J. DIFFERENTIAL GEOMETRY
34 (1991) 115–141

LIMIT VOLUMES OF HYPERBOLIC THREE-ORBIFOLDS

COLIN C. ADAMS

Abstract

We prove that the three smallest limit volumes for hyperbolic 3-orbifolds are $0.3053218\ldots$, $0.4444514\cdots$, and $0.4579827\ldots$. The corresponding unique orbifolds are given. We also show that an n-fold limit point of volumes of hyperbolic 3-orbifolds is bounded below by $(n-1/2)v/2$ where $v = 1.01494\ldots$ is the volume of an ideal regular tetrahedron in hyperbolic 3-space. Applications to the order of the isometry groups of hyperbolic 3-manifolds are also given.

1. Introduction

A hyperbolic 3-orbifold is the quotient of hyperbolic 3-space by a discrete group of isometries of hyperbolic 3-space. If the group has no elliptic isometries, the quotient will be a hyperbolic 3-manifold. In all that follows, we will assume that the manifolds and orbifolds are all orientable.

The set of volumes of hyperbolic 3-manifolds are known to be well-ordered by the work of Thurston and Jørgensen (cf. [8]). In particular, given any specified set of hyperbolic 3-manifolds, there is a smallest volume among the set of volumes of the elements in the set. Results on the smallest volumes for hyperbolic 3-manifolds have been obtained in [1], [2], and [7].

Similarly, it is accepted folklore that the volumes of hyperbolic 3-orbifolds are well-ordered. A previous result on small volumes for hyperbolic 3-orbifolds was obtained by Meyerhoff in [6], where he found the smallest volume orientable cusped orbifold. However, unlike what occurs for orientable manifolds, the volume of this cusped orbifold is not a limit of volumes of closed orbifolds. Hence, it remains to find the smallest limit volume for a hyperbolic 3-orbifold.

In this paper, we show that a certain noncompact orbifold which is a quotient of the Borromean rings complement is the unique orbifold with the smallest volume that is a limit of volumes. Its volume is $0.3053218\ldots$. We also find the unique orbifolds with the second

Received August 24, 1989 and, in revised form, February 20, 1990. Supported in part by NSF Grant DMS-8711495.

and third smallest limit volumes, those volumes being $0.4444514\ldots$ and $0.4579827\ldots$.

Restating this result in terms of group theory, we have determined the three nonelementary Kleinian groups of least finite volume which are themselves the geometric limits of finite volume Kleinian groups.

In order to prove this result, we first prove a theorem which shows that the limit of a sequence of volumes of hyperbolic 3-orbifolds is the volume of an orbifold from which a subsequence of the orbifolds can be obtained by Dehn filling. This theorem also was folklore. The proof of this theorem is the content of §2.

Let a cusp in an orbifold be called *nonrigid* if Dehn filling can be performed on this cusp, and otherwise call it *rigid*. In §3, we use methods similar to those in [2] in order to show that $\mathrm{vol}(O) \geq (n-1/2)v/2$ for any orbifold O with n nonrigid cusps where $v = 1.01494\ldots$ is the volume of an ideal regular tetrahedron in hyperbolic 3-space. Such an O has a volume which is an n-fold limit point.

§§4-7 are devoted to improving this lower bound in the case of one nonrigid cusp and result in finding the three smallest limit volumes for hyperbolic 3-orbifolds. In §8, we give some applications to the orders of the diffeomorphism groups of knots in the 3-sphere. Throughout the following, all real numbers given in decimal notation have been rounded off to the number of decimal places shown.

2. Limit volumes and Dehn filling

In this section, we prove the following theorem.

Theorem 2.1. *Let $\{O_i\}$ be a sequence of hyperbolic 3-orbifolds with volumes bounded by a constant α. Then there exists a subsequence, all the orbifolds of which come from Dehn filling the cusps of a single hyperbolic 3-orbifold O.*

The proof of Theorem 2.1 will be similar to the proof in the manifold case appearing in Chapter 5 of [8]. We first define the thin and thick parts of an orbifold O. Let x be any point in O, and let $l(x)$ be the length of the shortest loop in O which passes through x and lies on either the Euclidean boundary of a cusp or the Euclidean boundary of an equidistant neighborhood of the projection of the axis of a hyperbolic isometry. Let S_n be that subset of the singular set with angle $2\pi/n$. Let $m_n(x)$ be the length of the shortest loop in O which passes through x and lies on the boundary of an equidistant neighborhood of S_n.

Definition. $O_{[0,\varepsilon)} = \{x \in O: l(x) < \varepsilon$ or, if $1/n < \varepsilon$, then $m_n(x) < \varepsilon$ or $x \in S_n\}$. Let $O_{[\varepsilon,\infty)} = O - O_{[O,\varepsilon)}$.

Vertices in the singular set of a hyperbolic 3-orbifold can be one of four types. If $v(a, b, c)$ represents a vertex with three singular axes of angles $2\pi/a$, $2\pi/b$, and $2\pi/c$, then the possibilities are $v(2, 3, 4)$, $v(2, 3, 3)$, $v(2, 3, 5)$, and $v(2, 2, m)$ where $m \geq 2$.

Let Γ be the fundamental group of the orbifold acting as a discrete group of isometries on H^3. Let x' be a lift of a point x in O to a point in H^3. Let $\Gamma_\varepsilon(x')$ be the set of elements in Γ which move x' a distance less than ε. The Margulis Lemma states that there exists an ε independent of the particular group Γ such that $\Gamma_\varepsilon(x')$ must have an abelian subgroup of finite index.

Lemma 2.2. *If v_1 and v_2 are vertices in the singular set of a hyperbolic 3-orbifold O which are within ε of each other, then v_1 and v_2 are both type $(2, 2, n)$ vertices and they share the axis of order n.*

Proof. Let x be a point a distance less than $\varepsilon/2$ from each of v_1 and v_2. Lift x to a point x' in H^3. Then $\Gamma_\varepsilon(x')$ contains in its generating set elliptic isometries corresponding to each of the singular axes intersecting v_1 and v_2. Since $\Gamma_\varepsilon(x')$ cannot then be finite, there are two possibilities remaining.

The first possibility is that $\Gamma_\varepsilon(x')$ consists of parabolic and elliptic isometries which all share a fixed point on the sphere at ∞. However, as there is more than one singular axis through v_1, the corresponding elliptic isometries cannot share a fixed point on the sphere at ∞.

The second possibility is that $\Gamma_\varepsilon(x')$ consists of elliptic isometries of order two which are perpendicular to the axis of a hyperbolic isometry together with possible elliptic rotations about the axis of the hyperbolic isometry. Since v_1 and v_2 are vertices, there must in fact be elliptic rotations about the axis of the hyperbolic isometry, making v_1 and v_2 both type $(2, 2, n)$ vertices which share the axis of order n.

Proof of Theorem 2.1. Let ε be a positive constant which is less than the smaller of the Margulis constant or 0.1. Let O_i be any one of the orbifolds in $\{O_i\}$.

Define a *standard μ-ball* in $O_{i[\varepsilon,\infty)}$ to be an embedded ball of radius μ which, if it intersects the singular set in $O_{i[\varepsilon,\infty)}$, does so in a set of radii and diameters. Our goal is to cover all of $O_{i[\varepsilon,\infty)}$ by standard balls and tubes around the singular axes. Note that all of the vertices in the singular set of $O_{i[\varepsilon,\infty)}$ are pairwise a distance at least ε apart. Place a ball of radius $\varepsilon/4$ at each of the vertices in the singular set of $O_{i[\varepsilon,\infty)}$.

These balls are each embedded by the definition of $O_{i[\varepsilon,\infty)}$ and standard by an argument similar to the proof of Lemma 2.2. Cover that part of each singular axis of order m in $O_{i[\varepsilon,\infty)}$ which is not already covered by the $\varepsilon/4$-balls by a tubular neighborhood of the singular set of radius $\sinh^{-1}(m\varepsilon/40)$.

All of these tubes will be disjoint as follows. If two tubes corresponding to singular sets of orders m and n did share a point x, then isometries corresponding to rotation about these axes would move a lift x' of x a distance less than ε. Both of the rotations would be in $\Gamma_\varepsilon(x')$.

One possibility is that the two axes of rotations share a finite vertex. But this contradicts the fact that the tubes lie outside the $\varepsilon/4$-balls at the vertices.

It could also be that the two axes of rotation share an endpoint at ∞. However, for each of the cases where this could occur, the subgroup fixing that point at ∞ would contain a parabolic element moving x' a distance less than ε. By the definition of $O_{i[\varepsilon,\infty)}$, x would not be in $O_{i[\varepsilon,\infty)}$.

The last possibility is that the two axes of rotation are both order 2. However, the distance between them would be less than $\varepsilon/2$ and they would generate a hyperbolic isometry which moves x' a distance less than ε. Again, x would not be in $O_{i[\varepsilon,\infty)}$.

The circumference of each tube is $\pi\varepsilon/20$. We will cover the remainder of $O_{i[\varepsilon,\infty)}$ by balls of radius $\pi\varepsilon/40$ such that their centers stay a distance $\pi\varepsilon/40$ from each other and a distance $\varepsilon/4+\pi\varepsilon/80$ from the vertices in the singular set and such that their centers stay a distance $\pi\varepsilon/80$ from the tube boundaries.

Since we have found a covering of $O_{i[\varepsilon,\infty)}$ by tubes and balls, we can choose V to be a maximal set of points in $O_{i[\varepsilon,\infty)}$ such that the points are the centers of the $\pi\varepsilon/40$-balls in such a cover. The $\pi\varepsilon/80$-balls centered at the points in V are all embedded, disjoint from one another, disjoint from the $\varepsilon/4$-balls centered at the singular vertices and disjoint from the tubes of radius $\sinh^{-1}(m\varepsilon/40)$.

Since the volume of O_i is bounded by α, there is a bound on the number of disjoint embedded $\pi\varepsilon/40$-balls that we can have. Similarly, there is a bound on the number of $\varepsilon/4$-balls that we can have at the singular vertices and hence a limit on the number of singular vertices that we can have.

Additionally, there is a bound on how much tubing we can have, since the tubes contribute to the total volume.

Combinatorially, there is a only a finite number of ways that the tubing can fit together with the finite number of $\varepsilon/4$-balls at the singular vertices to form a neighborhood of the singular set. Since there are only a finite number of $\pi\varepsilon/40$-balls, and since there are only a finite number of ways to glue together a finite set of balls with the neighborhood of the singular set, there are only a finite number of $O_{i[\varepsilon,\infty)}$'s possible. Hence we can choose a subsequence of $\{O_i\}$ such that the corresponding $O_{i[\varepsilon,\infty)}$'s are all homeomorphic via homeomorphisms which preserve the singular sets.

In fact, since the set of possible gluing maps of the $\pi\varepsilon/40$-balls is compact, there is a convergent sequence of gluing maps for the $O_{i[\varepsilon,\infty)}$'s which converge geometrically to the gluing map for an orbifold $O_{[\varepsilon,\infty)}$.

Of the four types of vertices in the singular set, our choice of ε insures that the vertices of the first three types which occur in O_i also occur in $O_{i[\varepsilon,\infty)}$. Vertices of the fourth type with $m > 1/\varepsilon$ will be removed from O_i when we form $O_{i[\varepsilon,\infty)}$.

By definition, $O_{i[\varepsilon,\infty)}$ is obtained from O_i by removing:

(i) balls which are neighborhoods of edges in the singular set with order at least $1/\varepsilon$ (which is greater than 9);
(ii) balls which are neighborhoods of short edges of order m in the singular set connecting two $(2, 2, m)$ vertices;
(iii) balls which are neighborhoods of short geodesic edges perpendicular to and connecting two order-two singular axes;
(iv) solid tori which are neighborhoods of short geodesics;
(v) solid tori which are neighborhoods of geodesics in the singular set of order at least $1/\varepsilon$;
(vi) cusps.

All of the sets which we remove from O_i are either concentric or disjoint by the Margulis Lemma. As the lengths of the geodesics and geodesic edges which we remove from the sequence of orbifolds get shorter, the radii of their neighborhoods get larger. Similarly, as the orders on those pieces of the singular set which we remove get larger, the radii of the corresponding tubes get larger. Hence, there is a complete orbifold from which the sequence of orbifolds is obtained by doing Dehn filling on the cusps. q.e.d.

We state a lemma, the proof of which is contained in the proof above.

Lemma 2.3. *Let C be a nonrigid cusp in a hyperbolic 3-orbifold. Then the only singular axes that can intersect C are order-two axes which go directly out the cusp.*

3. Lower bounds on volumes

Let C be a cusp in a hyperbolic 3-orbifold or 3-manifold O. We expand C to a maximal cusp in the following sense. Lifting the cusp to H^3, we have a disjoint set of horoballs, all of which are identified by $\pi_1(O)$. Expand them equivariantly until two first touch. The projection of these expanded balls down to O is called a *maximal cusp*. For more details, see [1] and [2].

We work in the upper-half-space model of H^3, choosing ∞ to be a parabolic fixed point, and normalizing so that one of the horoballs covering C, denoted H_∞, is centered at ∞ and has boundary the Euclidean plane at height 1. The *center* of a horoball which is not H_∞ is defined to be that point where the horoball is tangent to the x-y plane. Horoballs covering C which are tangent to H_∞ are of Euclidean diameter one and are called *full-sized horoballs*. The subgroup of isometries in $\pi_1(O)$ which fix ∞ will consist only of parabolic and elliptic isometries. Call this subgroup G_∞. In the case that C is a nonrigid cusp, the elliptic isometries will all be of order 2. Let P_∞ be the subgroup of G_∞ consisting of all its parabolic isometries.

Lemma 3.1. *A nonrigid maximal cusp has volume at least* $\sqrt{3}/8$.

Proof. Let C be a nonrigid maximal cusp in O and normalize so that ∞ is a parabolic fixed point for a parabolic isometry corresponding to this cusp.

A fundamental domain F for the action of P_∞ acting on the x-y plane must contain the center of at least one full-sized horoball. By the circle-packing argument as in [1], F has an area of at least $\sqrt{3}/2$. A rotation of order 2 can lower the area by at most a factor of 2. Any two such rotations sharing a fixed point at ∞ have product a parabolic isometry, so a fundamental domain for the action of G_∞ on the x-y plane must have an area of at least $\sqrt{3}/4$. Hence the maximal cusp C has a volume of at least $\sqrt{3}/8$. q.e.d.

In the case that the orbifold has more than one cusp, we can improve this lower bound considerably.

Lemma 3.2. *Let C be a nonrigid maximal cusp in a hyperbolic orbifold with more than one cusp. Then* $\mathrm{vol}(C) \geq \sqrt{3}/4$.

Proof. Since the cusp C is maximal, it touches itself on its boundary. Lifting to the upper-half-space model of H^3, where ∞ is the center of a horoball covering this cusp, we find at least one full-sized horoball in the fundamental domain of the action of G_∞. The existence of the other cusp forces the existence of a disk of no tangency in this fundamental domain, which is equivalent in terms of the needed area to having a second full-sized ball in the domain. Hence, this forces the volume up to $\sqrt{3}/4$. (See the proof of Lemma 4.1 of [2] for more details.) Note that none of the other cusps which are present need to be nonrigid cusps. q.e.d.

Rather than maximizing the volume in a single cusp, we can maximize the sum of the volumes in a disjoint set of cusps.

Lemma 3.3. *Let O be a hyperbolic 3-orbifold which has n nonrigid cusps. Then there is a choice of cusps with disjoint interiors such that the total volume contained in the cusps is strictly greater than $(n-1/2)\sqrt{3}/4$.*

Proof. Starting with all cusps disjoint, expand the first cusp until it becomes maximal. Then expand each cusp in turn until it touches itself or it touches one of the previously expanded cusps.

Suppose now that after expansion, the union of the resultant cusps is not connected. Each component will contain at least one maximal cusp. For a given component, each of the other components generates a disk of no tangency on one of the cusps in the first component. Note that different disks of no tangency may overlap, so we are only assured of an extra $\sqrt{3}/8$ in volume per component. If there are k cusps in a given component, then that component has a volume of at least $(k-1/2)\sqrt{3}/4$ without counting the disk of no tangency and $k\sqrt{3}/4$ if we add in the disk of no tangency contribution. Hence, in this case, we get a volume in all the cusps of at least $n\sqrt{3}/4$.

Suppose now that the union of all the resultant cusps is connected. Then there are at least $n-1$ tangency points between distinct cusps. Each of these $n-1$ tangency points will contribute $\sqrt{3}/8$ to each of the volumes of the two cusps which are tangent there. Since the first cusp which was expanded contributes $\sqrt{3}/8$ to its own volume, we get a total volume in the cusps of at least $(n-1/2)\sqrt{3}/4$.

If the total volume was exactly $(n-1/2)\sqrt{3}/4$, then put a horoball corresponding to the first cusp at ∞. The resulting circle packing in the x-y plane corresponding to the projection of the horoballs must then be the hexagonal packing.

Suppose first that two full-sized horoballs corresponding to this first cusp touch each other, so the shortest translation length is one in this cusp. Then every full-sized horoball corresponding to a distinct cusp must touch a second full-sized horoball corresponding to the same cusp, the second ball being obtained by translating the first. But then the second cusp touches itself and is therefore maximal. This yields an extra point of tangency and forces the total volume in the cusps to be at least $n\sqrt{3}/4$.

So we may suppose that no two full-sized balls corresponding to the first cusp touch one another. Hence a full-sized ball corresponding to the first cusp must be surrounded by six full-sized balls which do not correspond to this first cusp. Pick two of them which are adjacent. If they both correspond to the same cusp, we again have a second cusp which is maximal, yielding an extra point of tangency. If the two adjacent horoballs correspond to distinct cusps, then the three cusps corresponding to these three horoballs are pairwise tangent. However, this will mean there are more than $n - 1$ points of tangency between distinct cusps, causing the volume to be at least $(n + 1/2)\sqrt{3}/4$. q.e.d.

Just as in the case of manifolds, where we can relate the volume in the cusps to the entire volume in the manifold (Lemma 2.1 of [2]), we have such a relation for orbifolds. Let v be the volume in an ideal regular tetrahedron in H^3, that is, 1.01494 to five decimal places.

Lemma 3.4. *If O is a hyperbolic 3-orbifold, then*

$$\mathrm{vol}(O) \geq (2v/\sqrt{3})\,\mathrm{vol}(C),$$

where C is the union of a set of cusps in O with disjoint interiors.

Theorem 3.5. *If O is a hyperbolic 3-orbifold with n nonrigid cusps, then* $\mathrm{vol}(O) > (n - 1/2)v/2$.

Proof. The proof is immediate from Lemmas 3.3 and 3.4. q.e.d.

In the case $n = 1$, the next few sections are devoted to finding the correct lower bound. In the case $n = 2$, we conjecture that the orbifold with volume approximately 3.6638/4, obtained by taking a quotient of the Whitehead link complement, is the hyperbolic 3-orbifold with two nonrigid cusps of least volume.

4. One cusp case: overview

Our goal in the next three sections will be to demonstrate the following improvement of Lemma 3.2 in the case of $n = 1$.

Theorem 4.1. *The volume of a maximal nonrigid cusp in a hyperbolic 3-orbifold is either* $1/4$, $\sqrt{7}/8$, $\sqrt{2}/4$ *or greater than* 0.3969.

Once we have proved Theorem 4.1, we will prove the following corollary.

Corollary 4.2. *The three smallest limit volumes for hyperbolic 3-orbifolds are* 0.3053218..., 0.4444574..., *and* 0.4579827... .

The proofs of Theorem 4.1 and Corollary 4.2 will appear in §7. The proof of Theorem 4.1 will require a careful analysis of the horoball diagrams which can occur. (See [4] for pictures of examples of horoball diagrams in the manifold case.) Basically, the idea is that if the horoballs in the diagram stay far apart, this forces the area of a fundamental domain for P_∞ up and hence the volume of the cusp up. But if the horoballs in the diagram get close to one another, then putting one of the horoballs at ∞ forces the other horoball to be a large horoball in the horoball diagram, again forcing the area and hence the volume up.

We will begin with some basic geometric facts. Throughout what follows, H_z will represent a horoball in the upper-half-space model of H^3, with center at the point z in the x-y plane.

Lemma 4.3. *Let H_X be a horoball of diameter 1, centered about the point x in the boundary plane and tangent to the horoball H_∞ centered about ∞. Let β be a geodesic with one endpoint at x and its other endpoint at y, some other point in the plane. Let μ be the geodesic from y to ∞. If d is the distance on ∂H_∞ from $H_x \cap H_\infty$ to $\partial H_\infty \cap \mu$, then $1/d$ is the distance on ∂H_X from $H_x \cap H_\infty$ to $\partial H_X \cap \beta$.*

Proof. Let Ω be a geodesic with one endpoint at y and such that its point which lies the farthest above the boundary plane is directly above x. Let g be the hyperbolic isometry given by a 180° rotation about Ω. Then $g(H_X)$ is a horoball centered at ∞, $g(H_\infty)$ is a horoball centered at x, $g(\beta)$ is a geodesic from y to ∞, and $g(\mu)$ is a geodesic from x to y. Since the hyperbolic distance from the top of H_X to Ω is $\ln(d/1)$, the hyperbolic distance from the top of $g(H_X)$ to Ω is $\ln(d/1)$. Hence if z is the diameter of $g(H_X)$, then $\ln(z/d) = \ln(d/1)$ and thus $z = d^2$. The distance from the top of $g(H_X)$ to $g(H_X) \cap g(\beta)$ on $\partial g(Hx)$ is the Euclidean distance divided by the height above the boundary plane, namely $d/d^2 = 1/d$. However this distance is isometric to the distance on ∂H_X from $H_x \cap H_\infty$ to $\partial H_X \cap \beta$.

Lemma 4.4. *Two tangent horoballs of radii r_1 and r_2 have centers a Euclidean distance $2\sqrt{r_1 r_2}$ apart.*

Proof. Apply the Pythagorean theorem. q.e.d.

In the next two lemmas, we assume that C is a cusp covered by a horoball H_∞ centered at ∞ with boundary plane given by $z = 1$. All of the horoballs which are discussed cover the cusp C. Both of the proofs are similar to the proof of Lemma 4.3

Lemma 4.5. *Let H_x be a horoball of diameter 1, centered about the point x in the boundary plane and tangent to the horoball H_∞ centered about ∞. Let H_y be a horoball of diameter a which is less than 1, centered at the point y in the boundary plane. If H_x and H_y are not tangent and if the distance from x to y is less than 1, then there exists a horoball covering C with diameter greater than a but less than 1.*

Lemma 4.6. *Let H_x and H_y be two horoballs which are not tangent which have Euclidean diameters a and b respectively. Let c be the distance between their centers. Then there exists a horoball with Euclidean diameter ab/c^2.*

We assume from now on that C is a nonrigid maximal cusp in a hyperbolic 3-orbifold O. There is a set of horoballs in H^3 with disjoint interiors which covers C. All of the horoballs that we mention from now on are in this set. We again assume that the horoball H_∞, which is centered at ∞ and has boundary the horizontal plane at height 1, projects to C. The subgroup of $\pi_1(O)$ which fixes ∞ is denoted G_∞ and its parabolic subgroup is denoted P_∞.

Let d be the translation length of an element T_d of P_∞ and suppose $d > 1$. Then for a given tangency point on ∂H_∞ corresponding to a full-sized horoball H_x, there is a pair of tangency points generated by T_d and T_d^{-1} such that each of the pair is a distance d from the original tangency point on ∂H_∞ and such that the three tangency points all lie in a line. Since there exist elements of $\pi_1(O)$ identifying any two horoballs covering C, there is an element which sends H_x to H_∞. This isometry will then send H_∞ to a full-sized horoball H_y. The two points of tangency on H_∞ generated by T_d and T_d^{-1} force the existence of two tangency points a distance d from the top of H_y along the boundary of H_y (where distance is being measured in the induced Euclidean metric on the horosphere). That is, there must be two smaller horoballs on opposite sides of H_y such that they are tangent to H_y and such that the points of tangency are exactly a distance d along the full-sized horosphere from the topmost point of that horosphere. It again must be the case that the three centers of the horoballs all lie in a line. We call these smaller horoballs *$1/d$-balls* as it follows from Lemma 4.3 that their centers are a distance $1/d$ from the center of H_y. The Euclidean diameter of a $1/d$-ball is $1/d^2$ by Lemma 4.4.

Since every full-sized horoball is the image of H_∞ under an isometry of the group which sends some other full-sized horoball to H_∞, every full-sized horoball has a pair of adjacent $1/d$-balls. Note that if we draw the line segment from the center of one $1/d$-ball adjacent to H_x through the center of H_x and to the center of the other adjacent $1/d$-ball, the resulting line segment will be parallel to the line segment obtained by applying the same process to the $1/d$-balls adjacent to H_y.

Lemma 4.7. *If $\mathrm{vol}(C) < \sqrt{3}/4$, then P_∞ must identify all full-sized horoballs in the cusp diagram and there must be a fundamental domain for*

P_∞ which is a parallelogram with elliptic isometries of order two from G_∞ corresponding to rotations about vertical geodesics above each of the four vertices and above the midpoints of each pair of vertices of the parallelogram.

Proof. By Lemma 2.3, all of the elliptic isometries which occur in G_∞ must be order two. The proof of Lemma 3.1 shows that a fundamental domain for P_∞ can contain the center of at most one full-sized horoball. In order that $\mathrm{vol}(C) < \sqrt{3}/4$, G_∞ must properly contain P_∞. Hence G_∞ is the semidirect product of $Z + Z$ with Z_2. The lemma then follows. q.e.d.

Lemma 4.8. *If* $\mathrm{vol}(C) < \sqrt{3}/4$, *then every horoball which is not full-sized is tangent to a larger horoball.*

Proof. Suppose H_x is smaller than a full-sized ball and is not tangent to any larger horoball. Then there are no points of tangency anywhere in the interior of its upper hemisphere. This upper hemisphere is a disk of radius 1 on ∂H_x. Sending H_x to H_∞ by a group element, we have a disk of radius 1 on ∂H_∞ such that it contains no points of tangency with full-sized horoballs. This keeps the full-sized horoballs apart in the same way that an extra full-sized horoball which did have a tangency point at the center of the disk would. This forces the volume in the cusp to be at least $\sqrt{3}/4$. (See the proof of Lemma 4.1 of [2] for more details.) q.e.d.

From now on, we will assume all horoballs which are smaller than the full-sized balls are tangent to larger balls.

Lemma 4.9. *If* $\mathrm{vol}(C) < \sqrt{3}/4$, *then each point of tangency between two horoballs covering* C *is the triple intersection point of three mutually orthogonal axes of elliptic isometries of order two in* $\pi_1(O)$.

Proof. Let H_x be a full-sized horoball tangent to H_∞ at the point w. Lemma 4.7 implies that there is a vertical geodesic through w corresponding to an order two elliptic isometry. Since $\pi_1(O)$ identifies all horoballs covering C, there is an element g of $\pi_1(O)$ which sends H_x to H_∞. Since H_∞ touches H_x, g sends H_∞ to a full-sized ball. Since all full-sized balls are identified by P_∞, there is a parabolic element p fixing ∞ which identifies $g(H_\infty)$ with H_x. Thus, pg switches H_x and H_∞ and fixes w. Hence pg must be an elliptic rotation about a geodesic which is tangent to the two horoballs and which passes through w. So we now have two elliptic isometries with perpendicular axes passing through w. Their product generates the third elliptic isometry. q.e.d.

Define the *minimum tangency length* on the boundary of a maximal cusp to be the shortest length on the surface of a horoball between points of tangency. By Lemma 4.7, the minimum tangency length is realized by a parabolic isometry when we have a maximal nonrigid cusp C such that $\mathrm{vol}(C) < \sqrt{3}/4$.

Let d be the minimum translation length of a parabolic isometry fixing ∞ and let e be the minimum translation length of a parabolic isometry fixing ∞ which is linearly independent from the first.

5. Tangency length one

The first situation that we will deal with is the case that the minimum tangency length is 1.

Lemma 5.1. *If the minimum tangency length is* 1, *and the cusp volume is less than* $\sqrt{3}/4$, *then* $\pi_1(O)$ *contains a Fuchsian subgroup which is isomorphic to the modular group.*

Proof. In the case that the minimum tangency length is 1, there are three horoballs which are pairwise tangent, one of them centered at ∞, one of them a full-sized horoball, and the third a translate of the second. The three points of tangency between these horoballs must be identified by the group $\pi_1(O)$. The pair of points of tangency on the surface of a given one of the three horoballs must be identified by a parabolic isometry fixing the center of that horoball. Any two of the three corresponding parabolic isometries generate a rotation of order three about the point which is equidistant from the three points of tangency in the plane that they define. This order three rotation together with the parabolics generates the modular group acting on this plane. The rotations of order three in the plane are the restrictions of elliptic rotations of order three about a geodesic perpendicular to the plane. Similarly, there are order two rotations about geodesics perpendicular to the plane and through the points of tangency.

Lemma 5.2. *If* $d = 1$ *and* $e = 1$, *then* $\mathrm{vol}(C) = 1/4$.

Proof. Let H_x be a full-sized ball. Let T be a translation of length 1 corresponding to d and Q a translation corresponding to e. Then Lemma 5.1 implies that there must be an axis of rotation of order two which lies in the plane defined by x and the translation T and which passes through the point of tangency between H_x and H_∞. It must be that $Q(H_x)$ is sent by this rotation to a full-sized ball on the other side of H_x. But since all full-sized balls on the other side of H_x are of the form $Q^{-1}T^n(H_x)$, it must be the case that the angle between the T translation and the Q translation is either $\pi/3$ or $\pi/2$.

In the case that it is $\pi/3$, we have a full-sized ball centered at the end of the geodesic around which we were going to do order three rotations.

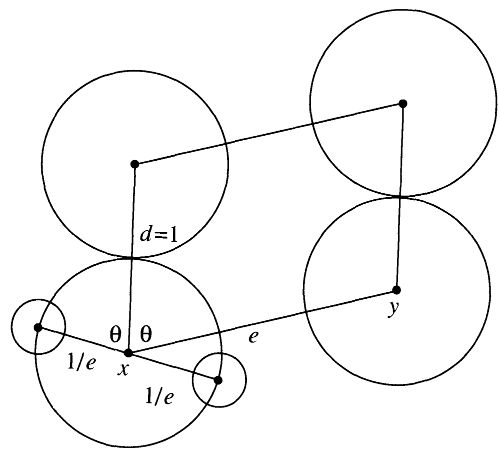


FIGURE 1

The resulting cusp would therefore not be a nonrigid cusp. Hence the only possible angle is $\pi/2$. This yields a cusp volume of $1/4$. q.e.d.

We are now interested in the situation where the maximal nonrigid cusp has volume less than $\sqrt{3}/4$, $d = 1$ and $e > 1$. By Lemma 5.1, $\pi_1(O)$ contains a Fuchsian subgroup isomorphic to the modular group. Let P be the geodesic plane preserved by this subgroup. Lemma 4.9 states that if w is a point of tangency between H_∞ and a full-sized ball H_x such that w is in P, then it must be the case that three elliptic axes of order two pass through w, one of which is vertical. Lemma 5 forces one of the remaining two axes to lie in the plane P while the third axis is perpendicular to the plane. Rotating about the axis in the plane will send a full-sized horoball H_y with center a distance e from the center of H_x to a ball on the other side of H_x which is tangent to H_x. By Lemma 4.3, the center of this new ball is a distance $1/e$ from x. We call this new ball a $1/e$*-ball.* Note that its diameter is $1/e^2$ by Lemma 4.4 (see Figure 1).

Let θ be the angle which is no greater than $\pi/2$ between the translations corresponding to d and e. Note that since e is the shortest translation distance linearly independent from d, it must be the case $\cos(\theta) \leq 1/(2e)$. Note that θ is also the angle between the edge from the center of H_x in the direction corresponding to the translation d and the edge from the center of H_x to the center of the $1/e$-ball. Since the length e corresponds to a translation, the inverse of that translation induces the existence of a second $1/e$-ball touching H_x on the opposite side from the first $1/e$-ball. Thus, every full-sized ball touches two $1/e$-balls such that the centers of the three balls are in a line.

Lemma 5.3. *If $d = 1$, $e > 1$, and two full-sized balls a distance at least e apart share a $1/e$-ball such that the centers of these three balls lie in a line, then* $\mathrm{vol}(C)$ *is either* $\sqrt{2}/4$ *or* $\sqrt{7}/8$.

Proof. First, suppose that two full-sized balls which are a distance exactly e apart share a $1/e$-ball such that the three centers are in line. Then $2/e = e$, so $e = \sqrt{2}$. Let θ be the angle between the d and e translations where θ is chosen to be at most $\pi/2$. If $\theta = \pi/2$, then such a cusp exists, with cusp volume of $\sqrt{2}/4$. If θ does not equal $\pi/2$, the rotations which take full-sized balls to $1/e$-balls create an additional pair of $1/e$-balls. Hence, each full-sized ball is touched by four distinct $1/e$-balls. In order that there be room for the two $1/e$-balls on one side of a full-sized ball, it must be that θ is at most $\arccos(1/(2\sqrt{2}))$. This angle of θ is realized, yielding a cusp volume of $\sqrt{7}/8$. This angle is the least angle that θ could be, since $\cos(\theta) \leq 1/2e$ whenever $d = 1$.

Suppose now that two full-sized balls which are a distance f apart, where f is greater than e, share a $1/e$-ball such that the three centers are in a line. Then f must be the third shortest distance between centers of full-sized horoballs after d and e. Because the angle which the f translation makes with the d translation is different from the angle which the e translation makes with the d translation, there are at least four $1/e$-balls touching each full-sized horoball. Hence there must be four full-sized balls with centers at a distance e from the center of a given full-sized ball, forcing $f = e$, a contradiction.

Lemma 5.4. *If $d = 1$, $e > 1$, and each full-sized ball is touched by four $1/e$-balls, then* $\mathrm{vol}(C)$ *is either* $\sqrt{7}/8$ *or at least* $\sqrt{11}/8$.

Proof. Let H_x be a full-sized ball. The fact there are four $1/e$-balls implies there are four full-sized balls with centers at a distance e from x. This forces $\cos(\theta) = 1/(2e)$.

Let H_y be a full-sized ball with center a distance e from the center of H_x. Then the center of one of the $1/e$-balls touching H_x must be in line with the center of H_y and the center of one of the $1/e$-balls touching H_y. Note that the fact $\cos(\theta) = 1/(2e)$ forces the two $1/e$-balls on one side of H_x to touch one another.

If there is only a single $1/e$-ball being shared by H_x and H_y, then we are in the situation considered in Lemma 5.3, yielding a cusp volume of $\sqrt{7}/8$.

If there are two $1/e$-balls in between H_x and H_y and they do not touch each other, they will generate an intermediate sized ball if their centers are not a distance at least $1/e$ apart, by Lemma 4.6. Since the $1/e$-balls

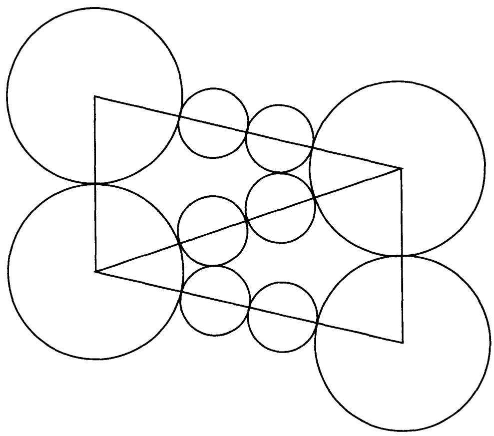


FIGURE 2

are the largest balls tangent to the full-sized balls, this would imply the existence of a ball tangent to no larger ball. Its upper hemisphere would then form a disk of no tangency, as in the proof of Lemma 4.1 of [2], forcing a cusp volume of at least $\sqrt{3}/4$. Thus, we will assume that the centers of the two $1/e$-balls are at least a distance $1/e$ apart. Therefore, $e \geq 3/e$ giving $e \geq \sqrt{3}$. This gives $\mathrm{vol}(C) \geq \sqrt{11}/8$.

If there are two $1/e$-balls between H_x and H_y, and the $1/e$-balls touch each other, then $e = 1/e + 1/e + 1/e^2$. Hence $e = (1+\sqrt{5})/2$, and $\mathrm{vol}(C) = \sqrt{(5+2\sqrt{5})}/8$. (See Figure 2 for a picture of the cusp diagram in this case.)

In fact, this last case cannot occur for the following reason. Let H_a and H_b be two tangent full-sized balls. They each have a point of tangency with H_∞. As in Lemma 5.1, there exists an elliptic isometry of order three which rotates about a geodesic, the points of which are equidistant from these three points of tangency. Rotating about this geodesic in one direction will send a to b, b to ∞, and ∞ to a. This rotation also sends d to c and c to g. To preserve points of tangency, e must go to f. This forces H_f to go to a ball which is tangent to H_d, H_e, H_f, and H_g. However, when we apply an isometry which takes this new ball to ∞, the four $1/e$-balls will be sent to four full-sized balls. These four full-sized balls will form a sequence such that any adjacent pair in the sequence are tangent. However, the centers of the four balls will not be collinear. This contradicts the pattern of full-sized balls which is present in this cusp diagram.

Lemma 5.5. *If $d = 1$ and two full-sized balls which are a distance at least e apart share a $1/e$-ball such that the centers of the three balls do not lie in a line, then* $\text{vol}(C) \geq \sqrt{11}/8$.

Proof. The rotations which take full-sized balls to $1/e$-balls cannot send the full-sized balls a distance e away to the $1/e$-balls mentioned in the hypotheses. Thus, the rotations create an additional pair of $1/e$-balls on each full-sized ball. Hence, each full-sized ball is touched by four distinct $1/e$-balls. Lemma 5.4 states that $\text{vol}(C) = \sqrt{7}/8$ or is at least $\sqrt{11}/8$. But the cusp diagram for the case when $\text{vol}(C) = \sqrt{7}/8$ does not fit the hypotheses of this lemma.

Lemma 5.6. *If $d = 1$, and two full-sized balls with centers a distance* 1 *apart share a $1/e$-ball, then* $\text{vol}(C) \geq \sqrt{3}/4$.

Proof. Suppose not. Since $1/e$-balls come in pairs which have their centers in a line with the center of the full-sized ball they touch, it must be the case that either $2/e = 1$ and a single $1/e$ ball has its center directly beneath the point of tangency of the two full-sized balls which it is tangent to or there are at least four $1/e$ balls touching each full-sized ball. In the first case, we have $e = 2$, which since $\cos(\theta) \leq 1/(2e)$, gives a volume in the cusp of at least $\sqrt{15}/8$. A volume of $\sqrt{15}/8$ occurs for the group $\text{PGL}_2(O_{15})$. (Note that in this case, the balls which we are calling $1/e$-balls in fact correspond to translations d^2 rather than to e, although the length of d^2 and e are the same.)

In the second case, we have four $1/e$-balls tangent to each full-sized ball. Because of the rotations of order two creating the correspondence between e-balls and $1/e$-balls, there must be four e-balls which are in the same directions from our full-sized ball as the four $1/e$ balls. This can only occur if a $1/e$-ball is shared by four full-sized balls. This means $e = \sqrt{2}$. If fact, e will not be the shortest translation linearly independent from the translation corresponding to d so this case cannot occur.

Lemma 5.7. *If $d = 1$, $e > 1$, and each $1/e$ ball touches one full-sized ball and the $1/e$-balls do not touch each other, then* $\text{vol}(C) \geq 0.3969$.

Proof. If a $1/e$-ball had center a distance less than 1 from a full-sized ball, Lemma 4.5 would imply that there was a ball intermediate in size between the full-sized balls and the $1/e$-balls. Since the $1/e$-balls are the largest balls tangent to the full-sized balls, this would imply the existence of a ball tangent to no larger ball. Its upper hemisphere would then form a disk of no tangency, as in the proof of Lemma 4.1 of [2] which would give a cusp volume of at least $\sqrt{3}/4$. Thus, each $1/e$-ball must stay a distance at least one from all the full-sized balls which it does not touch (see Figure 3).

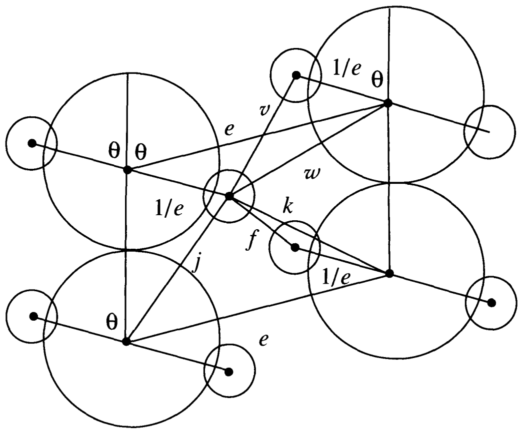


FIGURE 3

Hence the fact $j \geq 1$ and $w \geq 1$ imply respectively

(1) $$\cos(\theta) \leq 1/(2e),$$

(2) $$\cos(\theta) \geq (1/2)(\sqrt{}(3 - e^2 - 1/e^2)).$$

The fact $k \geq 1$ implies either

(3) $$\cos(\theta) \leq (1/4)(e + 1/e - \sqrt{}(10 - 3/e^2 - 3e^2))$$

or

(4) $$\cos(\theta) \geq (1/4)(e + 1/e + \sqrt{}(10 - 3/e^2 - 3e^2)).$$

(1) and (2) together force $e \geq \sqrt{2}$. (1) and (4) never hold simultaneously for $e > 1$, hence (3) instead of (4) must hold. Then comparing (3) with (2) shows that $e > 1.539$. This forces $\text{vol}(C) \geq 0.3746$. We can improve this lower bound.

We now look at the distances v and f. Since we are assuming the pairs of $1/e$-balls that v and f are measuring the distance between are not touching, each pair must stay apart a distance at least $1/e$ by Lemma 4.6. The fact $v \geq 1/e$ yields

(5) $$\cos(\theta) \geq \sqrt{}((4 - e^2 - 3/e^2)/8).$$

The fact $f \geq 1/e$ forces either

(6) $$\cos(\theta) \geq (1/8)(e + 2/e + \sqrt{}(28 - 7e^2 - 20/e^2))$$

or

(7) $$\cos(\theta) \leq (1/8)(e + 2/e - \sqrt{}(28 - 7e^2 - 20/e^2)).$$

(1) and (6) never hold simultaneously for $e \geq \sqrt{2}$, hence (7) instead of (6) must hold for $e \geq \sqrt{2}$. Comparing (7) and (5), we find that $e^8 - 7e^4 - 4e^2 + 12 \geq 0$, giving $e \geq 1.6126$ and $\text{vol}(C) \geq 0.3969$.

Lemma 5.8. *If $d = 1$, $e > 1$, each $1/e$ ball touches one full-sized ball, and the $1/e$-balls do touch each other, then* $\text{vol}(C) = (1 + \sqrt{5})/8$.

Proof. Suppose first that the two $1/e$-balls which are separated by a distance v touch each other. So $v = 1/e^2$. This forces

$$\cos(\theta) = \sqrt{(\tfrac{1}{8}(4 + 1/e^4 - e^2 - 4/e^2))}. \tag{8}$$

But comparing this with equation (2), we find $e \geq (1+\sqrt{5})/2$. Since if $e > (1 + \sqrt{5})/2$, the value inside the $\sqrt{\ }$ in (8) would be negative, it must be that $e = (1 + \sqrt{5})/2$ is the only possible e. Then $\theta = \pi/2$ and this yields a cusp volume of $(1 + \sqrt{5})/8$.

Suppose now that $f = 1/e^2$. Then

$$\cos(\theta) = (1/8)(e + 2/e \pm \sqrt{(28 - 7e^2 - 28/e^2 + 8/e^4)}). \tag{9}$$

By comparison with (1), we see that the correct formula when $e \geq \sqrt{2}$ is the one with the $-\sqrt{\ }$. When $e \geq \sqrt{2}$, (1) and (9) are both satisfied only in the range $\sqrt{2} \leq e \leq (1 + \sqrt{5})/2$. But by comparing (9) and (3), we see that $e \geq (1 + \sqrt{5})/2$. Hence the only possibility is $e = (1 + \sqrt{5})/2$ and $\cos(\theta) = 1/(1 + \sqrt{5})$. This example occurred in the proof of Lemma 5.4, where we showed that a nonrigid cusp with volume less than $\sqrt{3}/4$ could not have such a horoball pattern.

6. Tangency length greater than one

Let d be the shortest tangency length corresponding to a parabolic isometry fixing ∞, and let e be the shortest tangency length corresponding to a parabolic isometry which is linearly independent from the first. We are now interested in the case when $1 < d \leq e$.

We may again assume that the second to largest balls which appear in the horoball picture are themselves tangent to full-sized balls.

Lemma 6.1. *If $1 < d \leq e$, and the $1/d$-balls are each tangent to only one full-sized horoball, then* $\text{vol}(C) \geq (\sqrt{7} + 4)/16$.

Proof. Suppose $\text{vol}(C) < \sqrt{3}/4$. Let v_1, v_2, v_3, and v_4 be the four vertices of a parallelogram which forms a fundamental domain in the x-y plane for the action of the parabolic subgroup of $\pi_1(O)$ which fixes ∞. For convenience, choose coordinates so that v_1 occurs at the origin of the plane, v_2 occurs on the positive x-axis, and the parallelogram lies above the x-axis.

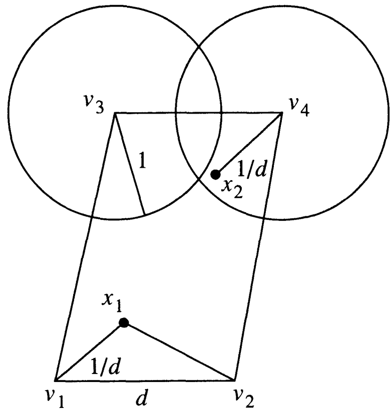


FIGURE 4

The four vertices of this parallelogram must all stay at least a distance d apart. Corresponding to the $1/d$ balls, there are two parallel line segments coming out of two opposite vertices into the parallelogram, each segment of length $1/d$. Let v_1 and v_4 be the vertices the line segments come out of. Call their endpoints x_1 and x_2. By Lemma 4.4, x_1 and x_2 must each stay at least a distance 1 from the three vertices of the parallelogram that each is not attached to. Since the vertices v_1, v_2, and x_1 form a triangle with edge lengths d, $1/d$, and an edge of length at least 1, the height h of x_1 above the x-axis is at least $(1/d)(1-(d^2+1/d^2-1)/2)^2)^{1/2}$. Additionally, since we have disks of radius 1 with centers a horizontal distance d apart which cannot intersect x_1, the edge from x_3 to x_4 must be a vertical distance above x_1 at least as large as $(1-(d/2)^2)^{1/2}$. Hence, it must be the case that

$$(1) \qquad \mathrm{vol}(C) \geq (d/4)(\tfrac{1}{d}(1-(d^2+1/d^2-1)/2)^2)^{1/2} + (1-(d/2)^2)^{1/2}.$$

At $d = 1$, this yields a cusp volume of $\sqrt{3}$. We need only examine this function up to $d = \sqrt{2}$, since the volume of the cusp is at least $d^2\sqrt{3}/8$ whenever $e \geq d$. For $1 \leq d \leq \sqrt{2}$, this function is decreasing. At $d = \sqrt{2}$, we obtain $\mathrm{vol}(C) \geq (\sqrt{7}+4)/16$. q.e.d.

Additional analysis could improve this lower bound. In particular, we did not use the fact that the $1/d$-balls must either touch or stay apart a distance at least $1/d$.

Now suppose a $1/d$-ball does touch more than one full-sized ball. Then there are two possible cases. Either there are two $1/d$-balls per full-sized ball or four $1/d$-balls per full-sized ball. In the second case, it must be that $d = e$.

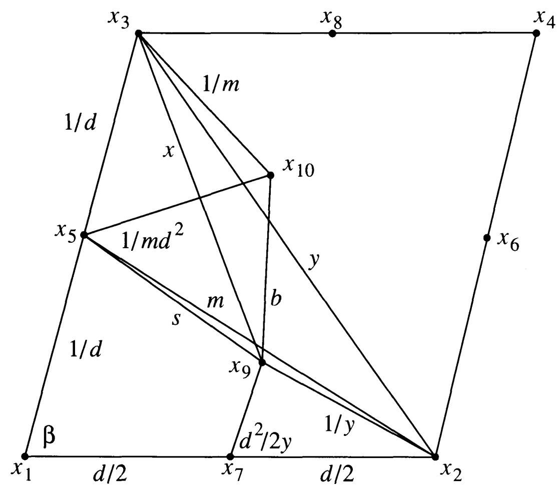


FIGURE 5

Lemma 6.2. *Suppose* $1 < d \leq e$ *and each* $1/d$*-ball touches two full-sized balls separated by a distance* e *and the centers of the three balls are all in a line. Then* $\mathrm{vol}(C) \geq \sqrt{3}/4$.

Proof. Call the two full-sized horoballs H_{X1} and H_{X3}. Then the distance from the center of H_{X1} to the center of H_{X3} is $2/d$. Let H_{X5} be the $1/d$-ball that they share. Let H_{X2} be a full-sized ball with center a distance d from x_1. There is an isometry which switches H_{x1} and H_∞ and sends H_{x5} to H_{x2}. Hence it takes H_{x3} to a ball which touches both H_{x1} and H_{x2}. This new ball, call it H_{x7}, has its center at the midpoint of the line segment from x_1 to x_2. Let β be the angle at x_1 between the line segment from x_1 to x_3 and the line segment from x_1 to x_2. We can choose x_2 so that $\beta \leq \pi/2$. Then $\mathrm{vol}(C) = \frac{1}{4} d(2/d) \sin\beta$. Hence it is enough to show that $\beta \geq \pi/3$ in order to show that $\mathrm{vol}(C) \geq \sqrt{3}/4$.

We label centers of horoballs as in Figure 5. We first note that on ∂H_∞, there are two full-sized horoballs, one a distance y from x_3 and one a distance $2/d$ from x_3. Hence the group element that takes H_∞ to H_{x2} and $H_\infty \cap H_{X3}$ to $H_{X2} \cap H_\infty$ will send the horoball at x_1 to a horoball at x_7 tangent to both H_{X1} and H_{X2}. The horoball at x_2 is sent to a horoball a distance $1/y$ from x_2. Note that the resulting triangle is similar to the triangle determined by x_1, x_2, and x_3 and is obtained by multiplying that triangle by $d/(2y)$. This new horoball at x_9 has a radius $1/(2y^2)$ by Lemma 4.6.

We also have a $1/d$-ball beneath a point w on ∂H_∞ which is at a distance m from x_2. Hence this $(m, 1/d, d)$ triangle must appear on H_{X3}. So there is a smaller horoball with center a distance $1/m$ from x_3 by Lemma 4.6. This new horoball will be tangent to the $1/d$-ball centered at x_5. The triangle defined by x_3, x_5, and the center of this new horoball is similar to the triangle defined by x_2, x_5, and x_1, but its size is $1/(md)$ times the size of the x_2, x_5, x_1 triangle.

Thus, we have horoballs of radius $1/2$ at x_1, x_2, x_3, and x_4. We have horoballs of radius $1/(2d^2)$ at x_5 and x_6. We have horoballs of radius $d^2/8$ at x_7 and x_8. We have a horoball of radius $1/(2y^2)$ at x_9 and a horoball of radius $1/(2m^2d^2)$ at x_{10}.

In order to make sure none of these horoballs overlap, it must at least be true by Lemma 4.4 that $x \geq 1/y$, $s \geq 1/(dy)$, and $b \geq 1/(mdy)$. Note that several other inequalities do exist, but it will turn out that these three will suffice. By the law of sines and the law of cosines we can formulate these inequalities entirely in terms of the variables d and y. The three inequalities yield respectively

$$2y^4 - (d^2 + 4/d^2)y^2 + (d^2 - 4/d^2)^2 \geq 0, \tag{1}$$

$$(dy)^4 - 4(dy)^2 - 4d^4 + d^8 - 2d^2 + 8 \geq 0, \tag{2}$$

$$\begin{aligned} &-32 + 16d^4 - 6d^8 + d^{12} + 40d^2y^2 - 10d^6y^2 \\ &+ 2d^{10}y^2 - 12d^4y^4 - d^8y^4 + 2d^6y^6 \geq 0.2 \end{aligned} \tag{3}$$

The calculation of the above inequalities was aided by the use of Mathematica. Note that the given inequalities are multiples of the original inequalities. Since our goal is to show that $\beta \geq \pi/3$, we take as a fourth inequality $\beta \leq \pi/3$, which yields

$$d^2 + 4/d^2 - y^2 - 2 \geq 0. \tag{4}$$

Any pair of values for y and d which does not satisfy all four of these inequalities cannot yield a counterexample to the theorem. In the case that $1 < d < \sqrt{2}$ and $1 < y$, at least one of the inequalities is not satisfied. At least one of the inequalities is also not satisfied in the case where $d = \sqrt{2}$ except when $y = \sqrt{2}$. In the case $d = y = \sqrt{2}$, we have a cusp volume of $\sqrt{3}/4$.

If $d > \sqrt{2}$, we can interchange d with $2/d$, and the above argument again works, although care must be taken in interchanging them appropriately. Note that since $y \geq d$, $\mathrm{vol}(C) > \sqrt{3}/4$ for all $d > \sqrt{2}$ anyway. q.e.d.

(Some helpful calculations: $s^2 = y^2/2 - 2/d^2 - 2/y^2 + 4/(y^2 d^4) + d^4/(2y^2)$, $m^2 = d^2/2 - 1/d^2 + y^2/2$.)

Lemma 6.3. *Suppose $d, e > 1$, a $1/d$-ball touches two full-sized balls which are separated by a distance d, and the centers of the three balls are all in a line. Then* $\mathrm{vol}(C) \geq \sqrt{3}/4$.

Proof. Then $d = 2/d$, so $d = \sqrt{2}$. But $\mathrm{vol}(C) \geq d^2\sqrt{3}/8$.

Lemma 6.4. *If $d, e > 1$, $\mathrm{vol}(C) < \sqrt{3}/4$, and a $1/d$-ball touches two or more full-sized balls such that if the $1/d$-ball touches exactly two full-sized balls, the three centers of the horoballs are not in a line, then* $\mathrm{vol}(C) = \sqrt{2}/4$.

Proof. This forces the existence of at least four $1/d$-balls per full-sized ball, which requires that $d = e$. Suppose first that the $1/d$-ball touches at least two full-sized balls which are themselves separated by a distance d. By the symmetry in the placement of $1/d$-balls, these two full-sized balls must share two $1/d$-balls. Let H_a and H_b be the two full-sized balls, and H_c and H_d the two $1/d$-balls that they share. There must be an order-two elliptic isometry T which rotates about a geodesic through the tangency point of H_a and H_∞ and switches H_a with H_∞. This rotation will send H_b to a $1/d$-ball tangent to H_a, and H_c and H_d to full-sized balls, both of which are also tangent to $T(H_b)$. Thus, the $1/d$-ball $T(H_b)$ is tangent to at least three full-sized balls.

If the $1/d$-ball touches four full-sized balls, then the symmetry forced by the elliptic rotations through the tangency points forces the four centers of the full-sized balls to form a square of side length d. The $1/d$-ball is at the center of the square, giving a length of $2/d$ to the diagonal. Hence, $d = \sqrt[4]{2}$ and $\mathrm{vol}(C) = \sqrt{2}/4$.

If the $1/d$-ball touches three full-sized balls, the symmetries caused by the elliptic elements forces the centers of the three balls to form an equilateral triangle with edge length d. The center of the $1/d$-ball is at the center of the triangle, and at a distance $1/d$ from each of the vertices. This forces $d = \sqrt[4]{3}$ and $\mathrm{vol}(C) = 3/8$.

This last case is not realized by an orbifold for the following reason. As in Figure 6 there must be vertical elliptic axes of order two at points a and b.

By Lemma 4.9, there are two order-two elliptic axes passing through $H_x \cap H_\infty$ which are both tangent to H_x and H_∞ at $H_x \cap H_\infty$. Rotation about one of these axes must send the $1/d$-ball at g to a full-sized ball centered at a distance d from x. Without loss of generality, we can take

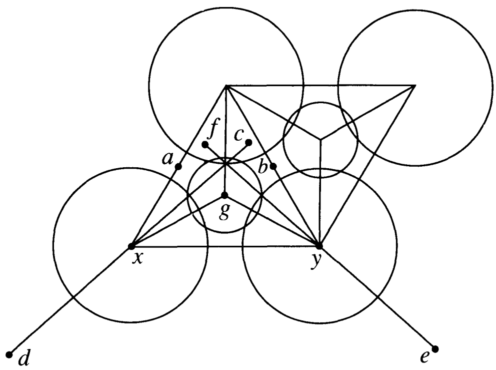


FIGURE 6

this axis to have endpoints at d and c. Then, translation of the other axis by a parabolic isometry will insure an order-two axis with endpoints at f and e.

Rotating about the axis with endpoints at d and c sends the vertical axis at a to an axis, which has one endpoint at x, and is perpendicular to the vertical geodesic at g. Similarly, a rotation about the axis with endpoints f and e sends the vertical geodesic at b to an axis, which has one endpoint at y, and is perpendicular to the vertical geodesic at g.

These two new order-two axes intersect one another at an angle of $2\pi/3$. Hence the vertical geodesic at g must be an elliptic axis of order divisible by 3, contradicting the fact only order-two axes go out the cusp.

Finally, it could also be the case that two full-sized balls separated by a distance other than d share the two $1/d$-balls. Let f be the shortest distance between the centers of two full-sized balls which are not separated by a distance d. Using an argument similar to the one above, we see that there must be a $1/f$-ball which touches three full-sized balls. By symmetry arguments, we can show that this ball must in fact be a $1/d$-ball, forcing $f = d$. This puts us in the previously examined case.

7. One cusp: conclusions

Proof of Theorem 4.1. The fact that a maximal nonrigid cusp of volume less than 0.3969 must have volume either $1/4$, $\sqrt{7}/8$, or $\sqrt{2}/4$ follows from the sequence of lemmas in §§4–6.

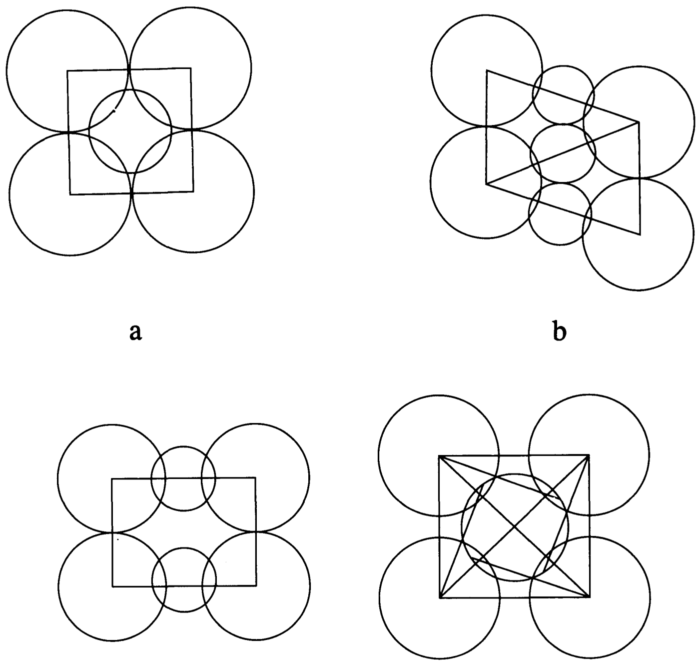
Lemma 7.1. *There exist two unique hyperbolic 3-orbifolds with maximal nonrigid cusp volumes* $1/4$ *and* $\sqrt{7}/8$ *respectively, and exactly two hyperbolic 3-orbifolds, each with a maximal nonrigid cusp volume of* $\sqrt{2}/4$.

Proof. Lemma 3.2 assures us that an orbifold with a maximal nonrigid cusp having one of these three volumes must be a 1-cusped orbifold. We first assume that we have a hyperbolic 3-orbifold with a maximal nonrigid cusp of volume $1/4$. The proofs of Lemmas 5.1 and 5.2 demonstrate that the cusp diagram for such an orbifold appears as in Figure 7a. Let F be the ideal octahedron in H^3 with five vertices at the centers of the five horoballs shown and the sixth vertex at ∞. The elliptic elements in $\pi_1(O)$ will include rotations of order two about each edge in the 1-skeleton together with order-three rotations about axes through opposite faces of the octahedron, and order-two rotations about axes through opposite vertices and about axes bisecting the faces respectively. The isometries generated by these elliptic isometries will tile all of H^3 by the images of this octahedron. A fundamental domain for the resultant group of isometries is given by taking $1/12$ of the original octahedron. The identifications and singular axes on this fundamental domain are completely determined, and hence the resultant orbifold is unique. Since an ideal regular octahedron has volume approximately 3.663862, this orbifold has volume $1/12$ of that, namely $0.30532\ldots$. The fundamental group of this orbifold is in fact $\mathrm{PSL}_2(O_1)$. This orbifold is covered by the Borromean rings complement.

In the case of the cusp volume $\sqrt{7}/8$, the proof of Lemma 5.3 yields a cusp diagram as in Figure 7b. An argument similar to the one above gives a tiling of H^3 by ideal prisms of volume approximately 2.6667. The resulting orbifold is again completely determined. Its volume will be $1/6$ of the volume of the ideal prism, namely $0.4444574\ldots$. The fundamental group of this orbifold is $\mathrm{PGL}_2(O_7)$, and this orbifold is covered by the complement of the link 6_1^3.

In the case of the cusp volume $\sqrt{2}/4$, Lemmas 5.3 and 6.4 yield cusp diagrams as in Figure 7c. In the first diagram, we can tile H^3 by ideal cuboctahedra. The resulting orbifold is completely determined and has volume approximately 0.5019204. The fundamental group of this orbifold is $\mathrm{PGL}_2(O_2)$.

In the second case, let F be the ideal octahedron with vertices at ∞ and at the centers of the five horoballs appearing in the figure. Again, the order-two rotations in the edges of the octahedron together with an order-two rotation in one bisector of each face will tile all of H^3 with copies



a b

c

FIGURE 7

of the octahedron. Cut the original octahedron into four $(\pi/4, \pi/4, \pi/2)$ ideal tetrahedra by cutting along the four ideal triangles which share the vertical edge through the center of the central horoball.

Corresponding to the elliptic axes in Lemma 4.9, there will be one elliptic axis of order two passing through two opposite edges of each tetrahedron. Without loss of generality, we can take these axes to be the ones which appear in the figure. Although rotation about these geodesics does not preserve the octahedral tiling, it does preserve the $(\pi/4, \pi/4, \pi/2)$ ideal tetrahedral tiling and completely determines the orbifold. The resulting orbifold is covered by the link 8^4_2, and has volume approximately 0.4579827.

Proof of Corollary 4.2. We have shown that if O has a maximal nonrigid cusp of volume less than 0.3969, then the cusp must have volume $1/4$, $\sqrt{7}/8$, or $\sqrt{2}/4$ and O must be the corresponding orbifold of volume approximately 0.3053 or 0.4444 in the first two cases and 0.5019 or 0.4579 in the last case. If O has a maximal nonrigid cusp with volume at least 0.3969, Lemma 3.4 implies $\mathrm{vol}(O) \geq 0.4651$. Since Theorem 2.1

states that every limit volume corresponds to an orbifold with a nonrigid cusp, the result follows.

8. Applications

Let M be a noncompact finite volume orientable hyperbolic 3-manifold, and let $\mathrm{Isom}^+(M)$ be the group of orientation-preserving isometries of M.

Corollary 8.1. *If all the elements of* $\mathrm{Isom}^+(M)$, *which fix points in the interior of any embedded cusp in* M *are of order two, then* $|\mathrm{Isom}^+(M)| \leq 3.276\,\mathrm{vol}(M)$.

Proof. Let O be the orbifold obtained by taking the quotient of M by $\mathrm{Isom}^+(M)$. Then $\mathrm{vol}(O) \geq 0.3053$ by Corollary 4.2. But $\mathrm{vol}(O) = \mathrm{vol}(M)/|\mathrm{Isom}^+(M)|$ which immediately yields the result. q.e.d.

Note that when we utilize the actual values for the numbers rather than the decimal approximations, this lower bound is exactly realized by the orientation-preserving isometry group of the Borromean rings complement.

Corollary 8.2. *Let* K *be a knot in* S^3 *with hyperbolic complement. Then the order of the group of periodic orientation-preserving diffeomorphisms from* S^3 *to* S^3 *leaving* K *invariant is bounded above by* $3.276\,\mathrm{vol}(S^3 - K)$.

Proof. Such a diffeomorphism restricts to a homeomorphism of the complement that must send the longitude I back to $\pm I$ since the longitude is trivial in $H_1(S^3 - K)$. This homeomorphism corresponds to an isometry of the hyperbolic complement which must also send the longitude I back to $\pm I$. Hence, there cannot be any elliptic axes of order greater than two going out the cusp. The result then follows from Corollary 8.1. q.e.d.

This last result generalizes.

Corollary 8.3. *Let* M *be a noncompact finite volume orientable hyperbolic 3-manifold with one cusp. Then* $|\mathrm{Isom}^+(M)| \leq 3.276\,\mathrm{vol}(M)$.

Proof. By Lemma 6.7 of [5], $H_1(M)$ is infinite. Let M' be the manifold obtained by removing the interior of an embedded cusp from M. By the proof of Lemma 6.8 of [5], $i^*: H_1(\partial M') \to H_1(M')$ has infinite image and nontrivial kernel. Hence, by trivializing any torsion, there is an epimorphism $\beta: H_1(\partial M') \to Z$ with nontrivial kernel. This kernel is generated by a single nontrivial simple closed curve in $\partial M'$ with either one of its two possible orientations. We call such a curve I a longitude of $\partial M'$. Let θ be an isometry of M, and θ' its restriction to M'. Then θ'^* must preserve the kernel of β, and hence θ' must send I to a curve homotopic to I or $-I$. Therefore, if θ is an elliptic isometry with axis going out the cusp, it must have order two. The result then follows immediately from Corollary 8.1.

Acknowledgment

I would like to thank the University of California–Santa Barbara, where much of this work was done.

References

[1] C. Adams, *The noncompact hyperbolic 3-manifold of minimal volume*, Proc. Amer. Math. Soc. **100** (1987) 601–606.

[2] ——, *Volumes of N-cusped hyperbolic 3-manifolds*, J. London Math. Soc. (2) **38** (1988) 555–565.

[3] C. Adams, M. Hildebrand & J. Weeks, *Hyperbolic invariants of knots and links*, Trans. Amer. Math. Soc. (to appear).

[4] C. Adams & W. Sherman, *Minimum ideal triangulations of hyperbolic 3-manifolds*, J. Discrete Computational Geometry (to appear).

[5] J. Hempel, *3-manifolds*, Annals of Math. Studies, No. 86, Princeton Univ. Press, Princeton, NJ, 1976.

[6] R. Meyerhoff, *The cusped hyperbolic 3-orbifold of minimum volume*, Bull. Amer. Math. Soc. (N.S.) **13** (1985) 154–156.

[7] ——, *Sphere packing and volume in hyperbolic 3-space*, Comment. Math. Helv. **61** (1986) 271–278.

[8] W. Thurston, *The geometry and topology of 3-manifolds*, Lecture notes, Princeton University, 1978.

WILLIAMS COLLEGE