in \Gamma,\ q \in Q$, and $(q, p) \in A$ one has $\gamma(B(q)) = B(\gamma q)$ and $\gamma(S(q, p)) = S(\gamma q, \gamma p)$. Let Γ_p be the stabilizer of a cusp point p and let $\Gamma_{(q,p)}$ be the stabilizer of $(q, p) \in A$ in Γ. As Γ_p preserves $B(p)$, it acts isometrically on $L - \{p\}$ with the distance function d_p. As Γ is discrete, this action is properly discontinuous. Let $\Lambda_p = (L - \{p\})/\Gamma_p$. Note also that $\Gamma_{(q,p)}$ is the stabilizer of q in Γ_p, so it is finite.

The action of Γ on SH^n commutes with the geodesic flow and preserves $\mathcal{R}$. So one obtains the *reduced geodesic flow* $\overline{\phi}_t$ on $\mathcal{R}/\Gamma \subset H^n/\Gamma$. If $[v]$ denotes the image in $\mathcal{R}/\Gamma$ of $v \in \mathcal{R}$, then $\overline{\phi}_t([v]) = [\phi_t(v)]$.

In order to study this flow, we introduce another correspondence. The symmetry group Γ acts properly discontinuously on $\mathcal{L}$ and $\mathcal{A}$. The natural injection $\mathcal{A} \to \mathcal{L} \times \mathcal{L}$ is equivariant and induces a natural mapping of quotient spaces

$$\mathcal{A}/\Gamma \to \mathcal{L}/\Gamma \times \mathcal{L}/\Gamma \tag{8.2}$$

that we call the *reduced aimed correspondence* $\mathcal{A}/\Gamma$ on $\mathcal{L}/\Gamma$. We will show that in many cases the reduced aimed correspondence is orbit-equivalent to $\overline{\phi}_t$.

For instance, take the packing of H^2 by Ford discs and let $\Gamma^+ = PSl(2, \mathbf{Z})$. As Γ^+ acts simply transitively on adjacent pairs of cusp points, we may identify $\mathcal{L}/\Gamma^+$ with $\mathbf{R}/\mathbf{Z}$. Then $\mathcal{A}/\Gamma^+$ is identified with the expanding map E on $\mathbf{R}/\mathbf{Z}$ with $E(x + \mathbf{Z}) = -1/x + \mathbf{Z}$ for $0 < |x| \leq 1/2$.

We say that a Γ-set X (that is, a set X with an action of Γ) is *Q-free* for Γ if the stabilizer in Γ of each point in X acts trivially on Q (and thus acts trivially on L). When Q is not contained in a proper subsphere of ∂H^n, any element of Γ that fixes Q must be the identity, and so X is Q-free only when the action of Γ on X is free. The distinction between free and Q-free is significant, however, in cases such as a Fuchsian group acting on H^3.

When the action of Γ on A is Q-free, we can describe the reduced aimed correspondence $\mathcal{A}/\Gamma$ as follows. Choose a set $Q' \subset Q$ that includes exactly one cusp point from each orbit of Γ on Q. We identify $\mathcal{L}/\Gamma$ with $\coprod \Lambda_p,\ p \in Q'$. Now choose for each $p \in Q'$ a set Q'_p of representatives for the orbits of Γ_p on the set of cusp points q with $(q, p) \in A$. As A is Q-free, each stabilizer $\Gamma_{(q,p)}$ acts trivially on $\mathcal{A}(q, p)$ so we may identify $\mathcal{A}/\Gamma$ with $\coprod \mathcal{A}(q, p),\ p \in Q', q \in Q'_p$. We choose $\gamma_q \in \Gamma,\ q \in Q$, so that $\gamma_q(q) = q' \in Q'$. The (q, p)th term in the self-correspondence on $\coprod \Lambda_p,\ p \in Q'$, is just the correspondence $\mathcal{A}(q,p) \to \Lambda_p \times \Lambda_{q'}$ induced by $(id, \gamma_q) : \mathcal{A}(q,p) \to (L - \{p\}) \times (L - \{q'\})$.

When $S(q,p)$ is the ridge shadow family, the same mappings γ_q give face pairings for H^n/Γ. Let $\kappa_p,\ p \in Q'$, be the natural projection from H^n to H^n/Γ_p. Then $H^n/\Gamma = \cup\, \kappa_p F(p)$, the interiors of these regions are disjoint, and their boundaries are identified by equating $\kappa_p(x),\ x \in R(q,p)$, with $\kappa_{q'}(\gamma_q(x)),\ q' = \gamma_q(q)$. Thus $\mathcal{A}/\Gamma$ is a dynamical system constructed from the same face pairings γ_q that glue the fan quotients $F(p)/\Gamma_p$ into the orbifold H^n/Γ.

We now use study when $\mathcal{A}/\Gamma$ is orbit-equivalent to the flow $\overline{\phi}_t$ on $\mathcal{R}/\Gamma$, where $\mathcal{R}/\Gamma$ is given the measure class induced by the projection $\mathcal{R} \to \mathcal{R}/\Gamma$.

Theorem 8.3. *When Q is null and $Q \times (L - Q)$ is Q-free for Γ, there is an orbit-equivalence Π_Γ from $\mathcal{A}/\Gamma$ to $\overline{\phi}_t$.*

V is Γ-equivariant and so induces a mapping $V/\Gamma : \widehat{\mathcal{A}}/\Gamma \to \mathcal{R}/\Gamma$. There is also a natural mapping $P_\Gamma : \widehat{\mathcal{A}}/\Gamma \to \widehat{\mathcal{A}/\Gamma}$ given by $P_\Gamma([l_i, i \in \mathbf{Z}]) = ([l_i], i \in \mathbf{Z})$. To define $\Pi_\Gamma : \widehat{\mathcal{A}/\Gamma} \to \mathcal{R}/\Gamma$, we must invert P_Γ.

Proposition 8.4. *P_Γ is continuous and onto and it commutes with the shift mappings. It is a homeomorphism when $Q \times (L - Q)$ is Q-free for Γ. This holds when the stabilizer of each cusp point is torsionfree.*

The first sentence is simple to verify. Let the *effective factor* Eff(Γ) be the image of Γ in the group of homeomorphisms of L. Assuming $Q \times (L - Q)$ is Q-free for Γ, we show that the natural mapping $\widehat{\mathcal{A}} \to \widehat{\mathcal{A}/\Gamma}$ is a regular covering mapping with group Eff(Γ). Consider two $\mathbf{Z}$-orbits (l_i) and (l_i') of $\mathcal{A}$, so that $l_i = (p_i, p_{i-1}, z)$ and $l_i' = (p_i', p_{i-1}', z')$ for all $i \in \mathbf{Z}$, where p_i is aimed at z and p_i' is aimed at z'. When these orbits have the same image in $\widehat{\mathcal{A}/\Gamma}$, there are $\gamma_i \in \Gamma$ so that $\gamma_i(l_i) = l_i'$, for each integer i. As $(p_i, z) \in Q \times (L - Q)$ and $\gamma_i(p_i, z) = \gamma_{i+1}(p_i, z)$, we see that all the γ_i have the same image in Eff(Γ). This shows that Eff(Γ) acts simply transitively on each level set in $\widehat{\mathcal{A}}$. Similar reasoning shows that $\widehat{\mathcal{A}} \to \widehat{\mathcal{A}/\Gamma}$ is a covering mapping, as desired. A torsionfree stabilizer Γ_p acts by translations on the Euclidean space $\partial H^n - \{p\}$, so it acts freely on $L - Q$. When this holds for all $p \in Q$, Γ acts freely on $Q \times (L - Q)$. The proposition is proved. The theorem follows from the last two propositions with $\Pi_\Gamma = (V/\Gamma) \circ (P_\Gamma)^{-1}$.

To illustrate Thm. 8.3, consider the "test case" when $n = 2$, H^2/Γ has finite area and just one cusp, and $\Gamma_\infty = \mathbf{Z}$. As (Int B_2)$/\mathbf{Z}$ injects into H^2/Γ ([MT], Prop. 0.7), we obtain an invariant horoball packing by the rule $B(\gamma(\infty)) = \gamma(B_2)$. Since a point w of H^2 lies closer to B_2 than to $\gamma(B_2)$ if and only if w is higher than $\gamma(w)$, the fan $F(\infty)$ is the Ford region. Each ridge $C = R(q, \infty)$ is transformed by an edge-pairing $g = \gamma_q$ preserving the height of its points. The ridge shadow $S(q, \infty)$ is the interval I under C. The reduced aimed correspondence $\mathcal{A}/\Gamma$ for the ridge shadow family is the piecewise smooth expanding mapping E of $\Lambda_\infty = \mathbf{R}/\mathbf{Z}$ given by $E([z]) = [g(z)],\ z \in I,\ g(z) \neq \infty$. By Thm. 8.3,

(8.5) the inverse limit of E is orbit-equivalent to the geodesic flow over H^2/Γ.

This result is related to the work of Bowen and Series [BS], [Se] although they work in another model of H^2 and produce a different expanding mapping.

Thm. 8.3 describes orbits of the reduced geodesic flow using orbits of the reduced aimed correspondence. This is an instance of a standard procedure in dynamical systems. For instance, one may study a 3-dimensional flow by passing to 2-dimensional transversals and then to 1-dimensional maps. The final step, which uses a shift on a 0-dimensional space, is known as *symbolic dynamics*.

One approach uses A/Γ as the space of *symbols*. Say that a symbol sequence $[(q_i, p_i)] \in A/\Gamma$, $i \in \mathbf{Z}$, is *allowed* if for each i there is an aimed segment p_i, q_i, r_i with (r_i, q_i) Γ-equivalent to (q_{i+1}, p_{i+1}). An allowed sequence is an orbit of a self-correspondence $B/\Gamma \to A/\Gamma \times A/\Gamma$ of discrete spaces, where B is the set of aimed segments $p,\ q,\ r$ with three terms. Each aimed sequence indexed by $\mathbf{Z}$ determines an allowed sequence and Γ-equivalent aimed sequences determine the same allowed sequence. When Γ acts freely on A, this mapping from Γ-orbits of $\mathbf{Z}$-aimed sequences to allowed sequences is one to one. Its image is a space of allowed sequences that models almost all of the flow $\overline{\phi_t}$, provided Q is null. In general one cannot identify this image in an explicit way. The next two sections, however, treat cases in which an explicit set of symbol sequences code geodesics. In the best cases, one obtains a correspondence on a finite space that is orbit-equivalent to $\overline{\phi_t}$, c.f. Cor. 11.4.

Section 9. Markov partitions for chain-finite shadow families

Let $B(q),\ q \in Q$, be a horoball packing and $S(q,p),\ (q,p) \in A$, a shadow family. Given $(q,p) \in A$, we consider the chains $S(p_1, p_0, ..., p_m)$, with $(p_1, p_0) = (q,p),\ m \leq 0$. We say that the shadow family is *chain-finite* when this collection of compact subsets of $S(q,p)$ is finite for every $(q,p) \in A$. We will show how one can then refine the flowbox cover to a Markov partition in a standard way. §10 uses this Markov partition to derive a symbolic dynamics for the geodesic flow on $\mathcal{R}$.

Following Sinai and Bowen, we define a *rectangle* as a compact subset of Ω that is a Cartesian product in the (z,a) coordinates. Fix $(q,p) \in A$. For each $z \in S(q,p)$, form the rectangle $R(z) = S(z) \times K_z(q,p)$. By Lemma 5.4, $K(q,p) = \cup\ R(z),\ z \in S(q,p)$, and $R(z) = R(z')$ if z and z' lie in the same chains $S(q,p,...)$. We obtain

Corollary 9.1. *If the shadow family $S(q,p),\ (q,p) \in A$, is chain-finite then each $K(q,p)$ is a finite union of rectangles.*

When the ridge shadow family for a horoball packing is chain-finite, we will also say this packing is chain-finite. For example, the Ford disc packing of H^2 is chain-finite. This reduces to the fact that the only chains $S(\infty, 0, ...)$ are $[2, \infty] = S(\infty, 0, -1/2),\ [\infty, -2] = S(\infty, 0, 1/2)$, and their union $S(\infty, 0)$. Cor. 9.1 is then consistent with (5.3), which expresses $K(\infty, 0)$ as the union of two rectangles.

The rectangles of Cor. 9.1 may overlap severely. To produce nonoverlapping rectangles, we proceed as in Bowen's theory for compact hyperbolic systems ([B3], Thm. 3.12). We need the topological version of the method of inclusion/exclusion given in the next proposition.

Given a topological space X and a subset $Y \subset X$, the *core* of Y in X is $\mathrm{Core}_X\ Y = \mathrm{Cl}(\mathrm{Int}\ Y)$. When X is clear from context, as when X is $L,\ L \times L$, or $\mathcal{R}$, this will be written Core Y. We often use that $\mathrm{Core}(Y - Z) = \mathrm{Core}\ Y$ if Int Z is empty. Y is *proper* if $Y =$ Core Y, that is if Y is closed and Int Y is dense in Y. For instance, Prop. 5.2 shows that each slice $K_z(r,q)$ is proper.

Let X be a locally compact metric space equipped with
1) a measure class of positive Borel measures with support equal to X and

2) a point finite family of closed sets $X_j,\ j \in J$, such that J is countable and Bd X_j is null for all j.
For each finite subset $F \subset J$, let $X(F) \subset X$ consist of the points that belong to X_j if and only if $j \in F$. Then the disjoint sets $X(F)$ cover $\cup X_j,\ j \in J$. Let $\mathcal{F}$ be the family of all F for which Int $X(F)$ is nonempty.

Proposition 9.2. *When the family $X_j,\ j \in J$, is locally finite, the family Core $X(F),\ F \in \mathcal{F}$, is a countable locally finite cover of X by proper sets with null boundary. Two of these cores can only meet at points in their boundaries. The core of $\cup X_j,\ j \in J$, is the union of the Core $X(F)$ for nonempty $F \subset J$.*

Note that the union $\mathcal{U} = \cup$ Bd $X_j,\ j \in J$, is closed and null and its complement is $\cup$Int $X(F),\ F \in \mathcal{F}$. By the Baire category theorem, $\mathcal{U}$ has empty interior and so the sets Core $X(F)$ cover X. Clearly this is a locally finite countable family of proper sets. For each F, Bd(Core $X(F)$) lies in $\mathcal{U}$, hence is null. For the second statement, suppose x lies in Core $X(F_i),\ i = 1,2$, with $F_1 \neq F_2$. Then each neighborhood of x meets the interior of $X(F_i),\ i = 1,2$, and so x lies in Bd(Core $X(F_i)), i = 1,2$. For the last statement, local finiteness implies that $\cup X_j$ is closed and that $\cup$ Bd $X_j \subset \mathcal{U}$. It follows that $\mathcal{O} = \cup X_j - \mathcal{U}$ is open in X and forms a dense subset of int $\cup X_j$. Thus $\overline{\mathcal{O}} =$ Core $(\cup X_j)$. But $\mathcal{O}$ is the locally finite union of the sets int $X(F)$ over all nonempty F, so $\overline{\mathcal{O}} = \cup$ Core $X(F)$ and the proposition is proved.

Fix a horoball packing $B(q),\ q \in Q$, and a chain-finite shadow family $S(q,p),\ (q,p) \in A$. Suppose L is the support of our measure μ, so that every open null set in ∂H^n is disjoint from L. We will apply Prop. 9.2 twice to construct new flowboxes that will form our Markov partition.

We first fix a cusp point p and let $X(p) = L - \{p\}$. Let $J(q,p),\ (q,p) \in A$, be disjoint finite index sets for the families of chains $S(q,p,...)$, so $S(q,p,...) = X(p)_j$ for some $j \in J(q,p)$, and let $J(p) = \cup J(q,p)$. By (3.1), the chains $X(p)_j,\ j \in J(p)$, form a countable locally finite cover of $X(p)$ by compact sets. As Bd $S(q,p,...)$ lies in a finite union of shadow boundaries, each of which is null, these chains obey the hypotheses of Prop. 9.2. For a finite set $F \subset J(p)$, $X(p)(F)$ is null (and so has empty interior) unless all the $j \in F$ are attached to one pair $(q,p) \in A$. Thus the family of finite subsets of $J(p)$ used in Prop. 9.2 can be written $\mathcal{F}(p) = \cup \mathcal{F}(q,p),\ (q,p) \in A$, where each $F \in \mathcal{F}(q,p)$ is a subset of $J(q,p)$. For $F \in \mathcal{F}(q,p)$, $X(p)(F) \subset S(q,p)$ consists of the points that belong to $X(p)_j$ if and only if $j \in F$. Define the *tile* $Z(F) =$ Core $X(p)(F)$. Two generic points $z,\ z' \in Z(F)$ must lie in the same chains and so $K_z(q,p) = K_{z'}(q,p)$. We let $A(F)$ denote this slice and we let $K(F) = Z(F) \times A(F) \subset K(q,p)$.

We now consider the flowbox family $\Phi(q,p) \subset \mathcal{R},\ (q,p) \in A$. By Thm. 7.3, this is a point finite family and its restriction to the open subset $\mathcal{R}_0$ is locally finite. $\mathcal{R}$ has a preferred positive Borel measure class of full support. For $j \in J(q,p)$ we let $\mathcal{R}_j$ consist of all vectors in $\Phi(q,p)$ tangent to $\overrightarrow{az}$ with $z \in S(q,p)_j$. By Cor. 7.4 and Thm. 7.3, $\mathcal{R}(F)$ is null (and so has empty interior) unless F is a nonempty subset of $J(q,p)$ for some $(q,p) \in A$. Thus Prop. 9.2 applies with $X = \mathcal{R}_0,\ J = \cup J(p),\ X_j = \mathcal{R}_j \cap \mathcal{R}_0,\ j \in J$, and $\mathcal{F} = \cup \mathcal{F}(p) = \cup \mathcal{F}(q,p)$. We let $M(F) =$ Core $\mathcal{R}(F)$.

Theorem 9.3. *Suppose $B(p),\ p \in Q$, is a horoball packing, that L is the support of μ, and that $S(q,p),\ (q,p) \in A$, is a chain-finite shadow family. As q varies over cusp points* *with $(q,p) \in A$ for a fixed $p \in Q$, the tiles $Z(F)$, $F \in \cup\, \mathcal{F}(q,p)$, form a locally finite cover of $L - \{p\}$. $M(F)$, $F \in \mathcal{F}$, is a point finite cover of $\mathcal{U}$ by proper subsets of $\mathcal{R}$. The $M(F)$'s have null boundaries and can only meet at points in their boundaries. This cover is locally finite at points of $\mathcal{R}_0$. $M(F)$ consists of all $v \in \Phi(q,p)$ tangent to $\overrightarrow{az}$ for some $(z,a) \in K(F) = Z(F) \times A(F)$.*

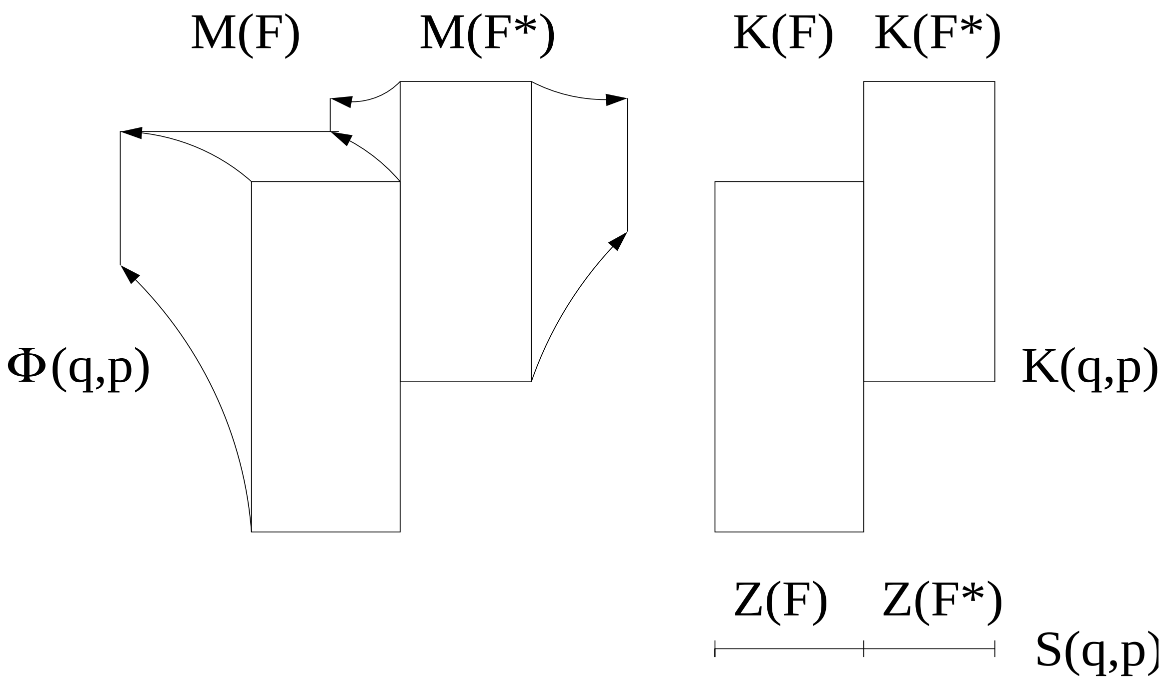

Figure 4. Markov boxes within a flowbox

Prop. 9.2 implies the tiles cover $L - \{p\}$ as stated. For $F \in \mathcal{F}(q,p)$ and v tangent to $\overrightarrow{az}$ we have $v \in \mathcal{R}(F)$ if and only if $v \in \Phi(q,p)$ and $z \in \mathcal{L}(F)$. Moreover v is an interior point of $\mathcal{R}(F)$ if and only if z is an interior point of $\mathcal{L}(F)$, a is an interior point of $A(F)$, and v is an interior point of $I(z,q,p,a)$. The last statement of the theorem follows.

Prop. 9.2 implies that the sets $M(F) \cap \mathcal{R}_0$ form a locally finite cover of $\mathcal{R}_0$. Thm. 7.3 shows that $\mathcal{R} - \mathcal{R}_0$ is closed, null, and has empty interior in $\mathcal{R}$. From this we see that the $M(F)$ form a point finite cover of $\mathcal{U}$ by proper sets. By Prop. 9.2, the sets $M(F) \cap \mathcal{R}_0$ have null boundaries and disjoint interiors in $\mathcal{R}_0$ and it follows that the same holds for the $M(F)$ as subsets of $\mathcal{R}$. This proves the theorem.

Taking into account Cor. 10.6 below, we see that the flowboxes $M(F)$ meet much like the flowboxes for a Markov partition for a basic set of an Axiom A flow [F1]. Accordingly we call the $M(F)$ *Markov boxes* and the family $M(F)$, $F \in \mathcal{F}$, a *Markov partition* for the invariant set $\mathcal{R}$. In Figure 4 we show the Markov boxes in a flowbox $\Phi(q,p) \subset \mathcal{R}$ with $\mathcal{F}(q,p) = \{F, F^*\}$ and the corresponding rectangles in $K(q,p) \subset \Omega$ and tiles in $S(q,p) \subset L$.

We now prove that the $A(F)$ depend on F in an order-preserving fashion.

Lemma 9.4. *If F and F' are distinct elements of $\mathcal{F}(q,p)$ with $F' \subset F \subset J(q,p)$ then $A(F')$ is a proper subset of $A(F)$.*

To prove this we take $z \in Z(F)$ and $z' \in Z(F')$ to be generic, so they do not belong to any chain boundaries. Since $F' \subset F$ we have $z \in S(z')$ and so by Lemma 5.4 we have $K_{z'}(q,p) \subset K_z(q,p)$. Thus $A(F') \subset A(F)$.

Suppose $j \in F - F'$ and so $z \in X(p)_j$ while $z' \notin X(p)_j$. For some $r \in Q$ we have $X(p)_j = S(q,p,...,r)$. By Continuation, $r \in K_z(q,p) = A(F)$. If $r \in K_{z'}(q,p) = A(F')$ then there is a sequence aimed at z' starting at r with successive terms p and q. Thus $z' \in S(q,p,...,r)' = X(p)_{j'}$ for some $j' \in J(q,p)$. But this implies $j' \in F'$. As $F' \subset F$ we find that $j' \in F$ and so $z \in X(p)_{j'}$. But z is generic and so there is only one sequence aimed at z starting at r, up to a shift. It follows that $j = j'$. Since $j' \in F'$ and $j \notin F'$, this is a contradiction. Thus $r \notin A(F')$ and so $A(F') \neq A(F)$, as desired.

We do not know whether the converse to Lemma 9.4 is false, but it seems likely that one may have $A(F) = A(F')$ in instances where and neither F nor F' is a subset of the other. This is suggested by Figure 2, where there are equal slices in adjacent shadows.

We now apply Lemma 9.4 in the context of §6.

Corollary 9.5. *When $n = 2$ and the shadow family is standard, each tile $Z(F)$, $F \in \mathcal{F}(q,p)$, is an interval in L.*

Choose an upper halfplane model with $p = \infty$, so $S(q,p) = \mathcal{I}(q,p) \cap L \subset \mathbf{R}$, and suppose $Z(F)$ is not an interval in L. There are then generic points $z_1, z_2 \in Z(F)$ and $z' \in S(q,p) - Z(F)$ with $z_1 < z' < z_2$. There is an $F' \in \mathcal{F}(q,p)$, $F' \neq F$, so that $z' \in Z(F')$.

Say $j \in F$. Then $Z(F) \subset X(p)_j$ and Lemma 6.5 implies that $X(p)_j$ is the intersection of L with a subinterval of $\mathcal{I}(q,p)$. As $X(p)_j$ contains z_1 and z_2 it contains all limit points in between. As z' is generic and is between z_1 and z_2, we see that $j \in F'$. Thus F is a subset of F'. By Lemma 9.4, $A(F)$ is a proper subset of $A(F')$.

Suppose $a \in A(F')$. We have $z' \in K^a(p)$ and $p \in K^a(p)$ and so by Lemma 6.7 we have $z_i \in K^a(p)$ where $i = 1$ if $a > z'$ and $i = 2$ if $a < z'$. As z_i is in $S(q,p)$ we find that $(z_i, a) \in K(q,p)$. As $z_i \in Z(F)$ is generic we see that $a \in A(F)$. Thus $A(F') \subset A(F)$. This is a contradiction, which proves the corollary.

We now give two formulas for the core of $K(q)$ as a subset of Ω for any cusp point q.

Proposition 9.6. *Core $K(q) = \cup\ K(F') \cup \{(z,q) : z \in L, z \neq q\}$ where $F' \in \mathcal{F}(q,p)$ for some $(q,p) \in A$. Also Core $K(q) = \cup\ K(F) \cup \{(q,a) : a \in L, a \neq q\}$ where $F \in \mathcal{F}(r,q)$ for some $(r,q) \in A$. In each case the intersections of any two terms in the union is null and nowhere dense.*

We fix q and prove the first equation in this lemma. Let $\mathcal{U}(q) = \cup$ Int $K(F')$ and suppose $(z,a) \in$ Int $K(q)$, $a \neq q$. Then by Prop. 5.2, there is some $p \in Q$ such that $(q,p) \in A$ and $(z,a) \in K(q,p)$ but only finitely many such p. Indeed any small compact neighborhood of (z,a) in Ω meets $K(q,p)$ for only finitely many p and hence meets $K(F')$ for finitely many F'. By the Baire category theorem, this neighborhood contains a nonempty open subset of one of these $K(F')$ and so it meets $\mathcal{U}(q)$. Thus $\mathcal{U}(q)$ is a dense open set in $\{(z,a) \in \text{Int } K(q) : a \neq q\}$. As each $\mathcal{F}(q,p)$ is finite and the family $K(q,p)$ is locally finite by Prop. 5.2, we can take the closure of $\mathcal{U}(q)$ term by term. But $K(F') = Z(F') \times A(F')$ and both factors are proper, so $K(F')$ is proper. Thus the union of the $K(F')$ is (Core

$K(q))- \{(z,q) : z \in L, z \neq q\}$. By Prop. 5.1, $(z,q) \in \text{Int } K(q)$ for $z \in L - \{q\}$, so the first equation follows.

The second equation can be proven in a similar fashion. The final assertion follows from Prop. 5.2 and Thm. 9.3, so the proposition is proved.

When $L = \partial H^n$, there is another way to form a Markov partition from a chain-finite family that avoids the use of the inclusion/exclusion method. One can remove from $S(q,p)$ all the chain boundaries Bd $S(q,p,...)$ to get an open subset of ∂H^n, take its connected components, and form their closures. One obtains a refinement of our tile family $Z(F), F \in \mathcal{F}(q,p)$, by finitely many proper connected sets. This is the approach taken in [F3].

Section 10. Symbolic dynamics

We now derive the "Markov property" for our Markov partition $M(F),\ F \in \mathcal{F}$.

Lemma 10.1. *Suppose $F \in \mathcal{F}(r,q)$ and $F' \in \mathcal{F}(q,p)$. If $Z(F) \cap Z(F')$ is not null then $Z(F) \subset Z(F')$ and $A(F') \subset A(F)$.*

Since the union of the chain boundaries Bd $S(r,q,...)$ is null, we may select a $z_0 \in Z(F) \cap Z(F')$ not belonging to such a chain boundary. Suppose that $z_1 \in Z(F)$ also lies outside this union, let i be 0 or 1, and fix some chain $S(q,p,...)$. Then the following are equivalent: $z_i \in S(q,p,...)$, $z_i \in S(r,q,p,...)$, $z_{1-i} \in S(r,q,p,...)$, $z_{1-i} \in S(q,p,...)$, where the second equivalence uses that $z_0,\ z_1$ are both in $X(q)(F)$. It follows that for some $F^* \in \mathcal{F}(q,p)$ we have that $z_0,\ z_1$ are both in $X(p)(F^*)$. As we vary z_0, we see that a set of full measure in $Z(F) \cap Z(F')$ is contained in $Z(F^*)$, so $F^* = F'$. If z_1 lies outside the union of the chain boundaries Bd $S(q,p,...)$ then $z \in \text{Int } X(p)(F') \subset Z(F')$. Thus $Z(F) - \cup$ Bd $S(q,p,...) \subset Z(F')$. As $Z(F)$ is proper and the chain boundaries have no interior, we see that $Z(F')$ contains a dense subset of $Z(F)$. As $Z(F')$ is closed, we see that $Z(F) \subset Z(F')$.

Next we take $z \in Z(F)$ not in any Bd $S(q,p,...)$ or Bd $S(r,q,...)$. We show that the slice $K_z(r,q) = A(F)$ contains the slice $K_z(q,p) = A(F')$. If $(z,a) \in K(q,p)$ then some aimed sequence p_i has $p_0 = q,\ p_{-1} = p,\ \lim^+ p_i = z$, and $\lim^- p_i = a$. As $z \in Z(F') \subset S(r,q)$, there is an aimed sequence p'_i with $p'_0 = q,\ p'_1 = r$, and $\lim^+ p'_i = z$. Splicing these sequences together gives an aimed sequence q_i with $q_i = p'_i$ for $i \geq 0$ and $q_i = p_i$ for $i \leq 0$. As $\lim^+ q_i = z$ and $\lim^- q_i = a$, we see that $(z,a) \in K(r,q)$, as desired, and the lemma is proven.

This Markov property leads in a standard way to a symbolic dynamics for the geodesic flow on $\mathcal{R}$, using the Markov boxes $M(F)$ in place of the flowboxes $\Phi(q,p)$ of §8. We define a *Markov transition* on $\mathcal{F}$ to be a pair (F',F) as in Lemma 10.1, so $F \in \mathcal{F}(r,q)$, $F' \in \mathcal{F}(q,p)$, and $Z(F) \subset Z(F')$. Let I be an interval of integers. A sequence $F_i \in \mathcal{F},\ i \in I$, is *allowed* if each pair (F_i, F_{i+1}) with $i \in I,\ i < \sup(I)$, is a Markov transition. Given such an allowed sequence, we see that there is an aimed sequence $p_i,\ i \in I$ such that $F_i \in \mathcal{F}(p_i, p_{i-1})$ for each $i \in I, i > \inf(I)$.

We let the *Markov correspondence* be the inclusion mapping $\mathcal{M} \to \mathcal{F} \times \mathcal{F}$, where $\mathcal{M}$ is the set of all Markov transitions and $\mathcal{M}$ and $\mathcal{F}$ are given the discrete topology. Then the inverse limit $\widehat{\mathcal{M}}$ is just the space of all allowed sequences with $I = \mathbf{Z}$, topologized as a subset of the product space $\mathcal{F}^{\mathbf{Z}}$. We define, by analogy with Section 8, $V : \widehat{\mathcal{M}} \to \mathcal{R}$ by $V(F_i,\ i \in \mathbf{Z}) = v(z, p_0, a)$ where (z,a) is the limit pair for this aimed sequence p_i.

The following theorem is analogous to Prop. 8.1, with aimed sequences replaced by allowed sequences throughout. This time, however, one can control the number of symbol sequences, up to shift, that represent a geodesic.

Theorem 10.2. *When Q is null, V is an orbit-equivalence from $\mathcal{M}$ to ϕ_t. A flowline with limit pair (z, a) meets the image of V if and only if z and a are not cusp points. If $v = V(F_i,\ i \in \mathbf{Z})$ then $\phi_t(v)$, $t \in \mathbf{R}$, is the unique flowline that lies in $\cup M(F_i), i \in \mathbf{Z}$. When the number of chains $S(q, p, \ldots)$ for $(q, p) \in A$ is bounded, there is a bound on the number of orbits of σ corresponding to a given geodesic.*

The shift orbit of $(F_i,\ i \in \mathbf{Z})$ maps by V to the set of all $v(z, p_i, a)$, $i \in \mathbf{Z}$, so shift orbits map to flowlines. Uniqueness and Thm 9.3 show that almost every flowline is covered by a single shift orbit on $\widehat{\mathcal{M}}$. Thus V is an orbit-equivalence. The image of V follows from Convergence and Existence. The flowline segment $I(z, p_i, p_{i-1}, a)$ lies in $\Phi(p_i, p_{i-1})$ and $z \in Z(F_i)$ so this segment lies in $M(F_i)$. Taking the union over i, we see that the flowline lies in the union of these Markov boxes. When a flowline lies in $\cup\Phi(p_i, p_{i-1})$, it is asymptotic to (z, a) with $(z, a) \in \cap K(p_i, p_{i-1})$. Then $z \in S(p_i, p_{i-1})$ for all $i \in \mathbf{Z}$ and so p_i is aimed at z by Convergence. Using (3.6) we see that $\cap K_z(p_i, p_{i-1})$ consists of a single point, so $a = \lim^- p_i$ and the third statement holds.

Now suppose there are at most N chains $S(q, p, \ldots)$ for each $(q, p) \in A$. Then the cardinality of $\mathcal{F}(q, p)$ is at most 2^N. Thm. 7.3 bounds the number of flowboxes $\Phi(q, p)$ in a neighborhood of a point of $\mathcal{R}_0$. As each of these contains at most 2^N Markov boxes, there is an integer M such that no more than M Markov boxes contain a point of $\mathcal{R}_0$. We will show that, up to a shift, no more than M^2 aimed sequences describe a given geodesic, which will imply the last statement of the theorem.

We follow a method used for compact hyperbolic systems (specifically the basic sets of Axiom A diffeomorphisms) by Bowen and Marcus, [B, p.14]. Choose an allowed sequence $F_m, \ldots, F_{m'}$ and points $z' \in \operatorname{Int} Z(F_{m'})$ and $a \in \operatorname{Int} A(F_m)$. By Lemma 10.1, (z', a) is an interior point of $K(F_j)$ for $m \leq j \leq m'$. Consider the orbit of the geodesic flow with limits (z', a) and the segment of this orbit from the point where it exits $M(F_m)$ to the point where it enters $M(F_{m'})$. This segment lies in the interior of $\cup\ M(F_j)$, $m \leq j \leq m'$, and so it does not meet any $M(F)$ for other $F \in \mathcal{F}$. Thus F_m and $F_{m'}$ determine all the F_j with $m \leq j \leq m'$.

Fix a flowline $\phi_t(v)$ and consider a pair of Markov boxes to which $\phi_s(v)$ and $\phi_{-s}(v)$ belong for some large positive s. There are at most M^2 such pairs. By the preceding paragraph, there is at most one way to extend this pair to a sequence of Markov boxes that covers the segment $\phi_t(v)$, $|t| \leq s$, in which consecutive boxes are related by a Markov transition. We let s tend to ∞ and see that at most M^2 allowed sequences correspond to this flowline, up to a shift. Thus the theorem is proved.

We now suppose that Γ is a symmetry group, so that Γ acts on $\mathcal{F}$, $\mathcal{M}$, and $\widehat{\mathcal{M}}$. Since V is equivariant, we obtain a mapping of quotient spaces $V/\Gamma : \widehat{\mathcal{M}}/\Gamma \to \mathcal{R}/\Gamma$. We define a positive function $T : \widehat{\mathcal{M}}/\Gamma \to \mathbf{R}$ by $T(x) = \tau(z, p_1, p_0)$, where $x = [F_i,\ i \in \mathbf{Z}]$, $F_i \in \mathcal{F}(p_i, p_{i-1})$ is an allowed sequence, and $z = \lim^+ p_i$. We can then build a *mapping torus* $(\widehat{\mathcal{M}}/\Gamma \times \mathbf{R})/\sim$ where the equivalence relation $\sim$ is generated by $(x, t+T(x)) \sim (\sigma(x), t)$, $x \in \widehat{\mathcal{M}/\Gamma}$, $t \in \mathbf{R}$, and where $\sigma(x) = [F_{i+1},\ i \in \mathbf{Z}]$. Thus σ is the homeomorphism on $\widehat{\mathcal{M}}/\Gamma$ induced by the shift on $\widehat{\mathcal{M}}$. The *suspension flow* ψ_t on the mapping torus is induced by translation on the factor $\mathbf{R}$, so $\psi_t([x,s]) = [x, s+t]$.

For $\tau = \tau(z, p_1, p_0)$, we have $\phi_{t+\tau}(v(z,p_0,a)) = v(z,p_1,a)$. It follows that the natural map $\widehat{\mathcal{M}}/\Gamma \times \mathbf{R} \to \mathcal{R}/\Gamma$ sending (x,t) to $\overline{\phi}_t((V/\Gamma)(x))$ induces a mapping Π_Γ from the mapping torus to $\mathcal{R}/\Gamma$. It is easy to see that $\overline{\phi}_t \circ \Pi_\Gamma = \Pi_\Gamma \circ \psi_t$, so Π_Γ is a semiconjugacy. Using Thm. 10.2, we obtain

Theorem 10.3. *There is a semiconjugacy Π_Γ from the suspension flow ψ_t on $(\widehat{\mathcal{M}}/\Gamma \times \mathbf{R})/\sim$ to the geodesic flow $\overline{\phi}_t$ on $\mathcal{R}/\Gamma$. The image of Π_Γ consists of all orbits which, when lifted to $\mathcal{R}$, are not asymptotic to a cusp point. Π_Γ is one-to-one except for a meager set in $\mathcal{R}/\Gamma$. This set is null whenever Q is null. When there is a bound on the number of chains $S(q,p,...)$ for $(q,p) \in A$, the level sets of Π_Γ have bounded cardinality.*

The analogous bound for compact hyperbolic systems plays a role in the enumeration of closed orbits, (see [B1], pp. 14-15, or [F1], §1) and the same is true here. Indeed the bound in Thm. 10.3 implies

(10.4) an orbit of σ is periodic if and only if the corresponding orbit of $\overline{\phi}_t$ is periodic.

To prove this, we first make a definition. We say that an allowed sequence F_i, $i \in \mathbf{Z}$, is Γ-*periodic* if there is a $\gamma \in \Gamma$ and a positive integer N such that $F_{i+N} = \gamma(F_i)$ for all integers i. Clearly such an allowed sequence determines a periodic orbit of $\overline{\phi}_t$ and hence a closed geodesic on H^n/Γ. Conversely, if the level sets of Π_Γ are bounded then each allowed sequence F_i, $i \in \mathbf{Z}$, that determines a periodic orbit of $\overline{\phi}_t$ must be Γ-periodic. Say $\overline{\phi}_t[v] = v$, $t > 0$, so $\phi_t(v) = \delta(v)$, for some hyperbolic element $\delta \in \Gamma$. The allowed sequences $\delta^m(F_i)$ can take only N values, up to a shift, if N is the cardinality of the level set $\Pi_\Gamma^{-1}[v]$. Thus we can take $\gamma = \delta^m$ for some integer m between 1 and N and (10.4) follows.

We now examine when the homeomorphism σ is a shift in its own right. The *reduced Markov correspondence* is the mapping of discrete spaces

$$\mathcal{M}/\Gamma \to \mathcal{F}/\Gamma \times \mathcal{F}/\Gamma \tag{10.5}$$

induced by the Markov correspondence. It is the analogue of the directed graph one uses to describe the transitions in a Markov partition for a compact hyperbolic system [B, F1]. If $\mathcal{F}$ is Q-free for Γ then, as in Prop. 8.4, the natural map $\widehat{\mathcal{M}}/\Gamma \to \widehat{\mathcal{M}/\Gamma}$ is a homeomorphism and so σ may be identified with the shift homeomorphism on $\widehat{\mathcal{M}/\Gamma}$. In this case, the reduced Markov correspondence is the analogue of the directed graph one uses to describe the transitions in a Markov partition for a compact hyperbolic system [B, F1]. Its inverse limit $\widehat{\mathcal{M}/\Gamma}$ is the analogue of the subshift of finite type used to describe a compact hyperbolic system. Indeed Thm. 10.3 implies that the reduced Markov correspondence is orbit-equivalent to the reduced geodesic flow, just as in the compact case. When $\mathcal{F}/\Gamma$ is finite, which is true for the geometrically finite groups treated in §11, this shift is expressed with a finite symbol set and a countable set of transitions.

When $\mathcal{F}$ is Q-free for Γ, the Markov partition $M(F)$, $F \in \mathcal{F}$, for ϕ_t can be pushed forward to a family of flowboxes $[M(F)]$ for $\overline{\phi_t}$ that cover $\mathcal{R}/\Gamma$ without overlap. The reduced Markov correspondence governs the dynamics of $\overline{\phi_t}$ in the sense that it determines the orbits that enter infinitely many flowboxes in forward and backward time. We will call these flowboxes for $\overline{\phi_t}$ the *reduced Markov partition*. Even when $\mathcal{F}$ is not Q-free, one can often subdivide the Markov boxes as in Appendix 3 to define a Markov partition for $\overline{\phi_t}$.

Prop. 9.6 and Lemma 10.1 give relations between the slices $A(F)$ that are dual to relations between the tiles $Z(F)$.

Corollary 10.6. *For $(r, q) \in A$ and $F \in \mathcal{F}(r, q)$ one has $A(F) - \{q\} = \cup\, A(F')$ where $F' \in \mathcal{F}$ and (F', F) is a Markov transition. For $(q, p) \in A$ and $F' \in \mathcal{F}(q, p)$ one has $Z(F') - \{q\} = \cup\, Z(F)$, where $F \in \mathcal{F}$ and (F', F) is a Markov transition. In either of these unions, the intersection of any two terms is null and nowhere dense.*

Consider the Ford disc packing of H^2 with $q = \infty$. Then Prop. 9.6 expresses the finite points in $K(\infty)$ as a union of two systems of infinite rectangles, the vertical rectangles $K(F)$ bounded by lines $z = m$ and $z = m \pm 1/2$ and the horizontal rectangles $K(F')$. We choose $F \in \mathcal{F}(0, \infty)$ so that $Z(F) = [0, 1/2]$ and $A(F) = [3 - r, r - 2]$, where here $r = (3 - \sqrt{5})/2$ as in §5. For $F' \in \mathcal{F}(\infty, m)$ we have $Z(F') = [m + 2, \infty]$ or $[\infty, m - 2]$ and $A(F') = [m + r - 1, m + r]$ or $[m - r, m + 1 - r]$, respectively. Taking $m \leq -2$ or $m \geq 3$, respectively, gives $Z(F) \subset Z(F')$ and $A(F') \subset A(F)$, otherwise $Z(F) \cap Z(F')$ contains at most one point. Thus we have identified all the Markov transitions (F', F). Notice that these $A(F')$ do not overlap and their union is $(\infty, r - 2] \cup [3 - r, \infty)$, that is $A(F) - \infty$. This is consistent with Cor. 10.6.

Suppose Γ is a symmetry group of a horoball packing and shadow family as in §8. Then the image under the face-pairing γ_q of a tile $Z(F)$, $F \in \mathcal{F}(q, p)$, meets $L - \{p'\}$, $p' = \gamma_q(q)$, in a union of tiles. This is immediate from Cor. 10.6 since $\gamma_q(Z(F)) = Z(F')$ for $F' = \gamma_q(F) \in \mathcal{F}$.

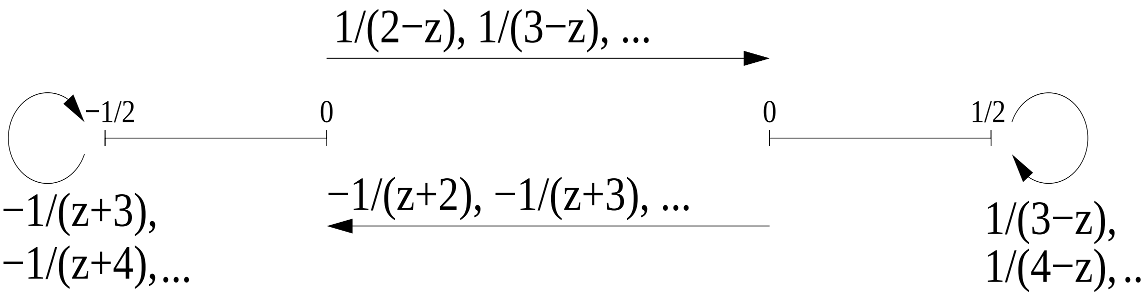

Figure 5. Interval maps for the modular group and Ford disc packing

For the Ford disc packing and its full symmetry group $\Gamma = PGL(2, \mathbf{Z})$ this is evident since the image of $[0, 1/2]$ by $1/x$ is $[2, \infty]$ and $[2, \infty)$ is a union of tiles $[m, m + 1/2]$, $m \geq 2$, and $[m - 1/2, m]$, $m \geq 3$. As Γ acts transitively on $\mathcal{F}$, the shift space $\widehat{\mathcal{M}/\Gamma}$ in this case has one symbol and a transition for each $m \leq -2$ or $m \geq 3$. The modular group $\Gamma^+ = PSl(2, \mathbf{Z})$ has index two in Γ and its shift space $\widehat{\mathcal{M}/\Gamma^+}$ can be described by two symbols, represented in Figure 5 by the tiles $Z(F)$ for $F \in \mathcal{F}(0, \infty)$. Each Markov transition (F', F) for $F \in \mathcal{F}(0, \infty)$ determines an arrow in this figure, representing the contraction mapping obtained by following the inclusion $Z(F) \subset Z(F')$, $F' \in \mathcal{F}(\infty, m)$, with the mapping $-1/(z - m) \in \Gamma^+$ that takes (∞, m) to $(0, \infty)$. It should be noted that the images of these maps fill the intervals without overlapping, as required by the second statement in Cor. 10.6. In the language of [F2], Figure 5 represents a *Markov system of interval maps*. If we identify the two points labeled 0 in Figure 5 and also identify $-1/2$ with $1/2$, we obtain a contracting correspondence on the circle that is the inverse of the reduced correspondence $\mathcal{A}/\Gamma^+$.

Section 11. Geometrically finite groups

A finitely generated discrete subgroup Γ of isometries of H^2 gives a quotient orbifold H^2/Γ with finitely many ends [FN]. A parabolic fixed point p determines a finite area neighborhood $B(p)/\Gamma_p$ of an end, where $B(p)$ is a sufficiently small horoball based at p. A component interval J of $\partial H^2 - L(\Gamma)$ with endpoints a, z and stabilizer Γ_J determines a halfplane X such that $\partial X = \overline{az}$ and X/Γ_J is an infinite area neighborhood of an end. All ends arise in one of these two ways, and the respective neighborhoods may be called *cusps* and *funnels.*

In order to control the ends of H^n/Γ for any $n \geq 2$, we assume Γ is geometrically finite. This condition can be defined in terms of fundamental domains, as in [Ra] (for $n \leq 3$ it means that Dirichlet domains are finite-sided), or by properties of the limit points, as below. These definitions are equivalent by [Ra], Thm. 12.3.5. Let $C(\Gamma) \subset H^n$ be the convex hull of the limit set. The *Nielsen kernel* or *convex core* $K(\Gamma) = C(\Gamma)/\Gamma$ is a deformation retract of H^n/Γ by [Ra] §12.1 ($K(\Gamma)$ is the complement of the interior of the funnels when $n = 2$). $K(\Gamma)$ has finitely many ends and each corresponds to a parabolic fixed point p, determined up to the action of Γ. For a sufficiently small horoball $B(p)$ based at p, $(B(p) \cap C(\Gamma))/\Gamma_p$ is a neighborhood of the end. This description of ends holds for a torsionfree Γ by [Ra], Thm. 12.6.6. Since Γ is finitely generated ([Ra], Cor. 5 of §7.5.), we can find a torsionfree subgroup of finite index by Selberg's Lemma ([Ra], Thm. 12.3.9) and derive the general case.

When $K(\Gamma)$ is noncompact, we choose such a horoball $B(p)$ for each end so that these neighborhoods are nonoverlapping in $K(\Gamma)$ (see [Ra] §12.6, Lemma 7 for a more general construction). Setting $B(\gamma(p)) = \gamma(B(p))$, $\gamma \in \Gamma$, gives an invariant horoball packing $B(q)$, $q \in Q(\Gamma)$, of H^n, with $Q(\Gamma)$ the set of parabolic fixed points (c.f. [S2]). This packing is crucial for our symbolic dynamics.

With this packing as motivation, we present some rather technical definitions. Let Γ be a nonelementary discrete group of isometries of H^n. A point $z \in L(\Gamma)$ is *conical* if for any two points w, $w' \in H^n$, there is a positive constant $\delta = \delta(w, w')$ so that infinitely many points $\gamma(w')$, $\gamma \in \Gamma$, lie in the δ-neighborhood of the ray $\overrightarrow{wz}$. Note that it is enough to check this condition for one pair (w, w'). In the literature, such a limit point is also called *radial* or *a point of approximation.* A point $q \in L(\Gamma)$ is *cusped* if the stabilizer Γ_q acts properly discontinuously on $L(\Gamma) - \{q\}$ with compact quotient Λ_q. For an equivalent definition in terms of a *cusped neighborhood* see [Ra], especially Thm. 12.6.1. Conical and cusped limit points have opposite behavior, in that each orbit of Γ accumulates rapidly at a conical limit point and slowly at a cusped limit point. Γ is *geometrically finite* if every limit point of Γ is conical or cusped. A geometrically finite group with $L(\Gamma) = \partial H^n$ is just a discrete group Γ of hyperbolic isometries for which H^n/Γ has finite volume ([Ra], Thm. 12.6.2). Every parabolic fixed point for a geometrically finite group is cusped ([Ra] Thm. 12.5.3).

Suppose Γ is nonelementary and geometrically finite. Let $\mathcal{R}(\Gamma) \subset SH^n$ consist of all the tangent vectors to lines $\overline{az}$ with a and z limit points of Γ. Let the *recurrent part* of the geodesic flow over H^n/Γ be the closed invariant set $\mathcal{R}(\Gamma)/\Gamma$. The recurrent part is *isolated* in the dynamical sense that it is the maximal invariant set in some open subset $U \subset SH^n/\Gamma$. Indeed one can take U to be any bounded neighborhood of the recurrent part, since $\overline{az}$ lies in a bounded neighborhood of $C(\Gamma)$ only when a and z are limit points.

By our description of the ends of $K(\Gamma)$, $\mathcal{R}(\Gamma)/\Gamma$ is compact if and only if $K(\Gamma)$ is compact, in which case one says Γ is *convex cocompact.* Say Γ acts freely on SH^n, as can be arranged by passing to a torsionfree subgroup of finite index, so SH^n/Γ is a manifold. As the geodesic flow over this manifold is Anosov and as $\mathcal{R}(\Gamma)/\Gamma$ is a compact isolated invariant set, Bowen's theory of Markov partitions ([B4] or [F3], §4) defines a symbolic dynamics on $\mathcal{R}(\Gamma)/\Gamma$. Thus the convex cocompact case is well understood, at least up to a finite cover.

So assume also that Γ is not convex cocompact. Then $Q(\Gamma)$ is nonempty and, as the action of Γ on $L(\Gamma)$ is minimal, $L(\Gamma)$ is the closure of $Q(\Gamma)$. Let $Q = Q(\Gamma)$, so $L = L(\Gamma)$, $\mathcal{R} = \mathcal{R}(\Gamma)$, and Q/Γ is finite. We will soon need the following lemma.

Lemma 11.1. *When no 2-plane in H^n is fixed by a nontrivial element of Γ, Γ acts freely on $Q \times (L - Q)$.*

The proof depends on an evident property of a properly discontinuous group G of isometries of a locally compact metric space X: the fixed sets $\mathrm{Fix}(g) = \{x \in X : g(x) = x\}$, $g \in G$, form a locally finite family in X. In particular, the sets $g(\mathrm{Fix}(h)) = \mathrm{Fix}(ghg^{-1})$, $g \in G/C(h)$, form a locally finite family, where $C(h)$ is the centralizer of h in G. We apply this to the action of Γ on SH^n. Suppose a nontrivial element $h \in \Gamma$ fixes a pair $(p, z) \in Q \times L$, $z \neq p$. The $\mathrm{Fix}(h)$ consists of two flowlines $\{\phi_t(\pm v) : t \in \mathbf{R}\}$, where v is tangent to $\overrightarrow{pz}$. Indeed these flowlines are fixed because h preserves $B(p)$ and no other vectors are fixed because $\pi(\mathrm{Fix}(h)) \subset H^n$ is fixed by h and so has dimension at most one. As the stabilizer of each flowline is the finite group $\Gamma_p \cap \Gamma_z$, we see that the family of flowlines $\Phi(\gamma) = \gamma^{-1}(\{\phi_t(v) : t \in \mathbf{R}\})$, $\gamma \in \Gamma$, is locally finite.

Suppose z is conical. Let $w = w' = \pi(v)$ and let $K \subset H^n$ be the closed ball of radius $\delta(w, w)$ centered at w. There is a sequence γ_m of distinct elements of Γ such that for each m the distance from $\gamma_m(w)$ to $\overrightarrow{pz}$ is at most $\delta(w, w)$. So the compact set $\pi^{-1}(K)$ meets $\Phi(\gamma_m)$ for each m, which contradicts local finiteness. Thus z is not conical and so, as Γ is geometrically finite, $z \in Q$. The lemma follows.

The case $n = 2$ of Lemma 11.1 is a classical result about the fixed line of a reflection in a finitely generated group of isometries of H^2, see [FN], p.118. There is, incidentally, an example in §16 with $n = 3$ where Γ contains reflections but still acts freely on $Q \times (L - Q)$.

As above, Γ is a symmetry group for a horoball packing $B(q)$, $q \in Q$. Moreover, the quotient by Γ of $\pi(\mathcal{R})-$ Int $\cup B(q), q \in Q$, is compact. It follows that this horoball packing is D-dense for some $D > 0$. Fix a shadow family $S(q, p)$, $(q, p) \in A$, such as the family of ridge shadows. §7 constructs transversals $T(p)$, $p \in Q$, and flowboxes $\Phi(q, p)$, $(q, p) \in A$ for the geodesic flow ϕ_t on $\mathcal{R}$. A/Γ is finite since Q/Γ is finite and each Λ_p, $p \in Q$, is compact. Thus there are only finitely many flowboxes up to the action of Γ. The index sets Q' and Q_p of §8 are finite and the reduced aimed correspondence $\mathcal{A}/\Gamma$ is defined on a compact space $\mathcal{L}/\Gamma$.

The next theorem shows that the results of previous sections apply to $B(q)$, $q \in Q$. We verify that conditions (A) and (B) hold for the *geometric measure class* of Sullivan, which was first defined and studied for $n = 2$ by Patterson [P]. This measure class generalizes the Lebesgue measure class, to which it reduces when H^n/Γ has finite volume. We will describe the geometric measure class in terms of certain standard objects in global analysis known as *fractional densities* (see [GS]). These are equivalent to the *conformal densities* used by Sullivan but are defined without a Riemannian metric.

Consider an expression $f(x)|dx|^s$ where f is a continuous real-valued function on an open set in $\mathbf{R}^m$, $|dx|$ is Lebesgue measure, and $s \in \mathbf{R}$. If we change coordinates by a C^1 diffeomorphism $x = \psi(y)$, we formally obtain $f(x)|dx|^s = g(y)|dy|^s$ where $g(y) = f(\psi(y))J(y)^s$ and $J(y) = |\det D\psi(y)|$, so J and g are continuous. Now fix a C^1 manifold M and define a *continuous s-density* on M to be a family of such expressions, one in each coordinate chart, that transform as indicated when one changes charts. A continuous 0-density is a continuous function on M whereas a continuous 1-density is a measure on M given by a continuous density function. Each continuous s-density on M acts as a continuous linear functional on the Frechet space of continuous $(1-s)$-densities with compact support by the rule "multiply and integrate over M." The full dual space consists of s-densities whose coefficients are certain generalized functions, namely those s-densities with local expressions of the form $\mu|dx|^{s-1}$ where μ is a Borel measure. We will be concerned with the case when all the μ are nonnegative Borel measures, in which case we will speak of a *positive s-density* on M.

Given a Riemannian metric g on M with its smooth measure $|dvol(g)|$, each positive s-density has the form $\mu(g)|dvol(g)|^{s-1}$ where $\mu(g)$ is a positive Borel measure on M. One can easily check that $\mu(\sigma^2 g) = \sigma^\delta \mu(g)$ for each positive scale function σ, where $\delta = m(1-s)$, $s = 1 - \delta/m$. This is the transformation law for a conformal density of dimension δ, as in [S1]. The natural dimension bound $0 \leq \delta \leq m$ translates into the condition $0 \leq s \leq 1$.

Let δ be the Hausdorff dimension of L, $m = n-1$, and $s = 1 - \delta/m$. Sullivan constructs a positive s-density on ∂H^n [S1, S2]. We may summarize this construction and several of Sullivan's results as follows. In the ball model of H^n, the asymptotic distribution of any Γ-orbit as viewed from the origin defines a certain Borel probability measure on ∂H^n whose support is L. The *geometric measure class* is the class of this measure. Coupled with the spherical metric on ∂H^n, this probability measure defines a positive s-density that is invariant by Γ.

Theorem 11.2. *Given a geometrically finite group Γ that is not convex cocompact, its geometric measure class satisfies (A) and (B) for any Γ-invariant horoball packing $B(q)$, $q \in Q = Q(\Gamma)$. Also Q is null. Let $S(q,p)$, $(q,p) \in A$, be a Γ-invariant shadow family, $\mathcal{A}/\Gamma$ its reduced aimed correspondence, and $\underline{\mathcal{A}}/\Gamma$ the reduced marked correspondence of Prop. A3.1. Let $\overline{\phi}_t$ be the reduced geodesic flow on $\mathcal{R}/\Gamma$.*
(a) $\underline{\mathcal{A}}/\Gamma$ is defined on a compact space $\underline{\mathcal{L}}/\Gamma$ and is orbit-equivalent to $\overline{\phi}_t$.
(b) $\mathcal{A}/\Gamma$ is defined on a compact space $\mathcal{L}/\Gamma$ and is orbit-equivalent to $\overline{\phi}_t$ when $Q \times (L - Q)$ is Q-free for Γ. This holds when the cusp stabilizers are torsionfree, when $n = 2$, or when Γ preserves orientation and $n = 3$.
(c) $\mathcal{A}/\Gamma$ has a finite invariant ergodic Borel measure $\overline{\nu}$ with support $\mathcal{L}/\Gamma$.

The proof of (A) uses an ergodic invariant measure for $\overline{\phi}_t$ found by Sullivan. Let $\mu|dx|^{-\delta/m}$ be the local expression for his positive s-density in an upper halfspace model of H^n. Let $\gamma \in \Gamma$ and $x' = \gamma(x) \in \mathbf{R}^m$, so $|dx'| = J(x)\gamma_*|dx|$, where $J(x) = |\det D\gamma(x)|$. As $\gamma_*(\mu|dx|^{-\delta/m}) = \mu|dx'|^{-\delta/m}$ by invariance, $\gamma_*(\mu) = J^{-\delta/m}\mu$. Now for $y' = \gamma(y)$ we have $|x'-y'|^2 = |x-y|^2(J(x)J(y))^{1/m}$, as is easily checked by expressing γ as a composition of similarity transformations and inversion in the unit sphere. Sullivan defines the Borel measure

$$\nu = (\mu \times \mu)/|x-y|^{2\delta} \tag{11.3}$$

on the space Λ of ordered pairs (x, y) of distinct points in ∂H^n. By the preceding formulas, $\gamma_*(\nu) = \nu$, $\gamma \in \Gamma$, and ν is supported on Ω. Since Ω is the orbit space of ϕ_t, ν defines a measure on $\mathcal{R}$ invariant by Γ and by ϕ_t. This descends to an invariant measure dm_μ for $\overline{\phi}_t$. Sullivan proves that dm_μ is finite and ergodic and has support equal to $\mathcal{R}/\Gamma$ so ν is ergodic for the action of Γ [S2]. If H^n/Γ has finite volume then $L = \partial H^n$, $\mu = |dx|$, dm_μ is Liouville measure on SH^n/Γ, and the ergodicity of dm_μ is a classical result of Hopf.

Let $\mathcal{O}_\rho \subset \cup B(q)$, $q \in Q$, be as in §1 and let $U_\rho = (\pi^{-1}\mathcal{O}_\rho)/\Gamma$, so $[v] \in U_\rho$ if some disc of radius $r > \rho$ centered at $\pi(v)$ lies in some $B(q)$. As the support of dm_μ is $\mathcal{R}/\Gamma$, and as U_ρ is open and nonempty, $dm_\mu(U_\rho) > 0$. By ergodicity, the set of all $[v] \in \mathcal{R}/\Gamma$ such that $\sup\{t : \overline{\phi}_t([v]) \in U_\rho\} = +\infty$ has full measure. (A) follows immediately.

Corollary A1.5 implies that (B) holds for the geometric measure class (but see also [Ru], Lemma 1). Q is null by [S2]. (a) and (b) now follow from Thm. 8.3, Lemma 11.1, and Prop. A3.1.

We now construct $\overline{\nu}$. Fix a cusp point q and note that Bd $K(q)$ is ν-null. It follows that we may restrict ν to $K(q)$. We project this measure forward to a measure ν_q on $L - \{q\}$ that is finite on compact sets. For $Z \subset L - \{q\}$, Prop. 5.2 implies $\nu_q(Z) = \Sigma\ \nu_p(Z \cap S(q,p))$, where p varies over the cusp points with $(q, p) \in A$. Thus $\mathcal{A}$ is measure-preserving for the measure $\nu_* = \coprod \nu_q$ on $\mathcal{L} = \coprod L - \{q\}$. Since Γ_q preserves $K(q)$ and since ν is Γ-invariant, ν_q is Γ_q-invariant and ν_* is Γ-invariant. Cor. A1.5 implies that the nontrivial fixed sets of elements of Γ_q are null. Thus there is a finite Borel measure $\overline{\nu}_q$ on $\Lambda_q = (L - \{q\})/\Gamma_q$ induced by ν_q. These measures $\overline{\nu}_q$ define a finite Borel measure $\overline{\nu}$ on $\mathcal{L}/\Gamma$. $\overline{\nu}$ is invariant by $\mathcal{A}/\Gamma$ since ν is Γ-invariant. Since ν is Γ-ergodic, $\overline{\nu}$ is ergodic for the reduced aimed correspondence. This proves (c) and the theorem is proved.

To illustrate (c), say $n = 2$ and $\Gamma = PSL(2, \mathbf{Z})$ with the Ford disc packing. One can take $\nu = (z - a)^{-2}|dz||da|$. Identifying $\mathcal{L}/\Gamma$ with $\mathbf{R}/\mathbf{Z}$, and using the endpoints in Figure 2, one finds $\overline{\nu} = \rho(z)|dz|$, where $\rho(z) = (|z| + 2 - r)^{-1} + (3 - r - |z|)^{-1}, |z| \leq 1/2$. $\overline{\nu}$ is a finite ergodic invariant measure for the transformation E of (0.1). It is the analogue for least-remainder continued fractions of the Gauss measure $(z + 1)^{-1}|dz|$ on $[0, 1]$ (c.f. [M], [Se], [F2]).

As A/Γ is finite, Thm. 10.2 and Thm. 10.3 imply

Corollary 11.4. *When the shadow family of Thm. 11.2 is chain-finite, the number of allowed sequences, up to shift, corresponding to a given $(z, a) \in \Omega$ is bounded. When Γ acts freely on $\mathcal{F}$, $\overline{\phi}_t$ has a Markov partition with finitely many Markov boxes and the reduced Markov correspondence $\mathcal{M}/\Gamma$ on the finite set $\mathcal{F}/\Gamma$ is orbit-equivalent to $\overline{\phi}_t$.*

These Markov partitions have an unexpected smoothness property when Γ is the symmetry group of a horoball packing, H^n/Γ has finite volume, and we use the family of ridge shadows. Each tile is then bounded by finitely many conformal spheres, so its boundary is piecewise smooth. This contrasts sharply with the situation for hyperbolic automorphisms of the 3-torus, say, where Bowen showed that the two-dimensional factor in a Markov partition cannot contain a smooth arc [B2]. Bowen's result was extended to higher dimensions by

Cawley [Ca] and one typically expects any Markov factor of dimension two or more to have fractal boundary (an exception must be made in such cases as the product of two automorphisms of T^2). In fact, for $n = 3$ the two-dimensional slices seem to have fractal boundary, as mentioned at the end of the next section.

Section 12. The slice transform

We assume in this section that Γ is a geometrically finite group, $Q = Q(\Gamma)$, $B(q),\ q \in Q$ is an invariant horoball packing, and $S(q,p),\ (q,p) \in A$, is an invariant chain-finite shadow family. Then up to the action of Γ there are only a finite number of shadows and chains and so we may treat the tiles $Z(F)$ as known. To find the Markov boxes $M(F)$, it only remains to describe the slices $A(F),\ F \in \mathcal{F}(r,q)$, as explicitly as possible. We will find the $A(F)$'s from their transformation properties as described in Cor. 10.6. That is, the formula $A(F) = \cup\ A(F')\ \cup \{q\}$, where $F \in \mathcal{F}(r,q)$ and (F', F) is a Markov transition, is a fixed point equation for the family of slices $A(F)$ that we will solve by iteration.

Given a family of sets $B(F) \subset L$ indexed by $F \in \mathcal{F}$, define its *slice transform* to be the family of sets $(TB)(F)$ given by the rule

$$(TB)(F) = \cup\ B(F')\ \cup \{q\}, \quad \text{where } (F', F) \in \mathcal{M}. \tag{12.1}$$

More explicitly, $F \in \mathcal{F}(r,q)$, p, q, r is an aimed segment, and $F' \in \mathcal{F}(q,p)$ with $Z(F) \subset Z(F')$. By Cor. 10.6, the family $A(F)$ is a fixed point of this transform.

To illustrate the use of the slice transform, consider the Ford disc packing with $(r,q) = (0,\infty)$ and $Z(F) = [0, 1/2]$. We fix $x > 1/2$ and $y < 0$ and let $B(\gamma(F)) = \gamma([x,y])$ for all symmetries γ of this packing. Thus $B(F)$ is a compact interval, disjoint from $Z(F)$, and a neighborhood of ∞. The most natural choice for $B(F)$ might be the interval $[1,-1]$ on which the face pairing $\gamma_0(z) = 1/z$ is nonexpansive. The Markov transitions (F', F) were calculated after Cor. 10.6. $B(F')$ is $m + [y^{-1}, x^{-1}],\ m \le -2$, or $m + [-x^{-1}, -y^{-1}],\ m \ge 3$. We say $[x,y]$ is *big* if $x^{-1} - y^{-1} > 1$, in which case $(TB)(F) = [3 - x^{-1}, x^{-1} - 2]$. If $(TB)(F)$ is also big, that is if $x + x^{-1} < 3$, then $(T^n B)(F) = [x_n, y_n]$ is a big interval for all $n \ge 0$. Here $x_0 = x,\ y_0 = y$ and $x_{n+1} = 3 - x_n^{-1},\ y_{n+1} = x_n^{-1} - 2$. One finds $x_n \to 3 - r$, so $y_n \to r - 2$ and $(T^n B)(F)$ converges exponentially to $[3 - r, r - 2] = A(F)$.

We will use the following properties of the slice transform.

(12.2) If $A_0(F)$ is empty for each F and $m \ge 1$ then $(T^m A_0)(F)$ consists of all cusp points p for which there is an allowed sequence $F_j, \ldots, F_0 = F$ with $-m \le j \le 0,\ F_i \in \mathcal{F}(p_{1+i}, p_i)$ for $j \le i \le 0,\ p = p_j,\ q = p_0$, and $r = p_1$. For any family $B(F)$ and $m \ge 1$, $(T^m B)(F)$ is the union of $(T^m A_0)(F)$ and the sets $B(F_{-m})$ for which there is an allowed sequence $F_{-m}, \ldots, F_0 = F$.

(12.3) T is monotone, that is if $C(F) \subset B(F)$ for all F then $(TC)(F) \subset (TB)(F)$ for all F. For instance $A_0(F) \subset A(F) = (TA)(F)$ so $\cup\,(T^m A_0)(F) \subset A(F)$. So when $(TB)(F) \subset B(F)$ for all F, each sequence $(T^m B)(F)$ is decreasing. Likewise when $C(F) \subset (TC)(F)$ for all F, each sequence $(T^m C)(F)$ is increasing.

(12.4) $(TB)(F)$ is compact when each $B(F')$ in (12.1) is compact and the $B(F') - \{q\}$ are a locally finite family in $L - \{q\}$. For the union of this family is then closed in $L - \{q\}$ and $(TB)(F)$ is closed in the compact set L. When each $C(F') \subset B(F')$ is also compact, $(TC)(F)$ is compact.

(12.5) When $B(F)$ is disjoint from $Z(F)$ for all F, the same holds for $(TB)(F)$. For $B(F') \subset L - Z(F') \subset L - Z(F)$ for each F'. Also $q \in A(F)$ so $q \notin Z(F)$.

For $(r,q) \in A$, $F \in \mathcal{F}(r,q)$, and $\epsilon > 0$ we let $A^\epsilon(F) = \{a \in L : d_q(a, Z(F)) \geq \epsilon\}$.

Theorem 12.6. *For any Γ-invariant family of compact sets $B(F) \subset L - Z(F)$, $F \in \mathcal{F}$, the sequence $(T^m B)(F)$, $m \geq 0$, consists of compact sets whose Hausdorff distance to $A(F)$ is $O(c^m)$. For ϵ sufficiently small, $(T^m A^\epsilon)(F)$, $m \geq 0$, is a decreasing sequence of compact sets whose intersection is $A(F)$.*

We first select ϵ sufficiently small that

(12.7) $A(F) \cup B(F) \subset A^\epsilon(F)$ and

(12.8) $d_p(x,z) < d_q(x,z)$ when $z \in S(q,p)$ and $d_q(x,z) < \epsilon$.

This is easily done for a fixed F or a fixed (q,p). As $\mathcal{F}/\Gamma$ and A/Γ are finite, we can choose one ϵ that works for all F and all (q,p). We next show

(12.9) $(TA^\epsilon)(F) \subset A^\epsilon(F)$, for $F \in \mathcal{F}(r,q)$.

By (12.7), $q \in A^\epsilon(F)$. Say $F' \in \mathcal{F}(q,p)$, $Z(F) \subset Z(F')$, and $a \in A^\epsilon(F') - \{q\}$. Choose $z \in Z(F')$ so that $d_q(a,z) = d_q(a, Z(F'))$. Since $d_p(a,z) \geq d_p(a, Z(F')) \geq \epsilon$, (12.8) gives $d_q(a,z) \geq \epsilon$. Thus $d_q(a, Z(F)) \geq d_q(a, Z(F')) \geq \epsilon$ so $a \in A^\epsilon(F)$, as desired.

(12.3) implies that the sequence $(T^m A^\epsilon)(F)$ is decreasing, the sequence $(T^m A_0)(F)$ is increasing, and $(T^m A_0)(F) \subset (T^m A^\epsilon)(F)$ for $m \geq 0$. The following lemma shows that these sets are close. For $q \in Q$ and $z \in \partial H^n$, $z \neq q$, let $d_{z,q}$ be the metric on $\partial H^n - \{z\}$ given by an upper halfspace model with $z = \infty$ and $B(z,q) = B_n$.

Lemma 12.10. *For $F \in \mathcal{F}(r,q)$, $z \in Z(F)$, and $m \geq 0$, $(T^m A_0)(F)$ is c^m/ϵ-dense in $(T^m A^\epsilon)(F)$ in the metric $d_{z,q}$.*

Choose an upper halfspace model with $q = \infty$, $B(q) = B_n$, and $z = 0$ and let $\zeta \in (T^m A^\epsilon)(F) - (T^m A_0)(F)$. By (12.2), there is an allowed sequence $F_i \in \mathcal{F}(p_{i+1}, p_i)$, $-m \leq i \leq 0$, with $p_0 = q$, $p_1 = r$, $F_0 = F$, and $\zeta \in A^\epsilon(F_{-m})$. Setting $p = p_{-m}$, we have $0 = z \in Z(F_0) \subset Z(F_{-m})$ and so

$$\epsilon \leq d_p(\zeta, Z(F_{-m})) \leq d_p(\zeta, 0) = h(B(p))|\zeta|/(|p||\zeta - p|). \tag{12.11}$$

By Monotonicity, $B(z,p)$ contains the $m|\log c|$-neighborhood of $B(z,q)$ and so $B(p)$ is disjoint from the horoball of diameter c^{-m} based at 0. As in Prop. 2.4, $h(B(p))c^{-m} \leq |p|^2$. Combining this with (12.11) gives $d_{z,q}(\zeta, p) = |\zeta - p|/(|p||\zeta|) \leq h(B(p))/(\epsilon|p|^2) \leq c^m/\epsilon$. As $p \in (T^m A_0)(F)$ by (12.2), the lemma follows. We now show

(12.12) $(TA^\epsilon)(F)$ is compact.

By (12.4), it suffices to show that each compact set $D \subset L - \{q\}$ meets $A^\epsilon(F')$ for only finitely many F'. The sets F' lie in only finitely many orbits of Γ_q and so we may restrict our attention to sets $F' = \gamma F'_0$ where $\gamma \in \Gamma_q$ and $F'_0 \in \mathcal{F}(q, p_0)$. Choose an upper halfplane model with $q = \infty$. If ∞ is not in $Z(F'_0)$ then $Z(F) \subset Z(F')$ implies that $\gamma^{-1}(Z(F)) \subset Z(F'_0)$, which can only hold for finitely many γ. Assume, then, that $\infty \in Z(F'_0)$. Then $\gamma^{-1}D$, $\gamma \in \Gamma_\infty$, converges to ∞ so $d_{p_0}(\gamma^{-1}D, \infty) < \epsilon$ except for finitely many γ. Thus for $p = \gamma p_0$ we have $d_p(D, \infty) < \epsilon$ except for finitely many γ. So D meets $A^\epsilon(F')$ for only finitely many γ, and (12.12) follows.

By (12.12), (12.9), and (12.4), each set $(T^m A^\epsilon)(F)$, $m \geq 0$, is compact. It follows from (12.12), (12.7), and (12.4) that each $(T^m B)(F)$ is compact. By (12.3) and (12.7), $(T^m A_0)(F) \subset (T^m B)(F) \subset (T^m A^\epsilon)(F)$, and the corresponding inclusions hold with $A(F) = (T^m A)(F)$ in place of $(T^m B)(F)$. By Lemma 12.10, the Hausdorff distance from $(T^m B)(F)$ to $A(F)$ is at most c^m/ϵ for the metric $d_{z,q}$. Thus for any Riemannian distance function on ∂H^n, the Hausdorff distance is $O(c^m)$. This finishes the proof of the theorem.

Consider, for instance, the Ford disc packing. The intervals $[x_m, y_m]$ converge at the rate $O(r^{2m})$, since r^2 is the derivative of $3 - x^{-1}$ at its fixed point $x = r^{-1}$. As $c = 1/4 > r^2$, this is consistent with Thm. 12.6.

$(T^m A_0)(F)$ is a countable closed subset of $A(F)$ that approximates $A(F)$ well for large m. Other approximations $(T^m B)(F)$, however, can be explicitly drawn in many cases. Say $n = 3$ and H^3/Γ is noncompact with finite volume, for instance. One can often choose the $B(F)$ so that each $(T^m B)(F)$ is bounded by finitely many circular arcs, the number of arcs grows rapidly with m, and the slice $A(F)$ appears to have fractal boundary, c.f. (18.4).

Section 13. Some chain-finite shadow families

We need a criterion for a shadow family to be chain-finite. Let G be the isometry group of H^n and let X^n be the space of all conformal balls $B \subset \partial H^n$, so X^n is a homogeneous n-manifold for G. In the hyperboloid model of H^n, X^n is parametrized by the hyperboloid of unit spacelike vectors, where the vector v determines the ball $B = \{w \in H^n : < w, v > \geq 0\}$. The action of G on X^n is not properly discontinuous, since the stabilizer H in G of a point $B \in X^n$ is noncompact, so the orbit $\{\gamma(B) : \gamma \in \Gamma\} \subset X^n$ of a discrete group $\Gamma \subset G$ may or may not be discrete. We say B is a Γ-*ball* when this orbit is discrete or, equivalently, when $H\Gamma$ is closed in G.

Theorem 13.1. *Let Γ be a discrete group of isometries of H^n such that H^n/Γ has finite volume but is noncompact. Let $B(q)$, $q \in Q = Q(\Gamma)$, be a Γ-invariant horoball packing of H^n and $S(q,p)$, $(q,p) \in A$, a $\Gamma-$invariant shadow family for this packing. If each shadow is an open subset of a finite intersection of Γ-balls then this shadow family is chain-finite.*

Suppose, for example, that $S(q,p)$ is a polyhedron in $\partial H^n - \{p\}$, as holds when $S(q,p)$ is a ridge shadow. Then $S(q,p)$ is a finite intersection of halfspaces, one for each smooth component of $\partial S(q,p)$, and so $S(q,p) \cup \{p\}$ is a finite intersection of balls. Clearly $S(q,p)$ is open in the latter set. The ridge shadows case of Thm 13.1 was treated in [F3], Thm. 12.

We note two ways to show that a given ball B is a Γ-ball. Let R_B be the reflection of H^n that fixes ∂B and let $P_B \subset H^n$ be the hyperplane fixed by R_B. If R_B commensurates Γ (that is, if $\Gamma \cap R_B \Gamma R_B^{-1}$ has finite index in Γ) then B is a Γ-ball, by Lemma 3 of [F3]. On the other hand, if $P_B/(\Gamma \cap H)$ has finite volume then B is a Γ-ball, by Thm. 1.13 of [Rag].

To prove the theorem, choose finite sets of Γ-balls $\mathcal{B}(q,p)$, $(q,p) \in A$, so that $S(q,p)$ is open in $\cap\{B : B \in \mathcal{B}(q,p)\}$. One may assume that ∂B meets $S(q,p)$ for each $B \in \mathcal{B}(q,p)$, that $\mathcal{B}(q,p)$ is invariant by the finite group $\Gamma_{(q,p)}$, and that the disjoint union $\mathcal{B} = \{(q,p,B) \in A \times \mathcal{B} : B \in \mathcal{B}(q,p)\}$ is Γ-invariant. Since A/Γ is finite, $\mathcal{B}/\Gamma$ is also finite. Thus the union $\mathcal{U} = \cup\mathcal{B}(q,p)$, $(q,p) \in A$, is a discrete set in X^n.

For each chain $S(p_{m+1}, p_m, ..., p_k)$, let $\mathcal{B}(p_{m+1}, p_m, ..., p_k) \subset X^n$ consist of those B such that ∂B meets $S(p_{m+1}, p_m, ..., p_k)$ and $B \in \mathcal{B}(p_{l+1}, p_l)$ for some l, $k \leq l \leq m$. Then $S(p_{m+1}, p_m, ..., p_k)$ is open in $\cap\{B : B \in \mathcal{B}(p_{m+1}, p_m, ..., p_k)\}$.

For $(q,p) \in A$, let $\mathcal{B}(q,p)^* = \cup \mathcal{B}(q,p,...)$. Then $\mathcal{B}(q,p) \subset \mathcal{B}(q,p)^* \subset \mathcal{U}$ and so $\mathcal{B}(q,p)^*$ is discrete in X^n. We now show $\mathcal{B}(q,p)^*$ is finite by showing that it lies in a compact subset of X^n. This will use the curvature estimate (2.3). For each compact set $K \subset \partial H^n - \{p\}$ and each $\beta > 0$, the set of all B such that ∂B meets K and $\kappa(\partial B, p) \leq \beta$ is compact in X^n. Thus $\mathcal{B}(q,p)^*$ is finite if $\kappa(\partial B, p)$ is bounded for $B \in \mathcal{B}(q,p)^*$.

We may suppose $k \leq 0 \leq m$, $B \in \mathcal{B}(p_1, p_0)$, and ∂B meets $S(p_{m+1}, p_m, ..., p_k)$. Let $\kappa_i = \kappa(\partial B, p_i)$, $0 \leq i \leq m$. (2.3) gives $\kappa_{i+1} \leq \max(\kappa_i, \kappa^+)$, where $0 \leq i < m$ and $\kappa^+ = 2e^D/(c^{-1/2} - c^{1/2})$. This implies $\kappa_m \leq \max(\kappa_0, \kappa^+)$. As $\mathcal{B}/\Gamma$ is finite, κ_0 is bounded. This gives a uniform bound β for κ_m, as desired.

Since $\mathcal{B}(q,p)^*$ is finite, it has only finitely many subsets. For each subset S the set $\cap\{B : B \in S\}$ has only finitely many components, so there are only finitely many chains $S(q,p,...)$, and the theorem is proved.

Note that the preceding proof works up until the last step when Γ is geometrically finite but not convex cocompact, $Q = Q(\Gamma)$, and each $S(q,p)$ is open in the intersection of L with finitely many Γ-balls. The "finitely many components" assertion fails in general, but it may hold if L is sufficiently connected. For such an example, with $n = 3$ and L a proper subset of ∂H^3, see §16.

We can use Thm. 13.1 to construct invariant chain-finite shadow families for certain discrete groups. Let P^* be a hyperbolic polyhedron of finite volume with one ideal vertex and one compact face C. Suppose there is an isometry ι of order two with $\iota(P^*) \cap P^* = C$ such that the group Γ generated by ι and the reflections in the noncompact faces of P^* is discrete and P^* is a fundamental domain for Γ. Choose an upper halfspace model so that ∞ is the ideal vertex of P^*. Then $P = \pi_\infty(P^*) = \pi_\infty(C)$ is a compact Euclidean polyhedron, Γ_∞ is a discrete group of Euclidean motions generated by reflections in the faces of P, and P is a fundamental domain for Γ_∞. Also H^n/Γ has finite volume and $Q(\Gamma)$ is the orbit of ∞. We may suppose that $0 = \iota(\infty)$. Let N be the hyperplane normal to $\overline{0\infty}$ stabilized by ι.

Choose a horoball B based at ∞ disjoint from C and define an invariant horoball packing by $B(p) = \gamma(B)$ for $p = \gamma(\infty) \in Q(\Gamma)$, $\gamma \in \Gamma$. Let A be the Γ-orbit of $(0,\infty)$. For $(q,p) \in A$ we define the *P-shadow* $S(q,p) = \cup\gamma(P)$, where γ varies over the elements of Γ with $\gamma(0,\infty) = (q,p)$. This family obeys (3.1) and (3.3). ι is expanding on P (in the metric d_p) if and only if P lies in the open disc centered at 0 bounded by ∂N. In this case (3.2*) holds and we have a Γ-invariant shadow family. If $C \subset N$, for instance if ι is a reflection, then ι is expanding on P and each P-shadow is a ridge shadow.

Corollary 13.2. *When ι is expanding on P, the P-shadows form a chain-finite shadow family.*

First, all P-shadows are Γ-equivalent. $S(0,\infty)$ is a compact Euclidean polyhedron and each of its faces is fixed by a reflection in Γ_∞. The corollary now follows from Thm 13.1.

Note that [Ra] §7.3 gives simplices $P^* \subset H^n$, $n \leq 9$, to which this corollary applies, with ι a reflection. We will examine some of these simplex reflection groups in §15 and §17. A survey of hyperbolic reflection groups can be found in [VS].

Consider a triangle P^* in the upper halfplane with vertices $\infty, w \in H^2$, and $-1/w$. Let L and R be the reflections in the vertical sides of P^* and let $\iota(z) = -1/z$. $\Gamma =< L,\ R,\ \iota >$ is discrete with fundamental domain P^* if and only if the angle sum of P^* is π/k for some integer $k \geq 2$. For a fixed k there is a curve of possible w's, and when $k = 3$ these are the groups defined in (6.9). $C \subset N$ when $|w| = 1$ (in this case Γ is an index two subgroup of the triangle group $(2, 2k, \infty)$ of §15) so ι is expanding on P for $|w|$ near 1. Thus for each $k \geq 2$, Cor. 13.2 gives a continuous family of finite area quotients, each of which admits a chain-finite shadow family.

Section 14. Markov partitions for $n = 2$.

Now fix x with $0 < x < 1$ and consider the Ford disc packing with the shadow family $S(\gamma(0), \gamma(\infty)) = \gamma([x-1, x])$, $\gamma \in \Gamma = PSl(2, \mathbf{Z})$. An aimed sequence $p_0, p_1, \ldots$ with $p_0 = \infty$ corresponds to a continued fraction expansion (2.6) with remainder interval $[x-1, x]$. If x is rational and $B = [x-1, x]$ then R_B commensurates Γ, so Thm 13.1 applies. If x has degree 2 and x' is its algebraic conjugate, then $[x', x]$ and $[x, x']$ are Γ-balls and the integer translates of these intervals are also Γ-balls. If $|x' - x| > 1$ we see that $[x-1, x]$ is a finite intersection of Γ-balls and Thm 13.1 again applies. Even when $|x - x'| < 1$, however, one can use the proof of Thm 13.1 to show that only a finite number of points can occur in the chain boundaries for a fixed pair (q, p). This shows that this shadow family is still chain-finite.

So each x of degree at most 2 with $0 < x < 1$ gives a Markov partition for the geodesic flow over H^2/Γ. Taking $x = 1/2$ gives the ridge shadow family and the Markov partition based on least-remainder continued fractions that we discussed in earlier sections of this paper. One expects that as x approaches 1 these Markov partitions converge, in some geometric sense, to the partition based on simple continued fractions [F2]. If so, this interpolates a family of reduction theories between the theories of Hurwitz and of Gauss.

One can generalize the chain-finite shadow families for x rational to other packings of H^2, as in the next theorem. Let $B(q)$, $q \in Q$, be a horoball packing of H^2, $S(q, p)$, $(q, p) \in A$, a standard shadow family, and Γ a symmetry group. For each $(q, p) \in A$ we have an interval $\mathcal{I}(q, p)$ such that $S(q, p) = L \cap \mathcal{I}(q, p)$. If we choose this interval as small as possible then its endpoints belong to L and may be called the endpoints of $S(q, p)$. These minimal intervals are permuted by Γ and so are the endpoints of the shadows. Lemma 6.5 shows that the chains for a standard family are intervals in L and so they have well-defined endpoints, too. We call a component interval of $\partial H^2 - L$ a *gap* and its endpoints *gap points.* Let G be the set of all gap points.

Theorem 14.1. *Suppose Γ is a geometrically finite group of isometries of H^2 that is not convex cocompact, $B(q)$, $q \in Q = Q(\Gamma)$, is an invariant horoball packing of H^2, and $S(q, p)$, $(q, p) \in A$, is an invariant standard shadow family. If the endpoints of the shadows are all in $Q \cup G$ then this family is chain-finite. Moreover such a shadow family can be constructed by perturbing the ridge shadow family.*

We mention that for $n = 2$, Γ is geometrically finite if and only if Γ is discrete with a finite set of generators ([Ra], Thm. 12.3.9). We will use some classical results concerning such groups, due to Koebe and Nielsen. First, any gap is stabilized by a hyperbolic element of Γ ([FN], p.119). This implies that a gap point cannot be a cusp point ([Ra], Thm. 5.5.4). Second, if a cusp point is fixed by a hyperbolic reflection in Γ then the other fixed point in ∂H^2 is either a cusp point or lies in a gap ([FN], p.118). This implies that Γ acts freely on $Q \times (L - Q)$, as noted after Lemma 11.1.

For the second assertion of the theorem, we must perturb the $\mathcal{I}(q,p)$'s to intervals with endpoints in $Q \cup G$ in a Γ-invariant fashion. Say e is an endpoint of the ridge shadow $S(q,p)$. If e lies in no $S(q',p),\ q' \neq q$, then $e \in G$ and we let $\epsilon(p,e) = e$. Otherwise e is an endpoint of some $\mathcal{I}(q',p),\ q' \neq q$. If some nontrivial $\rho \in \Gamma_p$ stabilizes e then $\rho(q) = q'$ and ρ is the hyperbolic reflection in the line $\overline{pe}$. As $e \in L$, the preceding paragraph implies that $e \in Q$. In this case we also let $\epsilon(p,e) = e$. In the remaining case, when e is an endpoint of two shadows and the stabilizer of e in Γ_p is trivial, we may perturb e to $\epsilon(p,e) \in Q$ and let $\epsilon(\gamma(p),\gamma(e)) = \gamma(\epsilon(p,e))$, for all $\gamma \in \Gamma$. We use these perturbed endpoints $\epsilon(p,e)$ to define an invariant family of intervals $\mathcal{I}'(q,p)$ that approximate the $\mathcal{I}(q,p)$.

For a sufficiently small perturbation, the $S'(q,p) = L \cap \mathcal{I}'(q,p),\ (q,p) \in A$, form a shadow family that satisfies conditions (6.1) and (6.2) of a standard family. This relies on the fact that A/Γ is finite. To arrange that (6.3) holds as well, one should choose the $\epsilon(p,e)$ so that when $(r,p) \in A$ and $r \notin S(q,p)$ then $r \notin S'(q,p)$. This is possible since the r's form a discrete set in $\partial H^2 - \{p\}$. Then if $S'(q,p) \cup S'(r,q) = L$ we find $q \in S'(q,p)$ and $p \in S'(r,q)$ so $q \in S(q,p)$ and $p \in S(r,q)$. Thus the triple (r,q,p) is such that $\overline{pq}$ meets $R(p,q) \cap R(r,q)$. Since there are only finitely many triples satisfying this condition up to the action of Γ, and since $S(q,p) \cap S(r,q)$ is empty for such a triple, we can choose our perturbation so small that $S'(q,p)$ is disjoint from $S'(r,q)$ for all such triples and assure that (6.3) holds.

For the first assertion of the theorem, we let $\mathcal{T}$ be the set of all triples (q,p,e) such that $(q,p) \in A$ and e is an endpoint of some chain $S(q,p,\ldots)$. We claim that $\mathcal{T}/\Gamma$ is finite. Granting this claim, we find for each $(q,p) \in A$ that there are only finitely many e with $(q,p,e) \in \mathcal{T}$, since $\Gamma_{(q,p)}$ is finite. So the endpoints of the chains $S(q,p,\ldots)$ form a finite set and so there are only finitely many such chains, as desired.

To prove the claim, note that when $(q,p,e) \in \mathcal{T}$ we may choose a chain $S(p_{i+1},p_i,\ldots,p_0)$ with endpoint e and $(p_{i+1},p_i) = (q,p)$ and with $i \geq 0$ as small as possible. When $i > 0$, this implies that e is an endpoint of $S(p_1,p_0)$ but not of $S(p_{i+1},p_i,\ldots,p_1)$. By Theorem 1, we can extend $p_0,\ p_1,\ldots$ to a sequence aimed at e. We let $\mathcal{Q} \subset A \times \mathcal{T}$ be the set of all quintuples (q',p',q,p,e) such that e is the endpoint of a chain $S(p_{i+1},p_i,\ldots,p_0)$ with $(p_{i+1},p_i) = (q,p)$ and $(p_1,p_0) = (q',p')$, e is an endpoint of $S(p_1,p_0)$, but e is not an endpoint of $S(p_{i+1},p_i,\ldots,p_1)$ when $i > 0$. Thus the projection from $\mathcal{Q}$ to $\mathcal{T}$ is onto. As this projection is equivariant, it suffices to show that $\mathcal{Q}/\Gamma$ is finite. Since A/Γ is finite and the projection from $\mathcal{Q}$ to A is also equivariant, it suffices to show that the latter projection is finite-to-one.

So fix (q',p') and fix one of the two endpoints e of $S(q',p')$. When $e \in Q$, there are only finitely many aimed sequences $p_j,\ j \in J$, with limits (e,p') and with $\inf(J) = 0$. For, as in the proof of Convergence, we refer to an upper halfplane model with $e = \infty$. Then the diameters $h(B(p_i))$ increase by a factor of at least $1/c$ at each step but stay bounded. This bounds the length of the aimed sequence. As successive horoballs are at most 2Δ apart, the number of aimed sequences is finite. Thus there are only finitely many possible (q,p).

If $e \notin Q$ then $e \in G$ and we may suppose $i > 0$. The gap with endpoint e is then a subinterval of $\mathcal{I}(p_{i+1},p_i,\ldots,p_1)$, since the latter contains e but does not have e as an endpoint. But in the metric $d_{q'}$ we know that the diameter of $\mathcal{I}(p_{i+1},p_i,\ldots,p_1)$ decreases like $c^{i/2}$. This bounds i from above and so once again there are only finitely many possible (q,p). Thus the projection from $\mathcal{Q}$ to A is finite-to-one and the theorem follows.

There is a dynamical way to construct the endpoints of the chains in Thm 14.1. Consider the reduced aimed correspondence $\mathcal{A}/\Gamma$ on $\mathcal{L}/\Gamma$ defined in §8 and apply it repeatedly to the classes $[e] \in \Lambda_p$, where e is an endpoint of $S(q,p)$, using all the pertinent branches of the correspondence. When e is a gap point, one only obtains classes $[g]$ where g is a gap point. As $\mathcal{A}/\Gamma$ is expanding, the corresponding gaps have at least a certain size. These classes thus form a finite set in $\mathcal{L}/\Gamma$. If e is a cusp point then, since there are only finitely many aimed sequences with limits (e,p), one again obtains only a finite set of classes. So when we apply $\mathcal{A}/\Gamma$ repeatedly to all of the endpoint classes $[e]$ we get only a finite set in $\mathcal{L}/\Gamma$ altogether. But this set contains $[z]$ whenever $z \in L - \{p\}$ is an endpoint of a chain $S(q,p,...)$. This method will be used in the next section to calculate chains and tiles.

Thm. 14.1 leads to Markov partitions.

Corollary 14.2. *Suppose Γ is a finitely generated discrete group of isometries of H^2 that is not convex cocompact and let $L = L(\Gamma)$ and $\mathcal{R} = \mathcal{R}(\Gamma)$. There is a Γ-invariant Markov partition for $\phi_t : \mathcal{R} \to \mathcal{R}$ indexed by $\mathcal{F}$ such that each tile and slice is an interval in L. When Γ acts freely on $\mathcal{F}$, for instance when Γ is orientation preserving, this gives a Markov partition for $\overline{\phi}_t$ indexed by the finite set $\mathcal{F}/\Gamma$. The reduced Markov correspondence on this set is then orbit-equivalent to $\overline{\phi}_t$.*

The first assertion follows from Thm. 14.1, Cor. 9.5, and Lemma 6.6. If Γ preserves orientation then each cusp stabilizer Γ_q acts freely on $L - \{q\}$ and so Γ acts freely on $\mathcal{F}$. The final assertion follows from Cor. 11.4.

Consider the shadow family of Thm. 14.1 when Γ does not act freely on $\mathcal{F}$. One can easily construct a marked shadow family (see Appendix 3) on which Γ acts freely. If some reflection $\rho \in \Gamma$ fixes $(q,p) \in A$ then one subdivides $S(q,p)$ into two marked shadows, each an interval in L with q as an endpoint. Otherwise one takes $S(q,p)$ as a marked shadow. It is easy to see that this marked shadow family is chain-finite. But Γ acts freely on the family of Markov boxes, so one obtains a Markov partition for $\overline{\phi}_t$. Thus marked shadows give an extension of Cor. 14.2 to the case when Γ does not preserve orientation.

The Markov partition of Cor. 14.2 depends analytically on parameters. Suppose Γ_λ is a family of geometrically finite groups with $\Gamma_0 = \Gamma$ and there are isomorphisms $\theta_\lambda : \Gamma \to \Gamma_\lambda$ such that $\theta_0 = id$, each path $\theta_\lambda(\gamma)$ is real analytic, and $\theta_\lambda(\gamma)$ is parabolic if and only if γ is parabolic. For each nontrivial $\gamma \in \Gamma$, the fixed points of $\theta_\lambda(\gamma)$ vary analytically in λ. But each endpoint e of each shadow in Thm. 14.1 is fixed by a nontrivial $\gamma \in \Gamma$. So by analytic continuation one may define a family of intervals $\mathcal{I}_\lambda(q,p)$ with $\mathcal{I}_0(q,p) = \mathcal{I}'(q,p)$ such that the endpoint e_λ corresponding to e is fixed by $\theta_\lambda(\gamma)$. The limits sets $L(\Gamma_\lambda)$ vary continuously, indeed there is a continuous family of homeomorphisms $h_\lambda : L(\Gamma) \to L(\Gamma_\lambda)$ with $h_0(z) = z$ such that $\theta_\lambda(\gamma)h_\lambda(z) = h_\lambda(\gamma(z))$. This is shown using the λ-lemma of Mane, Sad, and Sullivan, see [MT] Lemma 5.43. It follows that the sets $S_\lambda(q,p) = \mathcal{I}_\lambda(q,p) \cap L(\Gamma_\lambda)$ form a standard Γ_λ-invariant chain-finite shadow family for small λ. The endpoints of the corresponding tiles $Z_\lambda(F)$ vary analytically. When the endpoints of the slices $A(F)$ are fixed by hyperbolic elements of Γ, as seems always to be the case, the endpoints of the slices $A_\lambda(F)$ also vary analytically and so the Markov boxes $M_\lambda(F)$ are defined by intersecting $\mathcal{R}(\Gamma_\lambda)$ with analytically varying regions in SH^2. As $|\lambda|$ grows, however, the constant $c = c(\lambda)$ of (3.2) may approach 1 and the Markov boxes may degenerate in the direction of the flow.

Section 15. Triangle groups

We will now examine several examples, each related to the Ford disc packing and $PGl(2, \mathbf{Z})$. We will treat triangle reflection groups and the groups (6.9) with triangular fundamental domain in this section and treat higher dimensional examples in the next three sections.

When Γ is geometrically finite and $Q(\Gamma)$ consists of a single Γ-orbit, we will say that Γ is chain-finite if an invariant horoball packing with $Q = Q(\Gamma)$ is chain-finite. Note that two invariant horoball packings determine the same ridges and shadows in this case, so the choice of packing is unimportant. For instance, Cor. 13.2 shows that a reflection group based on a polyhedron of finite volume with one ideal vertex is chain-finite.

Consider a triangle in H^2 with one ideal vertex and two finite vertices such that reflections in its sides generate a tiling of the hyperbolic plane. Choose an upper halfplane model with the ideal vertex at ∞ so that the reflection in the opposite side sends ∞ to 0. Then the three sides of the triangle lie on lines $\Re(z) = a$, $\Re(z) = b$, and $|z| = c$ that we denote by L_A, L_B, and L_C respectively (as in [F2]).The three side reflections act on ∂H^2 by $A(t) = 2a-t$, $B(t) = 2b-t$, and $C(t) = c^2/t$. Then there are integers p and q so that L_C meets L_A in an angle of π/p radians and L_C meets L_B in an angle of π/q radians. Then $p \geq 2$, $q \geq 2$, and $(p, q) \neq (2, 2)$ since the angle sum of a hyperbolic triangle is at most π. Conversely, these conditions are sufficient for the existence of a hyperbolic reflection group (p, q, ∞) whose fundamental domain is a triangle with angles 0, π/p, and π/q, see [Ra], Fig. 7.3.1. Note that $(2, 3, \infty)$ is just the group $PGl(2, \mathbf{Z})$ of all symmetries of the Ford disc packing. We will assume $q \geq p$ as well, so $q > 2$. The group of translations that fix ∞ is generated by AB and one may choose coordinates so that $AB(t) = t - 1$. With this choice (which differs from that used in [F2]) we find $a = -c\ \cos(\pi/p) \leq 0$, $b = a + 1/2 = c\ \cos(\pi/q) > 0$, and $c = (2\ \cos(\pi/p) + 2\ \cos(\pi/q))^{-1} > 0$. Let $(p, q, \infty)^+ \subset (p, q, \infty)$ be the index two subgroup of orientation-preserving isometries, which is the free product of a cyclic group of order p generated by $x = CA$ and a cyclic group of order q generated by $y = CB$.

Clearly $(p, q, \infty)^+$ is an instance of our "test case." The portion of the fan $F(\infty)$ between the vertical lines L_A and L_B is bounded by a smooth curve that projects to the interval $I = [a, b]$ of length 1/2. Take $I \cup AI$ as the fundamental domain for the action of $\mathbf{Z}$ by translations. The edge-pairings lead us to a piecewise smooth expanding map E of $\mathbf{R}/\mathbf{Z}$ given by $E(t + \mathbf{Z}) = x^{-1}(t) + \mathbf{Z}$ for $t \in I$ and $E(t + \mathbf{Z}) = x(t) + \mathbf{Z}$ for $t \in AI$. If $p > 2$ then $S(0, \infty) = [a, b]$, whereas $S(0, \infty) = [-1/2, 1/2] = [b - 1, b]$ when $p = 2$.

By Cor. 13.2, (p, q, ∞) is chain-finite. The chains and Markov boxes can be completely described for $p = 2$, that is for the right triangle reflection group $\Gamma = (2, q, \infty)$, $q \geq 3$. As above, $S(0, \infty) = [-1/2, 1/2]$, $A(t) = -t$, $B(t) = 1 - t$, $C(t) = 1/(4\cos^2(\pi/q)t)$, $x(t) = CA(t)$, $y(t) = CB(t)$, and $xy(t) = t-1$. Then $AC = CA$ and $A^2 = B^2 = C^2 = y^q = 1$. Let $N = N(q)$ be $(q-2)/2$ or $q-2$ as q is even or odd.

Theorem 15.1. *(a) The tiles $Z(F)$, $F \in \mathcal{F}(0, \infty)$, consist of $2N$ intervals with endpoints z_j, $|j| \leq N$, where $z_{-j} = -z_j$ and $0 = z_0 < z_1 < ... < z_N = 1/2$. The z_j, $0 < j < N$, are specified by the rules $y(z_j) = z_{j+1}$, $j \geq 0$, for even q and by the rules $y(z_j) = z_{j+2}$, $0 \leq j \leq N-2$, and $yxy(z_{N-1}) = z_1$ for odd q.*
(b) There are constants β and α_j, $j = 1, ..., N$, so that $0 < -\beta < \alpha_1 < \alpha_2 < ... < \alpha_N$ and $K_z(0, \infty) = [\alpha_j, \beta]$ for $z_{j-1} < z < z_j$. β is the unique negative fixed point of C or of

$y^{(q-1)/2}A$ as q is even or odd. $\alpha_N = B(\beta)$ and the other α_j are determined by the formulas

$$y(\alpha_j) = \alpha_{j+1},\ j < N,\ yxy(\alpha_N) = \alpha_1$$

for even q and

$$y(\alpha_j) = \alpha_{j+2},\ j < N-1,\ yxy(\alpha_j) = \alpha_{j+2-N},\ j \geq N-1$$

for odd $q \geq 5$.
(c) Let $Z_j = [z_{j-1}, z_j]$, $A_j = [\alpha_j, \beta]$, and $K_j = Z_j \times A_j$ for $j = 1, ..., N$. Then the sets $K(F)$, $F \in \mathcal{F}(0,\infty)$, consist of the K_j and $A(K_j) = -K_j$ for $j = 1, ..., N$. There is a reduced Markov partition for the geodesic flow over H^2/Γ with N Markov boxes whose transversals are parametrized by $K_1, ..., K_N$.

To follow the proof it is helpful to draw the disc model of the hyperbolic plane with the origin at the fixed point of y: this makes many equalities and inequalities obvious, and we will take them for granted in what follows. For two typical cases, each with $N = 3$, see Figure 6 (the regular polygons in this figure are just there for reference).

To find the tiles $Z(F)$, apply x to the known endpoints $\pm 1/2$ and use integer translations to bring the resulting points back into $[-1/2, 1/2]$. This produces a new endpoint which is 0 for $q = 3$ or the nonzero point $y^{-1}(1/2) = 1 + x(1/2)$ for $q > 3$. In the latter case, apply x again and use another translation to find another endpoint, etc. If q is even one obtains successively $1/2 > y^{-1}(1/2) > y^{-2}(1/2) > ... > y^{-N}(1/2) = 0$, which are the endpoints $z_N,\ z_{N-1},\ z_{N-2}, ...,\ z_0$. If $q = 2m+3 \geq 5$ is odd, however, the sequence $y^{-k}(1/2)$ is positive and decreasing for $0 \leq k \leq m$ but is negative for $k = m+1$. In fact the point $t = 1 + y^{-m-1}(1/2)$ lies between $y^{-1}(1/2)$ and $1/2$. Apply y^{-1} to t repeatedly to get a decreasing sequence with $y^{-(N-1)/2}(t) = 0$. This gives the endpoints z_j with $z_N = 1/2,\ z_{N-1} = t,\ z_{N-2} = y^{-1}(1/2),\ z_{N-3} = y^{-1}(t)$, and so on.

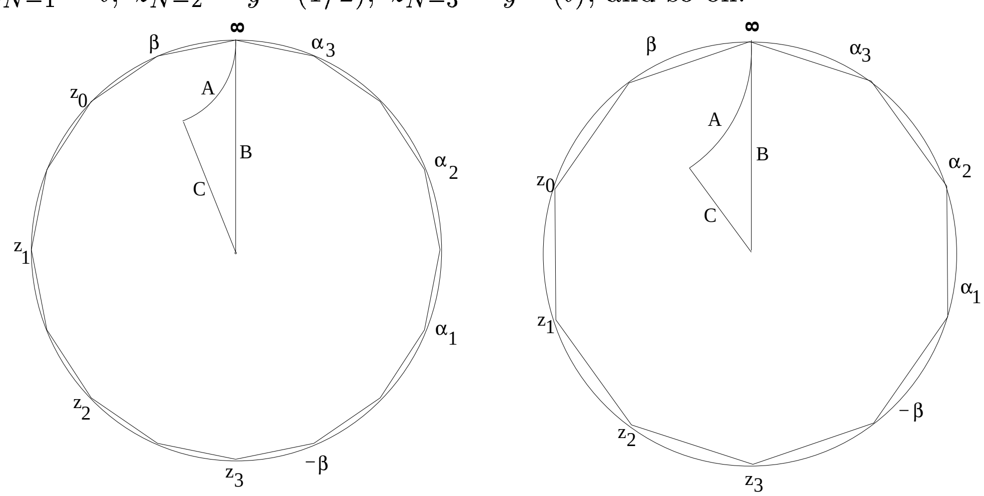

Figure 6. Tiles and slices for the Hecke triangle groups with q=8 and q=5

Examining each step, we see that each interval $[-1/2, z_k]$, $k \geq 0$, is a chain. First off, $[-1/2, y^{-k}(1/2)] = S(0, \infty, y^{-1}(\infty), ..., y^{-k}(\infty))$ if $0 \leq k \leq (q-2)/2$, and $y^{-1}(\infty) = 1$. For $q = 2m+3$ we get $[-1/2, t] = S(0, \infty, 1 + y^{-1}(\infty), ..., s)$ where $s = 1 + y^{-m-1}(\infty)$. It follows for $0 \le k \le m$ that $[-1/2, y^{-k}(t)] = S(0, \infty, y^{-k}(2), ..., y^{-k}(s))$. As $x(0) = \infty$ it follows that every chain $S(0, \infty, p, ...)$ is either $[-1/2, z_k],\ 0 \le k \le N$, if $p > 0$ or $[z_k, 1/2],\ -N \le k \le 0$, if $p < 0$. Applying inclusion/exclusion to these chains, we see that the tiles are as stated and (a) is proven.

We begin the proof of (b). For $j = 1, ..., N$ and $z_{j-1} < z < z_j$, Lemma 6.6 shows that the slice $K_z(0, \infty)$ is an interval $[\alpha_j, \beta_j]$. Since ∞ is an interior point of this slice we have $\beta_j < \alpha_j$. Thm. 6.8 shows that the sequences α_j and β_j are increasing. Lemma 9.4 shows that the sequence of sets $[\alpha_j, \beta_j]$ is strictly decreasing so the sequence α_j is strictly increasing and $\beta_j = \beta$ is independent of j. Using symmetry with respect to the origin, Thm. 6.8 gives $0 < -\beta_1 \le \alpha_1$.

The rest of (b) uses that $K(\infty)$ is tiled in 2 ways, both by the sets $K(m, \infty) = m + K(0, \infty)$ and by the sets $K(\infty, m) = m + K(\infty, 0)$, as in Prop. 5.2 and Figure 2, where we use that the cusp points adjacent to ∞ are just the integers m. Since $K(\infty, 0) = x(K(0, \infty))$, this forces relations between the endpoints β and α_j. In detail, we let $K^>$ consist of the finite points $(z, a) \in K(\infty)$ with $a > z$. Then $K^>$ is bounded below by the graph of the step function $\sigma(t)$ with values $m + \alpha_j$ on $m + Z_j$ and $m - \beta$ on $m - Z_j$, for $j = 1, ..., N$ and integer m. The subset K^+ of $K(0, \infty)$ with $z > 0$ transforms by x into $K^>$ and the integer translates $m + x(K^+)$ tile $K^>$. It follows that the portion of $x(K^+)$ with $x(1/2) - 1 < z < x(1/2)$ is bounded below by the graph of $\sigma(t)$.

Fix an odd $q \ge 5$. Then $x(z_0) = \infty,\ x(z_1) = xyxy(z_{N-1}) = z_{N-1} - 2$, and $x(z_j) = xy(z_{j-2}) = z_{j-2} - 1$ for $2 \le j \le N$. As $[z_{N-2} - 2, z_{N-1} - 2] = [x(z_n) - 1, x(z_1)] \subset x([z_0, z_1])$, one finds $\alpha_{N-1} - 2 = x(\alpha_1)$, hence $\alpha_1 = yxy(\alpha_{N-1})$. As $[z_{N-1} - 2, z_0 - 1] = x([z_1, z_2])$, one finds $\alpha_N - 2 = -\beta - 1 = x(\alpha_2)$. Thus $\alpha_2 = yxy(\alpha_N)$ and $\beta = B(\alpha_N)$. Finally $[z_{j-1} - 1, z_j - 1] = x([z_{j+1}, z_{j+2}]),\ j \ge 1$, so $\alpha_j - 1 = x(\alpha_{j+2})$ and $\alpha_{j+2} = y(\alpha_j)$. This proves the formulas for the α_j in (b). The proof for other q is similar, but simpler, and we will omit it.

The integer translates of the region $x(K^+)$ cover the region $K^>$. Since

$$x(K_z(0, \infty)) = K_{x(z)}(\infty, 0) = [x(\alpha_1), x(\beta)],\ 0 < z < z_1,$$

we have $x(\beta) = 1 + x(\alpha_1)$. Thus $\alpha_1 = yx(\beta)$. Now for q odd we have $\beta = B(\alpha_N) = By^{(q-3)/2}(\alpha_1) = By^{(q-1)/2}x(\beta)$ and $By^{(q-1)/2}x = y^{-(q-1)/2}Bx = y^{(q-1)/2}A$ since $By = yB^{-1}$, $y^q = 1$, and $yBx = A$. But $y^{(q-1)/2}A[0, \infty] = y^{(q-1)/2}[\infty, 0] \subset [0, \infty)$. As $y^{(q-1)/2}A$ reverses orientation, it follows that it has a positive sink and a negative source and no other fixed points. On the other hand, if q is even then $yx(\beta) = \alpha_1 = yxy(\alpha_N) = yxyB(\beta)$, so β is fixed by $yB = C$. The fixed points of C are $\pm c$. Thus the fixed point characterization of β is proven in both cases.

To show that $-\beta \ne \alpha_1$, recall that y has no fixed point on ∂H^2. When q is even, $\alpha_1 = yx(\beta) \ne x(\beta) = A(\beta)$. When q is odd, $Ay^{(q-1)/2}(\alpha_1) = Ay^{(q+1)/2}x(\beta) = xy^{(q-1)/2}A(\beta) = x(\beta) \ne yx(\beta) = \alpha_1$. As $Ay^{(q-1)/2}$ fixes $A(\beta)$ but not α_1, we find $A(\beta) \ne \alpha_1$. Thus $-\beta \ne \alpha_1$, which finishes the proof of (b).

We now show (c). $< A >= \Gamma_{(0,\infty)}$ and this group of order 2 acts freely on $\mathcal{F}(0, \infty)$. It follows that Γ acts freely on $\mathcal{F}$ and so Cor. 11.4 gives a Markov partition indexed by $\mathcal{F}/\Gamma = \mathcal{F}(0, \infty)/ < A >$. The flowboxes for this partition are the images in SH^2/Γ of the flowboxes $M(F_j)$ for $j = 1, ..., N$ where $K(F_j) = K_j$. This proves (c) and hence the theorem.

Thm. 15.1 implies that each endpoint of each slice is the unstable fixed point of an element of the triangle group. Since each α_j is in the orbit of β, it is enough to treat β. For q odd, β is the unstable fixed point of $y^{(q+1)/2}A$. For q even, $\beta = B(\alpha_N) = By^{N-1}(\alpha_1) = By^N xy(\alpha_N) = By^N xyB(\beta)$. But $By^N xyB = y^{-N}BxC = y^{-N-1}x = y^{q/2}x$ and $y^{q/2}x[0,\infty] = y^{q/2}[\infty, 0] \subset (0,\infty)$. So β is the unstable fixed point of $y^{q/2}x$, as desired.

For each hyperbolic triangle reflection group (p, q, ∞), [F2] gives a Markov partition with $pq - p - q$ Markov boxes that is not among those produced by the methods of this paper. For $p = 2$ and $q = 3$, for instance, both have a single Markov box but one is based on simple continued fractions and the other on least-remainder continued fractions.

Now consider the group Γ of (6.9) depending on a parameter μ with $0 < \mu \leq 3$. Suppose $\mu = 3$. Then Γ is isomorphic to the kernel of the homomorphism $h : (2, 6, \infty) \to \{\pm 1\}$ with $h(A) = 1 = h(B)$, $h(C) = -1$. Here $N = (6 - 2)/2 = 2$ so the geodesic flow over H^2/Γ admits a Markov partition with 4 Markov boxes.

On the other hand, suppose $2 < \mu < 3$. We refer to Figure 7, which uses the upper halfplane coordinates of §6. As Γ_∞ is generated by the reflections $L(t) = -3 - t = R_p(t)$ and $R(t) = \mu - t = \iota(R_q(\iota(t)))$, we may identify $\mathcal{L}/\Gamma = \mathbf{R}/\Gamma_\infty$ with the interval $[-3/2, \mu/2] = S(-1,\infty) \cup S(0,\infty)$. The expanding map E from the reduced aimed correspondence $\mathcal{A}/\Gamma$ has two smooth branches, namely $\iota(t) = -\mu/t$ on $S(0,\infty) = [-\mu/2, \mu/2]$ and $\eta(t) = -2 - 1/(t+1) = R_p(R_q(t))$ on $S(-1,\infty) = [-3/2, -\mu/2]$. It follows that for certain values of μ, Γ is chain-finite. Since $\iota(\mu/2) = L(-1)$, $\iota(-1) = R(0)$, $\iota(0) = \infty$, and $\eta(-3/2) = 0$, we need only examine the orbits of $-\mu/2$ under E. As $\iota(-\mu/2) = R(\mu - 2)$, $|\mu - 2| < \mu/2$, and $\iota(\mu - 2) = L\eta(-\mu/2)$, Γ is chain-finite if and only if the orbit of $\eta(-\mu/2)$ under E is finite. In geometric terms this means that the geodesic on H^2/Γ corresponding to the downward ray through ξ' either diverges to ∞ or is asymptotic to a closed geodesic. This appears to hold for a countable dense set of μ's.

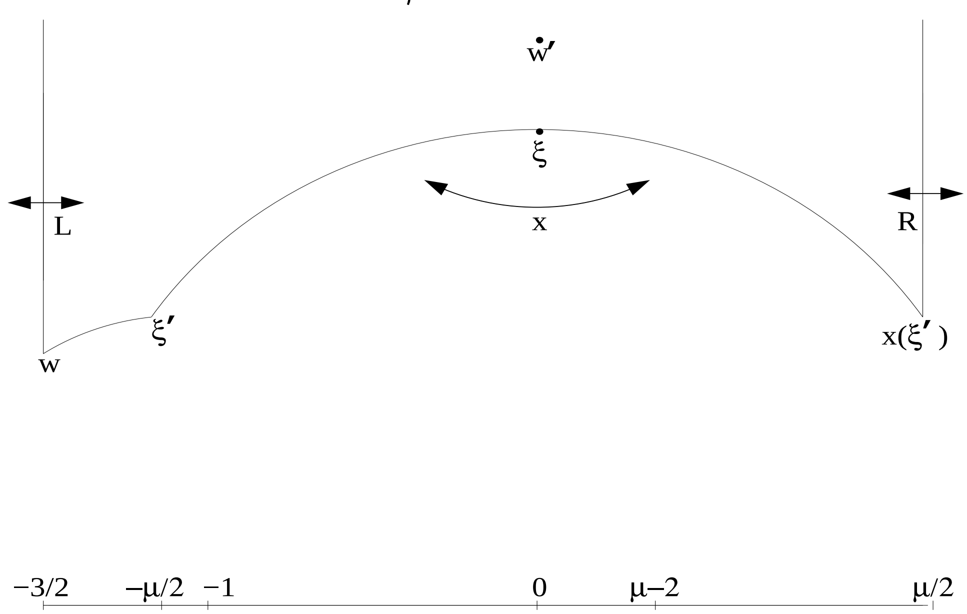

Figure 7. Shadows and tiles for 2< μ <3

We can now choose $\mu = \sqrt{3} + 1$ so that $y(-\mu/2) = \mu - 2$ and we get a chain-finite example. The only chain $S(-1, \infty, ...)$ is $S(-1, \infty)$ whereas the partition of $S(0, \infty)$ into tiles is

$$[-\mu/2, -1] \cup [-1, 0] \cup [0, \mu - 2] \cup [\mu - 2, \mu/2]. \tag{14.2}$$

The same tiles work for $\mu = \sqrt{13} - 1$, so that $y(-\mu/2) = \mu/2$, and for an infinite sequence of μ's that approaches 2, leading to Markov partitions with 5 Markov boxes.

Consider the triangle P^* with vertices ∞, w, and $\iota(w)$. This is a fundamental domain for Γ. When $9/4 < \mu \leq 3$, ι is expanding on $P = [-3/2, \mu/2]$ and so Cor. 13.2 gives an invariant chain-finite shadow family. These P-shadows are ridge shadows only for $\mu = 3$. There are 4 Markov boxes, corresponding to the decomposition of $S(0, \infty) = P = [-3/2, \mu/2]$ into the tiles $[-3/2, -1]$, $[-1, 0]$, $[0, \mu/3]$, and $[\mu/3, \mu/2]$. This illustrates the analytic dependence on parameters that was discussed at the end of §14.

Section 16. Bianchi groups and the Appolonian gasket

We now examine an important class of 3-dimensional examples. For a quadratic imaginary field k, whose ring of quadratic integers is O_k, the Bianchi group B_k is just $PSl(2, O_k)$, where O_k denotes the ring of quadratic integers in k. B_k is a discrete subgroup of $PGl(2, \mathbf{C})$, and so it acts properly discontinuously on H^3 by isometries. The quotient $X_k = H^3/B_k$ has finite volume but is noncompact, in fact there is one cusp in X_k for each element of the ideal class group I_k. As shown in [F3], there is a horoball packing invariant by B_k so that $h(B(p))$ is rational for all cusp points $p \neq \infty$ and such a packing is chain-finite since it obeys the hypotheses of Thm. 13.1. We also gave some examples in which the tiles $Z(F)$ were explicitly calculated. In particular 6 transversals suffice for a Markov partition for the geodesic flow over X_k for $k = \mathbf{Q}(\sqrt{-3})$ or $k = \mathbf{Q}(i)$, and only 3 are needed for $H^3/PGl(2, k)$ in these two cases.

We now give two examples of forked geodesics that are described by uncountably many aimed sequences, up to shift. For the first, let $(z_0, a_0) = ((1 + \sqrt{3}i)/2, (1 - \sqrt{3}i)/2)$ and let $\Gamma = PGl(2, \mathbf{Z}[i])$. We take the invariant horoball packing with $B(\infty) = B_3$. Let e and h be the elements of Γ with $e(z) = 1 - 1/z$ and $h(z) = (z - 1 + i)/((1 - i)z + i)$. Then e has order 3 and commutes with h and both fix (z_0, a_0). We see that the three horoballs $B(\infty)$, $B(0)$, and $B(1)$ are mutually tangent and are permuted by e. Applying h gives another such triple, namely $B((1 + i)/2)$, $B(i)$, and $B(1 + i)$. Note that i and $1 + i$ are each z_0-successors to ∞. It follows that any sequence $p_j = h^j(x_j)$, $j \in \mathbf{Z}$, with $x_j \in \{0, 1, \infty\}$ and $x_j \neq x_{j+1}$ for all j is aimed and has limit pair (z_0, a_0).

A similar discussion can be given for $(z_1, a_1) = (i, -i)$ and $\Gamma = PGl(2, \mathbf{Z}[z_0])$. One takes $e(z) = -1/z$ and $h(z) = (z + z_0)/(-z_0 z + 1)$. Then e has order 2 and commutes with h and both fix (z_1, a_1). One finds that any sequence $p_j = h^j(x_j)$, $j \in \mathbf{Z}$, with $x_j \in \{0, \infty\}$ is aimed and has limit pair (z_1, a_1).

In both instances, the semidirect product of Γ by $< \bar{z} >$ is a simplex reflection group and the forked geodesic corresponds to a periodic billiard path lying on an edge of the simplex. For these examples, the bounds in Cor. 11.4 imply that the natural map from allowed sequences F_j, $j \in \mathbf{Z}$, to aimed sequences p_j, $j \in \mathbf{Z}$, is not onto.

Next we consider a group with an infinite volume quotient. Let Γ^+ be the subgroup of $PGl(2, \mathbf{Z}[i])$ generated by the transformations $z + 1, -1/z$, and $i - z$. The limit set L of Γ^+

is known to be the *Appolonian gasket* $\mathcal{G}$. $\mathcal{G}$ is constructed by iteratively removing certain disjoint open discs from the extended complex plane. The first to go are the mutually tangent discs $\Im z > 1$, $\Im z < 0$, and $|z - i/2| < 1/2$. At each successive stage we remove the open discs that are tangent to three previously deleted discs. $\mathcal{G}$ is defined as the closed set left over after this infinite process. To see that $\mathcal{G}$ and L are equal, note that they both contain the extended real line $\hat{\mathbf{R}}$, which is the limit set of $PSl(2, \mathbf{Z}) =< z + 1, -1/z >\subset \Gamma^+$. The images of $\hat{\mathbf{R}}$ by Γ^+ then lie in L and one can check that they consist exactly of the boundaries of the deleted discs. It follows that L contains $\mathcal{G}$ and that the latter is invariant by Γ^+. Since L is a minimal Γ^+-invariant set, it equals $\mathcal{G}$. Γ^+ is the full group of conformal transformations of $\mathcal{G}$. Its normalizer $\Gamma = N(\Gamma^+)$ in the isometry group of H^3 is obtained by adjoining the reflection $-\overline{z}$ to Γ^+.

When we delete open discs from ∂H^3 to construct $\mathcal{G}$, we can also delete the corresponding open halfspaces from H^3 to obtain the convex hull $C(\Gamma)$. The Ford region R for Γ is found by removing the Euclidean unit balls centered at the points m and $m + i$, $m \in \mathbf{Z}$. As $(R \cap C(\Gamma))/\Gamma_\infty$ has only one end, one sees that Γ is geometrically finite and $Q = Q(\Gamma)$ is the orbit of ∞. Thus each cusp point lies in the closure of two deleted discs. Note that Γ_∞ acts freely on $L - Q$, in particular the reflections $-\overline{z}$, $1 - \overline{z}$, and $i + \overline{z}$ fix lines that meet L only in cusp points. So Γ acts freely on $Q \times (L - Q)$, even though Γ_∞ contains reflections.

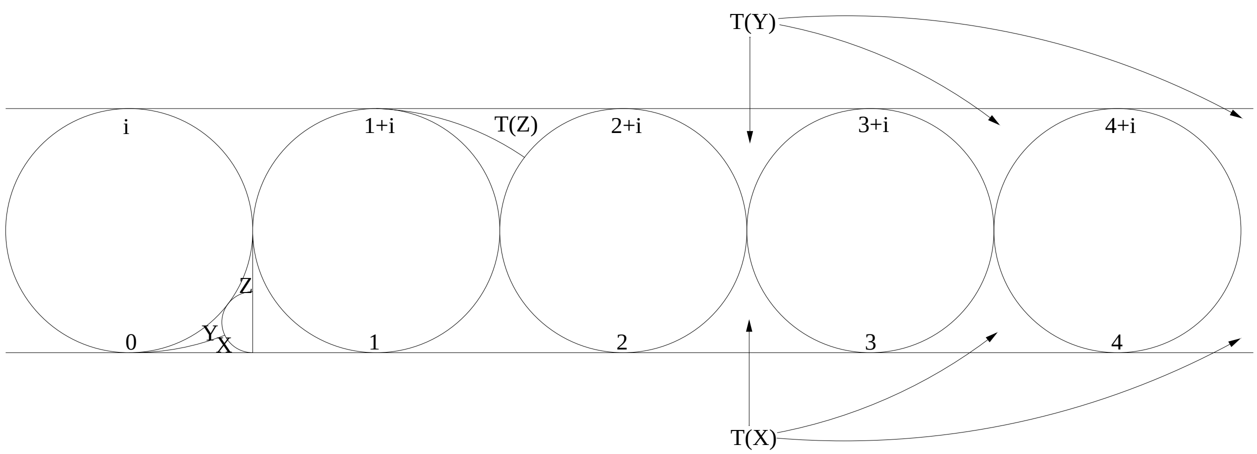

Figure 8. Tiles for the Appolonian gasket limit set

We will show that Γ is chain-finite, leading to a Markov partition for the recurrent part of the geodesic flow over H^3/Γ. The cusp points adjacent to ∞ are the points m and $m + i$ with integer m. Thus Γ^+ acts simply transitively on the set A of adjacent pairs of cusp points and $\Gamma_{(0,\infty)}$ is the order 2 subgroup generated by $-\overline{z}$. The ridge shadow $S(0, \infty)$ is the intersection of $\mathcal{G}$ with the halfstrip $|\Re(z)| \leq 1/2$, $\Im(z) \leq 1/2$.

We transform the 3 lines that bound the halfstrip by $T(z) = 1/\overline{z}$ to obtain unit circles centered at i and ± 1. These unit circles and those centered at $i \pm 1$ divide $S(0, \infty)$ into 6 compact sets. The three of these with nonnegative real part are shown in Figure 8 as X, Y, and Z. Since $\Gamma_\infty =< z + 1,\ i - z,\ -\overline{z} >$, we can identify $\mathcal{L}/\Gamma = \Lambda_\infty$ with $X \cup Y \cup Z$. The reduced correspondence $\mathcal{A}/\Gamma$ is given by applying T and reducing by the action of Γ_∞. But, except for the point ∞, each of $T(X)$, $T(Y)$, and $T(Z)$ is a union of images of these 3 sets by elements of Γ_∞, as can be seen from Figure 8. This establishes chain-finiteness.

Moreover the tiles $Z(F)$, $F \in \mathcal{F}(0,\infty)$, are just X, Y, and Z and their images by $-\overline{z}$. Say that $X = Z(F_1)$, $Y = Z(F_2)$, and $Z = Z(F_3)$.

As $\Gamma_{(0,\infty)} =< -\overline{z} >$ acts freely on $\mathcal{F}(0,\infty)$, Γ acts freely on $\mathcal{F}$. Thus we find

Theorem 16.1. *For the symmetry group Γ of the Appolonian gasket, the recurrent part of the geodesic flow over H^3/Γ has a Markov partition with Markov boxes $[M(F_j)]$, $j = 1,2,3$.*

Note that the inclusion $\hat{\mathbf{R}} \subset L$ defines an invariant set $\Lambda \subset \mathcal{R}/\Gamma$ for $\overline{\phi}_t$ isomorphic to the geodesic flow over $H^2/PGl(2,\mathbf{Z})$. Λ can also be characterized symbolically, as the union of all the flowlines that are contained in $[M(F_1)]$.

Section 17. The Hurwitz modular group

Just as linear fractional transformations over $\mathbf{R}$ or $\mathbf{C}$ give isometries of H^2 or H^3, linear fractional transformations of a quaternion variable z produce isometries of H^5. Note that the translations $z+q$, the similarity transformations qz, $q \neq 0$, and the involution $1/z$ are all conformal transformations of $\mathbf{H} \cup \{\infty\}$. Under composition, these generate all the transformations $\gamma_A(z) = (a_{11}z + a_{12})(a_{21}z + a_{22})^{-1}$, where $A = (a_{ij})$ is an invertible 2×2 quaternion matrix. Such matrices form a group $Gl(2,\mathbf{H})$ that acts conformally on ∂H^5 and only the real scalar matrices act trivially. Since the dimension of $Gl(2,\mathbf{H})$ is 16 and the dimension of the conformal group is 15, we obtain all conformal transformations of ∂H^5 and all orientation-preserving isometries of H^5 in this way.

A quaternion is *Hurwitz* if all four of its coefficients are integers or all four are half an odd integer. Let $\mathcal{H}$ be the ring of all Hurwitz quaternions. Let Γ be the *Hurwitz modular group* consisting of the linear fractional transformations corresponding to invertible matrices over $\mathcal{H}$ whose inverse matrix also has entries in $\mathcal{H}$. Thus $\Gamma = Gl(2,\mathcal{H})/\pm I$ is a discrete group of isometries of H^5. We will see shortly that H^5/Γ has finite volume and one cusp.

The geodesic flow $\overline{\phi}_t$ over H^5/Γ is closely related to the flows for Bianchi groups discussed in §16. Fix $v = v_1 i + v_2 j + v_3 k \in \mathcal{H}$ with relatively prime integers v_1, v_2, v_3. Then $v^2 = -m$ for some positive integer m not divisible by 4 and not congruent to -1 modulo 8. Conversely, by a result of Legendre, every such m can be expressed as a sum of three relatively prime squares. $\mathbf{Q}(v)$ is a quadratic imaginary field and $\mathcal{O}(v) = \mathbf{Q}(v) \cap \mathcal{H}$ is an order in $\mathbf{Q}(v)$, equal to $\mathbf{Z}[v]$ when m is congruent to 1 or 2 modulo 4 and to $\mathbf{Z}[(1+v)/2]$ when m is congruent to 3 modulo 8. Consider the 2-sphere $S(v) = \mathbf{R}(v) \cup \{\infty\}$, the hyperbolic 3-space $H(v) \subset H^5$ bounded by $S(v)$, and their stabilizer $\Gamma(v) \subset \Gamma$. The subgroup of $\Gamma(v)$ which preserves the orientation of $S(v)$ is $Gl(2,\mathcal{O}(v))/\pm I$, so $H(v)/\Gamma(v)$ has finite volume. Now let G be the isometry group of H^5 and $G(v)$ the stabilizer of $H(v)$ in G. Then $G(v)/\Gamma(v)$ and G/Γ have finite volume and so the natural injection $G(v)/\Gamma(v) \to G/\Gamma$ is proper by [Rag] Thm. 1.13. This implies that the natural immersions $i(v) : H(v)/\Gamma(v) \to H^5/\Gamma$ and $Si(v) : SH(v)/\Gamma(v) \to SH^5/\Gamma$ are proper. The geodesic flow $\psi(v)_t$ on $SH(v)/\Gamma(v)$ was studied in [F3] for the purpose of reducing quadratic forms over rings of quadratic imaginary integers. $\psi(v)_t$ corresponds under $Si(v)$ to $\overline{\phi}_t$, so $\psi(v)_t$ may be called a *properly immersed subflow* of $\overline{\phi}_t$.

We now build a Markov partition for $\overline{\phi}_t$ by following the procedure used for $PSl(2,\mathbf{Z})$. We will calculate the Ford region and ridge shadows and show this family is chain-finite. We will also find a simplex reflection group that contains Γ, much as the triangle group $(2,3,\infty)$ of §15 contains $PSl(2,\mathbf{Z})$.

We begin with some algebra. The units group in the ring $\mathcal{H}$ is $U = \{z \in \mathcal{H} : z\bar{z} = 1\}$. U is the semidirect product of $\mathcal{Q} = \{\pm 1, \pm i, \pm j, \pm k\}$ and the cyclic group $< u_0 >$ of order 3, where $u_0 = (1 + i + j + k)/2$. The next smallest Hurwitz integers are those of length $\sqrt{2}$, each of which has two zero components and two components equal to ± 1. Each of these can be uniquely written as $u(j + k),\ u \in U$. The Euclidean group Γ_∞ consists of the elements $\gamma_A(z)$ with $A \in Gl(2, \mathcal{H})$ and $a_{21} = 0$, so a_{11} and a_{22} are in U. Γ_∞ is the semidirect product of the group T of translations by Hurwitz integers (take $a_{11} = 1 = a_{22},\ a_{21} = 0$) and the stabilizer $\Gamma_{(0,\infty)}$ (take $a_{21} = 0 = a_{12}$ and $a_{11},\ a_{22} \in U$.)

Now let $\mathcal{C} = \{z \in \mathbf{H} : |z| \le |z - u|,\ u \in U\}$. The convex polyhedron with vertex set U is the self-dual *regular 24-cell* $\{3, 4, 3\}$ [Co], so $\mathcal{C}$ is also a regular 24-cell. The face of $\mathcal{C}$ with center $1/2$ is a regular octahedron $\mathcal{C}^+$ with vertices $(1 \pm i)/2, (1 \pm j)/2$, and $(1 \pm k)/2$. Thus each point of $\mathcal{C}$ has length at most $1/\sqrt{2} < 1$. Since every quaternion z can be written $z = h + r$ with $h \in \mathcal{H}$ and $r \in \mathcal{C}$, the Hurwitz integers form a Euclidean ring. The Voronoi cell for the lattice $\mathcal{H}$ is defined as $\{z \in \mathbf{H} : |z| \le |z - h|,\ h \in \mathcal{H},\ h \neq 0\}$. Since $|h| \ge 2/\sqrt{2}$ for $h \in \mathcal{H},\ h \neq 0,\ h \notin U$, this Voronoi cell equals $\mathcal{C}$. Thus $\mathcal{C}$ is a fundamental domain for T.

Next we find the isometric sphere for each element $\gamma_A \in \Gamma - \Gamma_\infty$. One may calculate the derivative of $\gamma_A(z)$ and see that it is expanding for $|a_{21}z + a_{22}|^2 < |a_{21}||a_{12} - a_{11}a_{21}^{-1}a_{22}|$. In fact the latter constant is 1 for $A \in Gl(2, \mathcal{H})$, as can be shown inductively by row reducing A to the identity matrix. Thus the isometric sphere is given by $|a_{21}z + a_{22}| = 1$. If $a_{21} \notin U$, the radius of this sphere is at most $1/\sqrt{2}$.

Let R be the closed region in the upper halfspace obtain by removing the Euclidean balls of unit radius centered at Hurwitz quaternions. If $w \in \partial R$ and $\pi_\infty w = h + r,\ h \in \mathcal{H},\ r \in \mathcal{C}$, then w has height $\sqrt{1 - |r|^2} \ge 1/\sqrt{2}$. It follows that R is the Ford region for Γ. The polyhedron $R \cap \pi_\infty^{-1}(\mathcal{C})$ maps onto H^5/Γ, so the latter has one end and finite volume. It follows that $Q = Q(\Gamma)$ is the Γ-orbit of ∞. Since $\mathcal{H}$ is Euclidean, this orbit consist of ∞ and all the quaternions with rational coefficients. Also the cusp points adjacent to ∞ are just the Hurwitz integers. Thus all ridges are Γ-equivalent.

We find that the ridge shadow $S(0, \infty)$ equals $\mathcal{C}$. Since all ridge shadows are equivalent to $\mathcal{C}$ and the faces of $\mathcal{C}$ are just $u(\mathcal{C}^+),\ u \in U$, it is easy to see that this shadow family is chain-finite. Let $\Sigma \subset \mathbf{H}$ be the hyperplane that bisects the interval from 0 to 1. We apply the transformations $h + 1/z,\ h \in \mathcal{H}$, to the spheres $u(\Sigma) \cup \{\infty\}$ that bound $\mathcal{C}$ to get the unit spheres centered at Hurwitz integers. Since $1 + 1/\sqrt{2} < \sqrt{3}$, such a sphere meets $\mathcal{C}$ only if its center has length 1 or $\sqrt{2}$. Applying $h + 1/z,\ h \in \mathcal{H}$, to these 48 unit spheres gives 12 new spheres that meet Int $S(0, \infty)$, namely the $u(\Sigma_0) \cup \{\infty\},\ u \in U$, where Σ_0 is the hyperplane of purely imaginary quaternions. We call the spheres $u(\Sigma) \cup \{\infty\}$, $u(\Sigma_0) \cup \{\infty\}$, and the unit spheres centered at Hurwitz integers of length 1 or $\sqrt{2}$ *special*. When we apply $E(z) = 1/z$ to any special sphere, we only get a translate of some special sphere by a Hurwitz integer. It follows that each chain $S(0, \infty, ...)$ is bounded by special spheres. This shows that Γ is chain-finite and that each tile $Z(F),\ F \in \mathcal{F}(0, \infty)$, is bounded by special spheres.

Unfortunately, the action of $\Gamma_{(0,\infty)}$ on $\mathcal{F}(0, \infty)$ is not free. If we partition $\mathcal{C}$ by cutting along the hyperplanes $u(\Sigma_0)$, each of the polyhedral pieces is a union of tiles $Z(F)$ for F in some subset of $\mathcal{F}(0, \infty)$. Now the only special spheres that contain $1/2$ are the hyperplanes $z_m = 0,\ m > 0$ and $z_0 = 0$. So there is a unique $F_0 \in \mathcal{F}(0, \infty)$ with $1/2 \in Z(F_0)$ such all the coordinates $z_m,\ m > 0$, are nonnegative on the tile $Z(F_0)$. Then $g_0(z) = u_0 z u_0^{-1}$ cyclically permutes the coordinates $z_m,\ m > 0$, and g_0 fixes $\mathbf{R}$ so, by uniqueness, $g_0(F_0) = F_0$.

Since the action of Γ on $\mathcal{F}$ is not free, one cannot form a reduced Markov partition for $\overline{\phi_t}$ just yet. There are two ways to proceed: one can either use a smaller group, and so settle for less, or find a finer Markov partition on which Γ acts freely.

We first produce a normal subgroup $\Gamma' \subset \Gamma$ of index three. The factor ring of $\mathcal{H}$ by the ideal generated by $j+k$ is a field $\mathbf{F}_4$ of order 4, with representatives $1, u_0, 1+u_0$, and 0. Reducing modulo this ideal gives a homomorphism $\rho : \Gamma \to Gl(2, \mathbf{F}_4)$ with $\rho(g_0) = u_0 I$. Let Γ' consist of those γ with $\det(\rho(\gamma)) = 1$. Then Γ/Γ' is isomorphic to the units group of $\mathbf{F}_4$ and $g_o \notin \Gamma'$.

The stabilizer of any $F \in \mathcal{F}$ is either trivial or conjugate to $< g_0 >$, as is seen by studying the action of $\Gamma_{(0,\infty)}$ on the polyhedral partition of $\mathcal{C}$ constructed above. It follows that Γ' acts freely on $\mathcal{F}$. Thus

Proposition 17.1. *The geodesic flow over H^5/Γ' has a reduced Markov partition. This flow is a 3-fold cyclic covering of $\overline{\phi_t}$.*

A finer Markov partition can be produced using a simplex reflection group Γ^* and marked chains, as in Appendix 3. Γ^* is defined as the group generated by Γ and two new generators $\alpha(z) = (j+k)z(j+k)^{-1}$ and $\beta(z) = \overline{z}$. Since α and β stabilize $\mathcal{Q}$, $\mathcal{H}$, and U, they normalize Γ. One can verify that Γ^* is the normalizer of Γ in the isometry group of H^5, but we will not need this fact. Since $\beta^2 = 1, \alpha\beta = \beta\alpha$, and $\alpha^2 \in \Gamma$, we see that Γ^*/Γ has order 4. Thus Γ^* is discrete. Since α and β fix ∞, Γ^* has the same Ford region R and the same set Q of cusped limit points as Γ.

We now show that a certain hyperbolic simplex $P^* \subset H^5$ is a fundamental domain for Γ^*. Since R is the Ford region for Γ^*, it suffices to take $P^* = R \cap \pi_\infty^{-1}(P)$, where $P \subset \mathcal{C}$ is a Euclidean simplex and a fundamental domain for Γ_∞ on $\mathcal{H}$. Since $\mathcal{C}$ is a fundamental domain for T, it is enough that P be a fundamental domain for the action of $\Gamma^*_{(0,\infty)}$ on $\mathcal{C}$. Since $\Gamma^*_{(0,\infty)}$ acts transitively on the faces of $\mathcal{C}$, it is enough that P lies in the cone $C\mathcal{C}^+ = \{tz : z \in \mathcal{C}^+, 0 \leq t \leq 1\}$ and is a fundamental domain for the stabilizer $\Gamma^*_{\mathbf{R}}$ of this cone in $\Gamma^*_{(0,\infty)}$. $\Gamma^*_{\mathbf{R}}$ is the subgroup of Γ^* that fixes $\mathbf{R}$. It contains α, β, and the group $\{uzu^{-1} : u \in U\}$ of order 12, so its order is at least 48. But $\Gamma^*_{\mathbf{R}}$ acts effectively on $\mathcal{C}^+$, which has only 48 symmetries. Thus $\Gamma^*_{\mathbf{R}}$ is the symmetry group of the regular octahedron $\mathcal{C}^+$, which has a simplex P_0 as fundamental domain. We find that the cone $P = CP_0$ is a simplex and a fundamental domain for the action of Γ^*_∞ on $\mathcal{H}$. We may define $P_0 \subset \mathcal{C}^+$ by the inequalities $0 \leq z_1 \leq z_2 \leq z_3$, so that $P = \{z \in \mathbf{H} : 0 \leq z_1 \leq z_2 \leq z_3,\ z_1 + z_2 + z_3 \leq z_0 \leq 1/2\}$.

But one can check that
(a) $\Gamma^*_{\mathbf{R}}$ is generated by reflections in the hyperplanes $z_1 = 0,\ z_2 = z_1$, and $z_3 = z_2$,
(b) $\Gamma^*_{(0,\infty)}$ is generated by $\Gamma^*_{\mathbf{R}}$ and the reflection $u_0\overline{z}$ in the hyperplane $z_1 + z_2 + z_3 = z_0$,
(c) Γ^*_∞ is generated by $\Gamma^*_{(0,\infty)}$ and the reflection $1 - \overline{z}$ in the hyperplane $z_0 = 1/2$, and
(d) Γ^* is generated by Γ^*_∞ and the inversion $1/\overline{z}$ that fixes the face $\partial R \cap \pi_\infty^{-1}(P)$ of P^*.
Thus Γ^* is a simplex reflection group, as desired.

Now Cor. 13.2 and its marked version from Appendix 3 apply. Using Lemma A3.2, we see that the marked tiles in P are bound by the special spheres and the faces of P. We find

Theorem 17.2. *There is a reduced Markov partition for the geodesic flows over H^5/Γ^* and*

H^5/Γ. *Each marked tile in* P *determines one Markov box in* SH^5/Γ^* *and four Markov boxes in* SH^5/Γ.

One can intersect the Markov boxes for $\phi_t : SH^5 \to SH^5$ with $SH(v)$ to get a $\Gamma(v)$-invariant flowbox family for the geodesic flow over $H(v)$. When $|v_1|$, $|v_2|$, and $|v_3|$ are distinct and nonzero, one can show that these flowboxes are nonoverlapping and give a Markov partition for $\psi(v)_t$ with a finite number of Markov boxes. For other v, however, $H(v)$ contains some tile boundaries and the flowboxes overlap. In these cases one may apply an inclusion/exclusion construction like that of §9 to produce a Markov partition with a finite number of Markov boxes.

The geodesic flow over H^5/Γ^* is just the billiard flow over the hyperbolic simplex P^*. The real points in P correspond to the *modular triangle* in P^* with angles $\pi/2$, $\pi/3$, and 0, c.f. §15. The billiard flow over this triangle was first described symbolically by Artin [Ar].

Section 18. Unit groups

Our final examples arise in all dimensions $n \geq 2$. Fix a nondegenerate quadratic form $f(x_0, ..., x_n)$ of index 1 with integer coefficients. A *unit* of f is an automorphism of the integer lattice Λ that preserves f. We use the projective model of H^n. If $w \in \mathbf{R}^{n+1}$ is timelike, so $f(w) < 0$, then w defines a point $[w] \in H^n$. If q is lightlike, so $q \neq 0$ and $f(q) = 0$, then $[q] \in \partial H^n$. Let $U(f)$ be the group of units of f and let $\Gamma = U(f)/\pm I$. Then Γ is a discrete group of isometries of H^n. These groups were studied for $n = 2$ and $n = 3$ by Fricke and Klein [FK], pp. 501-584. The case $n = 3$ is treated in Chapter 10 of [EGM], where it is shown that Γ is commensurable with a Bianchi group and where many examples are computed. Other examples, for various n, can be found in [Cas], Chapter 13 and pp. 166-7, and [CS], Chapters 27 and 28. When $f(x_0, x_1, x_2) = x_0^2 - x_1 x_2$, Γ can be identified with the group $PGl(2, \mathbf{Z})$ by identifying each point w in the upper halfplane with the timelike vector $(\Re(w), |w|^2, 1)$.

H^n/Γ has finite volume, as was shown for $n = 2$ and $n = 3$ by Fricke and Klein [FK] and in general by Venkov [Ve]. This was extended by Siegel to nondegenerate f of arbitrary index, with H^n is replaced by an appropriate symmetric space, and further generalized by Borel and Harish-Chandra in their reduction theory for arithmetic groups [Si], [BH]. $[q] \in \partial H^n$ is cusped for Γ if and only if q is a multiple of an integral vector. We assume such integral lightlike vectors exist, so H^n/Γ is noncompact.

Let $< v, w >$ be the bilinear pairing with $< w, w >= f(w)$. If q is lightlike we let $B(q) \subset H^n$ denote the horoball based at $[q]$ defined by $< w, q >^2 + < w, w > \leq 0$. We say this horoball is *rational* if q is a rational vector. Clearly the image of a rational horoball by a unit of f is a rational horoball. When a unit fixes $[q]$, it must send q to $\pm q$ and so it preserves $B(q)$. Thus we may may denote each cusped limit point $[q]$ so that all q are rational and lie in the same component of the lightcone and so the collection of all $\pm q$ is invariant by $U(f)$. Taking the q's sufficiently large, the rational horoballs $B([q]) = B(q)$ form a Γ-invariant horoball packing. We have, much as shown for Bianchi groups in [F3],

Theorem 18.1. *Each* Γ*-invariant packing by rational horoballs is chain-finite. Thus the action of* $U(f)$ *on in* H^n *has a reduction theory based on finitely many rectangles.*

Note that the hyperplane corresponding to the linear space $(p - q)^\perp$ contains the ridge $R([q], [p])$ for $([q], [p]) \in A$. This is easy to see in a coordinate system where $f(y) = -y_0^2 +$

$y_1^2 + ... + y_n^2$, $p = (1, 1, 0, ..., 0)$, and $q = (1, -1, 0, ..., 0)$. If two ridge shadows $S([q], [p])$ and $S([r], [p])$ share a face C (recall they are convex polyhedra in the metric d_p on $\partial H^n - \{p\}$) then $R([r], [p]) \cap R([q], [p])$ lies on the codimension two subspace corresponding to $(p-q)^\perp \cap (p-r)^\perp$. This subspace lies in the hyperplane $P \subset H^n$ corresponding to the linear space $N^\perp$, where $N =< r, p > (q-p) - < q, p > (r-p)$. $\partial P \subset \partial H^n$ passes through $[p]$ and so it contains C. Let R be the reflection in P. By Lemma 3 of [F3] and Thm. 13.1, the theorem will hold if the group $\Gamma' = R\Gamma R^{-1}$ is commensurable with Γ.

But R is the projective transformation induced by the linear transformation $\sigma(w) = w - (2 < w, N > / < N, N >)N$, the *symmetry in* N, which preserves f [Cas]. An element of $\Gamma \cap \Gamma'$ is given by a unit of f that stabilizes $\sigma(\Lambda)$. Since N is rational, $\sigma(\Lambda) \cap \Lambda$ has finite index in Λ. It follows easily that $\Gamma \cap \Gamma'$ has finite index in Γ and the theorem is proved.

The last paragraph proved a special case of a general fact about any algebraic group G defined over $\mathbf{Q}$, namely that the commensurator of $G_{\mathbf{Z}}$ contains the group $G_{\mathbf{Q}}$ of rational points. For our purposes, G is the orthogonal group of f, $G_{\mathbf{Z}} = U(f)$, and the rational points of interest are reflections in shadow boundaries.

Thm. 18.1 explains the structure of the conjugacy classes in the group $\Gamma = U(f)/\pm I$. For each hyperbolic element $\gamma \in \Gamma$ stabilizes an axis and so defines a closed geodesic on H^n/Γ. But the marked versions of Thm. 13.1 and Cor. 11.4 (see App. 3) give a correspondence $\underline{\mathcal{M}}/\Gamma$ on a finite set whose periodic orbits determine the closed geodesics in H^n/Γ. These periodic orbits are nearly in 1-1 correspondence with the hyperbolic conjugacy classes in Γ. Some exceptions arise because a closed geodesic may be described by more than one periodic orbit of $\underline{\mathcal{M}}/\Gamma$. However such a closed geodesic corresponds to a flowline that does not meet the interior of any Markov box. The only other exceptions arise when a geodesic in H^n has a stabilizer that is not cyclic, in which case the closed geodesic corresponds to a finite number of conjugacy classes. Roughly speaking, however, a generic conjugacy class in Γ corresponds to one periodic orbit of $\underline{\mathcal{M}}/\Gamma$.

Of course, a similar discussion can be given for other geometrically finite groups. In [F2], the known conjugacy class structure of triangle groups was used to construct symbolic dynamics.

Section 19. Open problems

(19.1) Given a geometrically finite group Γ, not convex cocompact, and an invariant horoball packing, find a shadow family that is chain-finite. Appendix 2 shows that ridge shadows are not sufficient in general. The case $n = 2$ is treated in §14. Take a ridge shadow family and successively alter it in some way to approach a chain-finite shadow family.

(19.2) Suppose $n = 2$ and one is given a standard Γ-invariant chain-finite shadow family as in Thm. 14.1. By Lemma 6.6, the slices $A(F)$ are intervals in L. Find their endpoints by a finite procedure. For Hecke triangle groups, §15 shows that each endpoint is fixed by a hyperbolic element of Γ. Show this is always the case.

(19.3) Compare and contrast the Markov partition for (p, q, ∞) of §15 with the one found in [F2]. The latter can be derived from the invariant checkerboard tiling of H^2 by ideal regular p-gons and q-gons whereas the former is derived from a dual structure, namely an invariant packing of H^2 by horoballs. For the case $(2, 3, \infty) = PGl(2, \mathbf{Z})$, the symbolic dynamics of [F2] is that of Artin [Ar], based on simple continued fractions, whereas the symbolic dynamics of this paper is based on least-remainder continued fractions. There is a rule for passing from a simple continued fraction to a least-remainder continued fraction, based on the identity $m + (1/(1 + (1/x))) = (m + 1) - (1/(x + 1))$, which shows that each least-remainder convergent is a simple convergent (see [A]). Find such rules for other values of p and q. For each Hecke triangle group, Rosen found a continued fraction theory that generalizes simple continued fractions [Ro]. Relate his convergents to those of this paper (aimed sequences with $p_0 = \infty$ for the ridge shadow family) by rules with a geometric basis.

(19.4) Program a computer to calculate slices. To fix ideas, take a finite volume quotient of H^3, an invariant horoball packing, and the ridge shadow family. Although the slice transform of §12 involves an infinite union, it seems that only finitely many arcs contribute to the boundary of each set $T^m A^\epsilon(F)$. One expects a fractal limit rather like the Koch snowflake curve. One would particularly like to generate images of the slices $A(F)$ for Bianchi groups, since these play a role in the reduction theory of [F3]. Show these slices are topological cells.

(19.5) Evaluate the dynamical zeta functions of a Bianchi group B_k using the symbolic dynamics of this paper. Using Cor. 11.4 and a method of Ruelle, see [R] or [F1], one can express such a zeta function as a finite alternating product of Fredholm determinants, provided that B_k acts freely on $\mathcal{F}$. The first term in this product is defined using transfer operators for the contracting holomorphic correspondence on the disjoint union of the tiles $Z(F)$, $F \in \mathcal{F}/B_k$, which is defined like the Markov family of interval maps in Figure 5. To calculate the kth term one needs to know which families of k distinct Markov boxes $M(F)$ have nonempty intersection. For $k > 1$, therefore, one needs some control over the slices $A(F)$, as in our previous question.

(19.6) The symbolic dynamics for the Hurwitz modular group gives a symbolic dynamics for the Bianchi-like groups $\Gamma(v)$, as sketched in §17. Estimate the number of Markov boxes needed for each of the geodesic flows $\psi(v)_t$ in terms of $|v|$.

(19.7) Extend the results of §17 to H^9 using integral octaves [Co].

(19.8) Suppose H^n/Γ, $n \geq 3$, has finite volume and some invariant horoball packing is chain-finite. Find a formula for the volume of the Markov boxes $M(F)$. One can compute these volumes with dilogarithms when $n = 2$ and the shadow family is standard. Expressing the volume of the unit sphere bundle SH^2/Γ as the sum of these volumes gives dilogarithm identities like those of [F2]. For $n = 3$, find corresponding trilogarithm identities.

(19.9) Apply the symbolic dynamics of Cor. 11.4 to the ergodic theory of the geodesic flow on $\mathcal{R}/\Gamma$. Produce a theory of equilibrium states comparable to the one known in the convex cocompact case through the work of Bowen and Ruelle [BR] and use this to describe the statistical distribution of closed geodesics. Interpret the measure $\overline{\nu}$ of Theorem 11.2 as an equilibrium state.

(19.10) Extend the theory of this paper to a complete simply connected Riemannian manifold whose sectional curvatures are bounded between two negative constants. Complex hyperbolic space is particularly important since it has interesting discrete groups of isometries with finite volume noncompact quotient. Show some of these groups have invariant chain-finite horoball packings.

(19.11) Find a theory of symbolic dynamics for suitable Anosov flows on noncompact manifolds that applies to the geodesic flows over finite volume quotients of hyperbolic space. In the compact quotient case, hyperbolic dynamics alone can be used to produce symbolic dynamics [B4], but here one should also assume some properties of the flow outside a large compact set. Such a theory might apply to Anosov geodesic flows over noncompact manifolds with small regions of positive curvature. Anosov geodesic flows over compact manifolds were studied by Eberlein [Eb].

(19.12) Does the theory of this paper extend to the product of two hyperbolic planes? Since the Hilbert modular group of a real quadratic field acts discretely on this space, this would have applications to number theory. Likewise, does the theory extend to the Teichmuller geodesic flow [V], replacing the fan tiling by Harer's tiling of Teichmuller space ([Ha], §2)? This would have applications to mapping class groups. These questions go well beyond (19.10) since they concern spaces of nonpositive curvature.

Appendix 1. Null sets for geometric measure

We will show that certain subsets of ∂H^n must be either full or null for the geometric measure class of a geometrically finite group Γ. The crucial property of such a geometric measure μ is that its Cartesian square restricts to a nontrivial ergodic measure on Ω, as we noted in Section 11. As this property holds for some groups that are not geometrically finite ([Ni], Thm. 8.3.5), we will work with a more general class of discrete groups.

To begin, we consider a locally compact σ-compact metric space X with a positive Borel measure μ. Suppose G is a countable group of homeomorphisms of X such that G preserves the family of null sets. We give the space $X \times X$ the Borel measure $\mu \times \mu$ and we let Λ be the open subset consisting of ordered pairs of distinct elements of X. We take ν to be a Borel measure on Λ equivalent to that obtained by restricting $\mu \times \mu$ (recall that measures are equivalent if they have the same null sets.) We assume that ν is nonzero (that is μ is not supported at one point) and that ν is ergodic for the natural action of G on Λ.

Let us define a *G-partition* of the G-space X to be a nonempty, countable family of nonnull Borel sets $X_j,\ j \in J$, that are essentially disjoint and that are essentially permuted by G. That is $X_j \cap X_{j'}$ is null for $j,\ j' \in J,\ j \neq j'$, and for each $g \in G$ and $j \in J$ there is some $j' \in J$ for which the symmetric difference $g(X_j)\ \Delta\ X_{j'}$ is null. Since j' is clearly unique, we obtain an action of G on J. If J contains only one element we say that the G-partition is trivial.

Lemma A1.1. *If our G-partition is nontrivial then the action of G on J is doubly transitive and the measure on each X_j is supported at one point.*

To prove this, we intersect Λ with the two sets $\cup\ X_j \times X_{j'},\ j,\ j' \in J,\ j \neq j'$, and $\cup\ X_j \times X_j,\ j \in J$. Each of the resulting sets is essentially invariant by G and they are essentially disjoint, so our ergodicity hypothesis implies that one of them is null. If the first set is null then J has only one element so we assume that the second set is null. Then ν vanishes on each term $(X_j \times X_j) \cap \Lambda$ and it follows easily that the measure on X_j is supported at one point. The ergodicity of ν implies that the permutation action of G on distinct pairs in J is transitive. But this is just the condition that the action of G on J be doubly transitive.

Suppose $X = \partial H^n$ and G is a discrete group of hyperbolic isometries with its natural action on X. Recall that G is *elementary* if it stabilizes a nonempty finite set in $H^n \cup \partial H^n$.

We prove

Lemma A1.2. *If $\mathcal{O}$ is an orbit of G on ∂H^n and G is is doubly transitive on $\mathcal{O}$ then G is elementary.*

Take $z \in \mathcal{O}$ and let G_z be its stabilizer, so G_z is elementary. We may suppose $G_z \neq G$ and choose $w \neq z, w \in \mathcal{O}$. Then double transitivity implies that the orbit of w by G_z is $\mathcal{O} - \{z\}$. If G_z acts properly discontinuously on $\partial H^n - \{z\}$ then either $\mathcal{O}$ is finite or its derived set $\mathcal{O}'$ equals $\{z\}$ and in either case G is elementary. Otherwise G_z acts properly discontinuously on $\partial H^n - \{a, z\}$ for some $a \neq z$ ([Ra], §5.5) and one may repeat the same argument to prove the lemma.

For the rest of this appendix, G denotes a nonelementary discrete group of isometries of H^n, $X = \partial H^n$, and μ is a positive Borel measure on Λ such that the measure class of μ is invariant and the action of G on Λ is ergodic. We do not assume that μ is a geometric measure but the reader may wish to do so. We now prove a dissipative property of μ.

Proposition A1.3. *Let Y be a Borel set in ∂H^n that is neither null nor full. There is an infinite chain of finite subsets $G(0) \subset G(1) \subset \ldots$ of G such that the sets $S_i = \cap\, g(Y), g \in G(i)$, are essentially decreasing, that is the differences $S_i - S_{i+1}$ are not null.*

One constructs these finite sets recursively. It is enough to show that there is some $g \in G$ so that neither $Y \cap gY$ nor $Y - gY$ is null. We will assume there is no such g and show that G is elementary, proving the proposition.

If $g \in G$ and $Y \cap gY$ is not null then $Y \cap g^{-1}Y$ is not null either. By our assumption, both $Y - g^{-1}Y$ and $Y - gY$ are null, hence $gY - Y$ is null, too. Let the *essential stabilizer* of Y in G be the subgroup G_Y of G consisting of those g such that gY is essentially equal to Y (that is, the symmetric difference of Y and gY is null). We have just shown that if $g \in G - G_Y$ then $Y \cap gY$ is null.

It follows that g_1Y and g_2Y are essentially equal or essentially disjoint according to whether $g_1^{-1}g_2$ is in G_Y or not. Thus the sets gY, for g varying over a set of coset representatives for G/G_Y, form a G-partition. As Y is not full and as μ is ergodic, Y is not essentially invariant so the G-partition is not trivial. By Lemma A1.1 $\mu|Y$ is supported on a point $y_0 \in Y$ and G is doubly transitive on the orbit of y_0. By Lemma A1.2, G is elementary, so the proposition is proved.

We now derive several corollaries.

Corollary A1.4. *μ has no atoms.*

For if $Y - gY$ and $Y \cap gY$ are not null then Y is not an atom.

In order to prove that the action of a geometrically finite group on Λ is ergodic, Sullivan showed that no cusped limit point is an atom for geometric measure [S2]. By Cor. A1.4, the same holds for radial limit points.

The following is a measure version of Lemma 1.1. It is used in $S11$.

Corollary A1.5. *If Y is a subsphere of ∂H^n then Y is null or full for μ. In the latter case, $L(G) \subset Y$.*

For suppose Y is neither null nor full and choose $G(i)$ and S_i as in Prop. A1.3. For each i, S_i is a subsphere of dimension between 1 and $n-1$ or it contains at most two points.

Thus the strictly decreasing chain S_i is finite, which contradicts Prop. A1.3. Thus Y is null or full.

Suppose Y is full. Then for $g \in G$, gY is full hence so is $Y' = Y \cap gY$. If $Y' = Y$ for all $g \in G$, then Y is invariant. If $Y' \neq Y$ for some g then Y' cannot be a point, by Cor. A1.4, and so Y' is a subsphere of lower dimension than Y. We choose a full subsphere Y^* of Y of least dimension and we find that Y^* is invariant. But then $L(G) \subset Y^*$ ([Ra], Thm. 12.1.4) and so $L(G) \subset Y$ and the corollary is proved.

Cor. A1.5 implies, paradoxically, that the countable dense family of circles in the Appolonian gasket must be a null set for the geometric measure associated to its symmetry group. We note that Cor. A1.5 includes Lemma 1 of [Ru].

Corollary A1.6. *If Y is a real analytic subset of ∂H^n then Y is null or full for μ. In the latter case, $L(G) \subset Y$.*

The proof is almost the same as for Cor. A1.5, since a strictly decreasing chain of real analytic sets is finite.

Appendix 2. Some chain-infinite examples

We will construct finite volume quotients H^3/Γ with one cusp such that Γ is not chain-finite. We first derive some of the consequences of a chain-finite invariant horoball packing.

Lemma A2.1. *Suppose a horoball packing $B(q)$, $q \in Q$, of H^n is Γ-invariant, the volume of H^n/Γ is finite, and the ridge shadow family $S(q,p)$, $(q,p) \in A$, is chain-finite. Suppose $n \geq 3$ and let Σ be a sphere that contains a smooth component of the boundary of some shadow. Then Σ is stabilized by a hyperbolic element of Γ.*

To see this, we refine the shadow family to the tile family $Z(F)$, $F \in \mathcal{F}$, of §9. Each $Z(F)$ is bounded by finitely many $(n-2)$-spheres S and we consider the set $\mathcal{P}$ of all such pairs (F, S). As $\mathcal{F}/\Gamma$ is finite we find that $\mathcal{P}/\Gamma$ is also finite. By our assumption, $(F_0, \Sigma) \in \mathcal{P}$ for some $F_0 \in \mathcal{F}$. We can iteratively construct an allowed sequence $F_i \in \mathcal{F}$ so that $(F_i, \Sigma) \in \mathcal{P}$ for $i = 0, 1, 2, \ldots$ There is a $\gamma \in \Gamma$ such that $\gamma(F_j, \Sigma) = (F_k, \Sigma)$ for some $k > j$. It follows that $\gamma(Z(F_j)) \subset Z(F_j)$. If we choose $(q,p) \in A$ so that $F_j \in \mathcal{F}(q,p)$, then γ contracts $Z(F_j)$ in the metric d_p. Thus γ is hyperbolic and stabilizes Σ.

Suppose $n = 3$. Select three cusp points p, q, and r and choose an upper half-space model of H^3 for which these points correspond to ∞, 0, and 1, respectively. Let $k = k\Gamma$ be the smallest subfield of $\mathbf{C}$ containing all the finite cusp points. This is called the *invariant trace field* of Γ ([MR], §5.5) and it is independent of our choice of three cusp points. Clearly k is not a subfield of $\mathbf{R}$. If $(az+b)/(cz+d) \in \Gamma$ then the ratios of a, b, c, and d lie in k.

If $k \cap \overline{k} \subset \mathbf{R}$ we say that the field k is *asymmetric*. This holds, for instance, if k has odd degree over $\mathbf{Q}$ (remarkably this degree is finite, see [MR], Thm. 3.3.7). We now show that asymmetry implies that some 1-spheres have small stabilizers.

Lemma A2.2. *Let Σ be a 1-sphere containing ∞ defined by an equation $\Re(z) = \sigma$ and suppose that Σ meets $k = k\Gamma$. Let Λ be the stabilizer of Σ in Γ. If k is asymmetric then Λ is finite.*

For asymmetry implies that the only purely imaginary number in k is 0, so Σ meets k in one point and meets $\hat{k} = k \cup \{\infty\}$ in two points. But Λ preserves $\Sigma \cap \hat{k}$ and Γ_∞ has finite point-stabilizers, since it acts properly discontinuously on $\mathbf{C}$. The lemma follows.

We say that a triple of mutually adjacent cusp points $p,\ q$, and q' is *isosceles* if the ridges $R(q,p)$ and $R(q',p)$ share an edge and $B(p,q) = B(p,q')$. In an upper halfplane model with $p = \infty$ the latter condition becomes $h(B(q)) = h(B(q'))$. This equality of diameters holds if q and q' are Γ_p-equivalent. Under this mild symmetry of the fan tiling, an asymmetric trace field is inconsistent with chain-finiteness.

Proposition A2.3. *Let $B(q),\ q \in Q$ be a Γ-invariant horoball packing of H^3 such that H^3/Γ has finite volume. If there is an isosceles triple and if $k\Gamma$ is asymmetric then this packing is not chain-finite.*

This follows from the lemmas by taking $p = \infty,\ q = 0,\ q' = 1$, and $\sigma = 1/2$.

For instance, Prop. A2.3 applies to the hyperbolic knot complement $S^3 - K$ where K is the two bridge knot $(7/3)$, denoted 5_2 in the usual knot tables. The hyperbolic structure on this knot complement is described in [MR], §4.5. $S^3 - K$ is diffeomorphic to H^3/Π, where Π has two generators $u(z) = z + 1$ and $v(z) = z/(\zeta z + 1)$ corresponding to the Wirtinger generators for the two bridges and where ζ is a root of $x^3 + x^2 + 2x + 1 = 0$. The field $k\Pi = \mathbf{Q}(\zeta)$ has degree 3, so it is asymmetric. The triple $\infty,\ 0, 1$ is isosceles, as can be checked by calculating the Ford region for this group. As H^3/Π has only one cusp, Π is chain-infinite.

Appendix 3. Marked sequences

It is useful to refine the theory of this paper by considering aimed sequences that are decorated by certain markings. For instance, say $n = 2$ and $p_i,\ i \in \mathbf{Z}$, is aimed at z. Let J_i be the interval containing z with endpoints p_i and p_{i+1} and mark p_{i+1} by a dot on the side containing J_i to form a *marked sequence.* Here each cusp point p admits two markings, one for each side of p.

In general, say $B(p),\ p \in Q$, is a horoball packing and $\epsilon : \underline{Q} \to Q$ is a mapping whose level sets $\epsilon^{-1}(q)$ are finite and nonempty. We use the simplified notation $\underline{q}$ for an element of $\underline{Q}$ with $\epsilon(\underline{q}) = q$, and we call $\underline{q}$ a *marking* of q.

Fix a shadow family $S(q,p),\ (q,p) \in A$, and choose a subset $\underline{A} \subset \underline{Q} \times Q$ and nonempty compact sets $S(\underline{q},p) \subset S(q,p)$ for $(\underline{q},p) \in \underline{A}$ so that the $S(\underline{q},p)$ with $\epsilon(\underline{q}) = q$ form a cover of $S(q,p)$ by essentially disjoint, nonoverlapping sets. We say the $S(\underline{q},p)$ are a *marked shadow family.* The marked shadows then satisfy marked versions of (3.1)-(3.3) and (B), where A is replaced by $\underline{A}$ and (q,p) by $(\underline{q},p)$.

From a marked shadow family we can construct a symbolic dynamics for ϕ_t. When ϵ is bijective, this will just be the theory developed earlier using aimed sequences. We need some definitions. A *marked chain* is a nonempty set of the form $S(\underline{p}_l, ..., \underline{p}_{k+1}, p_k) = \cap S(\underline{p}_i, p_{i-1})$, where $k < l$ and $k < i \leq l$, in which case we say that $p_k, \underline{p}_{k+1}, ..., \underline{p}_l$ is a *marked segment.* A *marked sequence* aimed at z is a sequence $p_i,\ i \in I$, aimed at z and markings $\underline{p}_i$ for $i > \inf(I)$ such that $z \in S(\underline{p}_i, p_{i-1})$ for all $i > \inf(I)$. A review of the theory of §3 − 5 shows that all the results can be extended to marked sequences. Continuation and Existence extend without difficulty. Marked versions of Uniqueness and Thm. 3.5 also hold.

We define $K(\underline{r}, q)$ to consist of the limit pairs of marked sequences with $p_0 = q$ and $\underline{p}_1 = \underline{r}$. Then $K(r,q)$ is the union of the sets $K(\underline{r}, q)$ over all markings of r. Indeed $(z,a) \in K(\underline{r}, q)$ if and only if $(z,a) \in K(r,q)$ and $z \in S(\underline{r}, q)$. This implies that the marked version of Prop. 5.2 holds.

Passing to §7, we define a flowbox $\Phi(\underline{q}, p)$ as the union of the $I(z,q,p,a)$ over $(z,a) \in K(\underline{q}, p)$. The union of these flowboxes for fixed $(q,p) \in A$ is $\Phi(q,p)$. A marked version of Thm. 7.3 holds provided that the level sets of ϵ are bounded. Prop. 7.4 easily extends to the flowbox cover $\Phi(\underline{q}, p)$, $(\underline{q}, p) \in \underline{A}$.

We can describe marked sequences with correspondences, much as in §8. Let $\underline{\mathcal{L}}(\underline{q}) = L - \{q\}$ for $\underline{q} \in \underline{Q}$. Let $\underline{\mathcal{L}} = \coprod \underline{\mathcal{L}}(\underline{q})$, so a point of $\underline{\mathcal{L}}$ is a pair $(\underline{q}, z) \in \underline{Q} \times L$ with $z \neq q$. Let $\underline{\mathcal{A}}(\underline{q}, \underline{p}) = S(\underline{q}, p) - \{q\} \subset \mathcal{A}(q,p)$ for $(q,p) \in A$. Let $\underline{\mathcal{A}} = \coprod \underline{\mathcal{A}}(\underline{q}, \underline{p})$, $(\underline{q}, p) \in \underline{A}$, so a point in $\underline{\mathcal{A}}$ is a triple $(\underline{q}, \underline{p}, z) \in \underline{Q} \times \underline{Q} \times L$ with $z \neq q$ and $z \in S(\underline{q}, p)$. The *marked correspondence* determined by ϵ and the marked shadow family is the self-correspondence $\underline{\mathcal{A}} \to \underline{\mathcal{L}} \times \underline{\mathcal{L}}$ that sends $(\underline{q}, \underline{p}, z)$ to $((\underline{p}, z), (\underline{q}, z))$. A $\mathbf{Z}$-orbit of this correspondence is a sequence $(\underline{p}_i, \underline{p}_{i-1}, z)$, $i \in \mathbf{Z}$, such that $\underline{p}_i$, $i \in \mathbf{Z}$, is a marked sequence aimed at z. Thus the inverse limit of $\underline{\mathcal{A}}$ can be identified with the space of marked sequences indexed by $\mathbf{Z}$.

We were able to interpret the aimed correspondence as a mapping on $\mathcal{L}$ since its first component function is nearly bijective. The same is not true for the marked correspondence, but it may be interpreted as a partially defined mapping on $\underline{\mathcal{L}}$ since its first component function is nearly injective.

There is a forgetful map $\epsilon \times \epsilon \times id : \underline{\mathcal{A}} \to \mathcal{A}$ that omits the markings. This defines a map of inverse limits $\widehat{\underline{\mathcal{A}}} \to \widehat{\mathcal{A}}$ that we compose with V to get a mapping $\underline{V} : \widehat{\underline{\mathcal{A}}} \to \mathcal{R}$. The marked version of Prop. 8.1 then holds, so $\underline{V}$ is an orbit-equivalence from the marked correspondence to ϕ_t.

We say that a symmetry group Γ is marked if it acts on $\underline{Q}$ so that ϵ is equivariant, $\underline{A}$ is preserved by Γ, and $\gamma(S(\underline{q}, p)) = S(\gamma(\underline{q}), \gamma(p))$ for all $\gamma \in \Gamma$. When Γ is a marked symmetry group we can form the reduced marked correspondence $\underline{\mathcal{A}}/\Gamma \to \underline{\mathcal{L}}/\Gamma \times \underline{\mathcal{L}}/\Gamma$. The marked version of Thm. 8.3 then holds, so the reduced marked correspondence is orbit-equivalent to $\overline{\phi}_t$ when Q is null and $\underline{Q} \times (L - Q)$ is Q-free for Γ. We now show that the latter condition can be achieved by a suitable marking in all cases of interest.

Proposition A3.1. *Suppose given a horoball packing with cusp set Q such that L has no seams and each subsphere that does not contain L is null. Given a shadow family and a symmetry group Γ, there is a family of markings and a marked shadow family so that Γ is marked and both $\underline{Q} \times (L - Q)$ and $\underline{A}$ are Q-free for Γ.*

Let $N \subset \Gamma$ be the finite normal subgroup of elements that fix Q, so $\mathrm{Eff}(\Gamma) = \Gamma/N$. For each $q \in Q$ there is a finite index subgroup $H_q \subset \Gamma_q$ so that $L - \{q\}$ is Q-free for Γ_q and $N \subset H_q$. For if S is the smallest subsphere of ∂H^n that contains Q, the group Γ_q/N is a discrete group of isometries of the Euclidean space $S - \{q\}$. Some finite index subgroup of Γ_q/N acts freely on $L - \{q\}$, for instance the free abelian group of [Ra], Thm. 5.4.5. We let H_q be the corresponding subgroup of Γ_q.

Choose Q' as in §8. Let $\underline{Q} = \coprod \Gamma/H_q$, $q \in Q'$, and let $\epsilon(\gamma H_q) = \gamma(q)$. Then Γ acts on $\underline{Q}$, ϵ is finite to one and equivariant, and $\underline{Q} \times (L - Q)$ is Q-free for Γ. For each $q \in Q'$ we choose a set of representatives Q^q for the action of Γ_q on $\{p \in Q : (q,p) \in A\}$. Let $\underline{A} \subset \underline{Q} \times Q$

consist of all pairs $(\gamma(H_q), \gamma(p))$, $\gamma \in \Gamma$, $q \in Q'$, $p \in Q^q$. The stabilizer in Γ of (H_q, p) is $H_q \cap \Gamma_p$, which acts trivially on Q. Thus $\underline{A}$ is Q-free for Γ.

It only remains to choose the marked shadows. By equivariance, it suffices to fix $q \in Q'$ and $p \in Q^q$ and to define $S(H_q, p)$ appropriately. Take S as above and note that the finite group $G = \Gamma_{(q,p)}/N$ is a group of isometries of the Euclidean space $S - \{p\}$. As L has no seams, we can choose a basepoint $r \in L - \{p\}$ on which G acts freely. Let $D(q, p)$ be the Dirichlet domain of r for G. Then $L \cap \partial D(q, p)$ has empty interior as L has no seams. $\partial D(q, p)$ is null as it lies in a finite union of proper subspheres that do not contain L. We define the marked shadows by $S(\gamma(H_q), \gamma(p)) = \gamma(S(q, p) \cap D(q, p)) \subset S(\gamma(q), \gamma(p))$. Since the cones $\gamma(D(q, p))$, $\gamma \in \Gamma_{(q,p)}/N$, are a nonoverlapping cover of $\partial H^n - \{p\}$ by essentially disjoint sets, we have defined a marked shadow family and the proposition is proved.

For S as above and $q \in Q$, Γ_q acts linearly on $T_q S$ and so defines a subgroup $G_q \subset Gl(T_q S)$. G_q is finite if the rank of the discrete Euclidean group Γ_q is $n-1$ or $n-2$. So when Γ is geometrically finite and $Q = Q(\Gamma)$, each G_q is finite provided either H^n/Γ has finite volume or $n \leq 3$.

When the G_q are finite, one may replace the coset markings used in the preceding proof with a more intuitive marking in which each $\underline{q}$ is a cone in $T_q S$. This is the sort of marking mentioned above for $n = 2$. For each $q \in Q'$, we choose a finite cover of $T_q S$ by nonoverlapping cones that are freely permuted by G_q. Let $\underline{Q}$ be the union over Q' of the Γ-orbits of these cones. The basepoint projection ϵ makes this a family of markings. When $(q, p) \in A$, let the *lune* $\lambda(\underline{q}, p) \subset S$ be the union of all the circular arcs joining q to p in S whose tangent vector at q lies in the cone $\underline{q}$. $\Gamma_{(q,p)}/N$ injects into G_q and so it freely permutes these lunes. Let $\underline{A}$ consist of all pairs $(\underline{q}, p)$ for which $S(\underline{q}, p) = \lambda(\underline{q}, p) \cap S(q, p)$ is nonempty. As in the preceding proof, this is a marked shadow family.

The Markov partitions of §9 and the symbolic dynamics of §10 can be generalized using marked sequences. Suppose a marked shadow family is chain-finite in the sense that there are only finitely many marked chains $S(\underline{q}, \underline{p}, ...)$ for each $(q, p) \in A$. It follows that the shadow family is also chain-finite. Using Prop. 9.2 on families of marked chains, one can define marked tiles $Z(\underline{F}) \subset S(\underline{q}, p)$, $\underline{F} \in \underline{\mathcal{F}}(\underline{q}, p)$. For each $\underline{F} \in \underline{\mathcal{F}}(\underline{q}, p)$ there is a unique $F \in \mathcal{F}(q, p)$ with $Z(F) \subset Z(\underline{F})$ and we call $\underline{F}$ a marking of F. We define $K(\underline{F}) = Z(\underline{F}) \times A(F) \subset K(\underline{q}, p) \cap K(F)$ and a Markov box $M(\underline{F}) \subset \Phi(\underline{q}, p) \cap M(F)$ by imposing the condition $(z, a) \in K(\underline{F})$. As each tile $Z(F)$ is the union of the marked tiles $Z(\underline{F})$ over all markings of F, marked versions of Thm. 9.3, Prop. 9.6, and Lemma 10.1 hold.

Give $\underline{\mathcal{F}} = \cup \underline{\mathcal{F}}(\underline{q}, p)$, $(\underline{q}, p) \in \underline{A}$, the discrete topology. Let $\underline{\mathcal{M}}$ be the discrete space consisting of all pairs $(\underline{F}', \underline{F})$ with $\underline{F} \in \underline{\mathcal{F}}(\underline{r}, q)$, $\underline{F}' \in \underline{\mathcal{F}}(\underline{q}, p)$, and $Z(\underline{F}) \subset Z(\underline{F}')$. We call $\underline{\mathcal{M}}$ the *marked Markov correspondence* on $\underline{\mathcal{F}}$. Define marked allowed sequences $\underline{F}_i$ by the conditions $(\underline{F}_i, \underline{F}_{i+1}) \in \underline{\mathcal{M}}$. We see that orbits of $\underline{\mathcal{M}}$ are marked allowed sequences with index set $\mathbf{Z}$, topologized as a subspace of the product space $\underline{\mathcal{F}}^{\mathbf{Z}}$. Define $\underline{V}$ on such orbits by $\underline{V}(\underline{F}_i) = V(F_i) \in \mathcal{R}$, that is by forgetting the markings and applying the mapping V of Thm. 10.2. The marked versions of Thm. 10.2 and Thm. 10.3 then hold provided ϵ has bounded level sets. When $\underline{A}$ is Q-free for Γ, as in Prop. A3.1, $\underline{\mathcal{F}}$ is also Q-free for Γ and we get a reduced Markov partition for $\overline{\phi}_t$ indexed by $\underline{\mathcal{F}}/\Gamma$.

Note that when Γ is geometrically finite and $Q = Q(\Gamma)$, the hypotheses of Prop. A3.1 hold. This permits us to apply this proposition in the proof of Thm. 11.2. Moreover Q/Γ

is finite so ϵ has bounded level sets. If in addition the marked shadow family is chain-finite then $\underline{\mathcal{F}}/\Gamma$ is finite, so the marked version of Cor. 11.4 holds.

Thm. 13.1 continues to hold for marked shadow families. In the setting of Cor. 13.2, one may define an invariant marked shadow family with $\underline{Q} = \Gamma$, $\epsilon(\gamma) = \gamma(0)$, $\underline{A} = \{(\gamma, p) : p = \gamma(\infty)\}$, and $S(\gamma, \gamma(\infty)) = \gamma(P)$ for all $\gamma \in \Gamma$. Since $\overline{P} \cup \{\infty\}$ is an intersection of Γ-balls, this marked shadow family is chain-finite. This is the marked version of Cor. 13.2. Γ acts freely on $\underline{\mathcal{F}}$, since it acts freely on $\underline{A}$, so we get a reduced Markov partition for $\overline{\phi}_t$. There is a very simple relation between tiles and marked tiles in this case.

Lemma A3.2. *Given $\underline{F} \in \underline{\mathcal{F}}(\underline{q}, p)$, $Z(\underline{F}) = Z(F) \cap S(\underline{q}, p)$.*

It is enough to show that each marked chain with nonempty interior $S(\underline{p}_l, \dots, \underline{p}_{k+1}, p_k)$ is equal to $S(\underline{p}_l, p_{l-1}) \cap S(p_l, \dots, p_k)$. This is proven by induction on $l - k$ and easily reduces to the case $l = 2,\ k = 0$. By equivariance, we may suppose $\underline{p}_1 = id \in \Gamma$ and $p_0 = \infty$. Setting $\gamma = \underline{p}_2 \in \Gamma$, we know that $\gamma(\infty) = 0$ and $\gamma(P) \cap P$ has nonempty interior and we must show $\gamma(P) \cap S(0, \infty) \subset P$.

Consider the reflections in $\Gamma_{(0,\infty)}$. Partition ∂H^n into lunes by deleting all the hyperspheres fixed by these reflections, taking connected components, and forming their closures. The lunes form a nonoverlapping cover of ∂H^n that is transitively permuted by $\Gamma_{(0,\infty)}$. Restricting this lune cover to $S(0, \infty)$ gives the family of polyhedra $g(P)$, $g \in \Gamma_{(0,\infty)}$, since $\Gamma_{(0,\infty)}$ is generated by the reflections in the faces of P through the origin. As ι normalizes $\Gamma_{(0,\infty)}$, it permutes the lunes. For each $h \in \Gamma_\infty$, no interior point of $h(P)$ is fixed by a reflection in Γ_∞, and so $h(P)$ is contained in a lune. Thus $\iota h(P) \cap S(0, \infty) \subset g(P)$ for some $g \in \Gamma_{(0,\infty)}$. Choose h so $\gamma = \iota \circ h$. Since $\gamma(P)$ overlaps P, we have $g(P) = P$ and the lemma follows.

§17 gives a case where Γ acts freely on $\underline{\mathcal{F}}$ but not on $\mathcal{F}$, so marked sequences produce a finite Markov partition for $\overline{\phi}_t$ even though aimed sequences do not.

References

[A] Adams, W.W. On a relationship between the convergents of the nearest integer and regular continued fractions. Math. Comp. 33 (1979) 1321-1331.

[Ad] Adler, R. Geodesic flows, interval maps, and symbolic dynamics. Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces. Oxford Univ. Press, 1991, 93-123.

[Ar] Artin, E. Ein Mechanisches System mit quasi-ergodischen Bahnen. Abh. Math. Sem. Univ. Hamburg 3 (1924) 170-175, or Collected Papers, Springer, Berlin, 1965, 499-504.

[As] Ash, A. Small dimensional classifying spaces for arithmetic subgroups of general linear groups. Duke Math. J. 51 (1984) 459-468.

[BH] Borel, A. and Harish-Chandra. Arithmetic subgroups of algebraic groups. Ann. of Math. 75 (1962) 485-535.

[B1] Bowen, R. On Axiom A Diffeomorphisms. CBMS Reg. Conf. 35, AMS, Providence, 1978.

[B2] Bowen, R. Markov partitions are not smooth. Proc. Amer. Math. Soc. 71 (1978) 130-132.

[B3] Bowen, R. Equilibrium States and the Ergodic Theory of Anosov Diffeomorphism. Springer-Verlag, Berlin, 1975.

[B4] Bowen, R. Symbolic dynamics for hyperbolic flows. Amer. J. Math. 95 (1973) 429-460.

[BR] Bowen, R. and D. Ruelle. The ergodic theory of Axiom A flows. Invent. Math. 29 (1975) 181-202.

[BS] Bowen, R. and C. Series. Markov maps associated with Fuchsian groups. Publ. Math. IHES 50 (1979) 153-170.

[Cas] Cassels, J.W.S. Rational Quadratic Forms. Academic Press, London, 1978.

[Ca] Cawley, E. Smooth Markov partitions and toral automorphisms. Erg. Th. Dyn. Syst. 11 (1991) 633-651.

[C] Conway, J.H. The Sensual (Quadratic) Form. Carus Monographs 26, MAA, 1997.

[CS] Conway, J.H. and Sloane, N.J.A. Sphere Packings, Lattices and Groups. Springer-Verlag, New York, 1988.

[Co] Coxeter, H.S.M. Geometry. Lectures on Modern Mathematics, Vol. III, T.L. Saaty, ed. Wiley, New York, 1965, 58-94.

[Eb] Eberlein, P. When is a geodesic flow of Anosov type? I, II. J. Diff. Geom. 8 (1973) 437-463, 565-577.

[E] Eichler, M. Grenzkreisgruppen und kettenbruchartige Algorithmen, Acta Arith. 11 (1965) 169-180.

[EGM] Elstrodt, J., Grunewald, F. and Mennicke, J. Groups Acting on Hyperbolic Space. Springer-Verlag, Berlin, 1998.

[EP] Epstein, D.B.A., and R.C. Penner. Euclidean decompositions of noncompact hyperbolic manifolds. J. Diff. Geom. 27 (1988) 67-80.

[FN] Fenchel, W. and J. Nielsen. Discontinuous Groups of Isometries in the Hyperbolic Plane. De Gruyter, Berlin, 2003.

[FK] Fricke, R. and F. Klein. Theorie der Automorphen Funktionen I. B.G. Teubner, Leipzig, 1897.

[F1] Fried, D. The zeta functions of Ruelle and Selberg, I. Ann. Scient. Ec. Norm. Sup. 19 (1986) 491-517.

[F2] Fried, D. Symbolic dynamics for triangle groups. Invent. Math. 125 (1996) 487-521.

[F3] Fried, D. Reduction theory over quadratic imaginary fields. J. Number Theory 110 (2005) 44-74.

[GS] Guillemin, V. and S. Sternberg. Geometric Asymptotics. Math. Surveys 14, AMS, Providence, 1977.

[Ha] Harer, J.L. The virtual cohomological dimension of the mapping class group of an orientable surface. Invent. Math. 84 (1986) 157-176.

[H] Hurwitz, A. Uber eine besondere Art der Kettenbruch-Entwicklung reeller Grossen. Acta Math. 12 (1889) 367-405, or Mathematische Werke, Birkhauser, Basel, 1933, Band II, 84-115.

[L] Leuzinger, E. An exhaustion of locally symmetric spaces by compact submanifolds with corners. Invent. Math. 121 (1995) 389-410.

[MR] Maclachlan, C. and A. Reid. The Arithmetic of Hyperbolic 3-Manifolds. Springer-Verlag, New York, 2003.

[MT] Matsuzaki, K. and M. Taniguchi. Hyperbolic Manifolds and Kleinian Groups. Oxford Univ. Press, 1998.

[M] Mayer, D. On the thermodynamic formalism for the Gauss map. Commun. Math. Phys. 130 (1990) 311-333.

[Me] Mendoza, E. Cohomology of PGl_2 over imaginary quadratic integers. Bonner Math. Schriften 128, 1979.

[Ni] Nicholls, P. The Ergodic Theory of Discrete Groups. LMS Lec. Notes 143, Cambridge Univ. Press, 1989.

[P] Patterson, S.J. The limit set of a Fuchsian group. Acta Math. 136 (1976) 241-273.

[Rag] Raghunathan, M.S. Discrete Subgroups of Lie Groups. Springer, New York, 1972.

[Ra] Ratcliffe, J. Foundations of Hyperbolic Manifolds. Springer-Verlag, New York, 1994.

[Ro] Rosen, D. A class of continued fractions associated with certain properly discontinuous groups. Duke Math. J. 21(1954), 549-563.

[Ru] Rudolph, D.J. Ergodic behaviour of Sullivan's geometric measure on a geometrically finite hyperbolic manifold. Erg. Th. Dyn. Syst. 2 (1982) 491-512.

[R] Ruelle, D. Zeta functions for expanding maps and Anosov flows. Inv. Math. 34 (1976) 231-242.

[Sa] Saper, L. Tilings and finite energy retractions of locally symmetric spaces. Comm. Math. Helv. 72 (1997) 167-202.

[Se] Series, C. Geometrical methods of symbolic coding. Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces. Oxford Univ. Press, 1991, 125-151.

[Si] Siegel, C.L. Einheiten quadratischer formen. Abh. Math. Sem. Hamburg 13 (1940) 209-239, or Gesammelte Abhandlungen II, No. 40.

[S1] Sullivan, D. The density at infinity of a discrete group of hyperbolic motions. Publ. Math. IHES 50 (1979) 171-202.

[S2] Sullivan, D. Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups. Acta Math. 153 (1984) 259-277.

[V] Veech, W. Dynamics over Teichmuller space. Bull. Amer. Math. Soc. 14 (1986) 103-106.

[Ve] Venkov, B.A. On arithmetical group of automorphisms of an indefinite quadratic form. Izv. Akad. Nauk SSSR 2 (1937) 139-170.

[VS] Vinberg, E.B. and Shvartsman, O.V. Discrete groups of motions of spaces of constant curvature. Geometry II, Encyclopedia Math. Sci. 29. Springer-Verlag, Berlin, 1993, 139-248.

[Vo] Vogtmann, K. Rational homology of Bianchi groups. Math. Ann. 272 (1985) 389-410.