Annals of Mathematics **176** (2012), 261–301
http://dx.doi.org/10.4007/annals.2012.176.1.4

Minimal co-volume hyperbolic lattices, II: Simple torsion in a Kleinian group

By T. H. Marshall and G. J. Martin

Abstract

This paper represents the final step in solving the problem, posed by Siegel in 1945, of determining the minimal co-volume lattices of hyperbolic 3-space $\mathbb{H}$ (also Problem 3.60 (F) in the Kirby problem list from 1993). Here we identify the two smallest co-volume lattices. Both these groups are two-generator arithmetic lattices, generated by two elements of finite orders 2 and 3. Their co-volumes are $0.0390\ldots$ and $0.0408\ldots$; the precise values are given in terms of the Dedekind zeta function of a number field via a formula of Borel.

Our earlier work dealt with the cases when there is a finite spherical subgroup or high order torsion in the lattice. Thus, here we are concerned with the study of simple torsion of low order and the geometric structure of Klein 4-subgroups of a Kleinian group. We also identify certain universal geometric constraints imposed by discreteness on Kleinian groups which are of independent interest.

To obtain these results we use a range of geometric and arithmetic criteria to obtain information on the structure of the singular set of the associated orbifold and then co-volume bounds by studying equivariant neighbourhoods of fixed point sets, together with a rigorous computer search of certain parameter spaces for two-generator Kleinian groups.

1. Introduction

A Kleinian group Γ is a discrete nonelementary subgroup of the orientation preserving isometry group of hyperbolic 3-space, $\mathrm{Isom}^+(\mathbb{H}^3)$. In this setting nonelementary means not virtually abelian. The orbit spaces of Kleinian groups are hyperbolic 3-*orbifolds*,

$$Q = \mathbb{H}^3/\Gamma.$$

Q is a hyperbolic manifold if Γ is torsion-free — that is if Γ has no elements of finite order. The Kleinian group Γ is called a *lattice* if the hyperbolic volume $\mathrm{vol}(Q) < \infty$.

Research supported in part by grants from the U.S. National Science Foundation, the Miller Foundation (Berkeley), the N.Z. Marsden Fund and the New Zealand Royal Society (James Cook Fellowship).

In 1945 Siegel [30] posed the question of identifying the numbers μ_n defined to be the infimum of co-volumes among all lattices acting on hyperbolic n-space. He solved this problem in dimension 2 ($\mu_2 = \pi/21$ from the (2,3,7)-triangle group) and suggested (since at that time the theory of covering spaces was not well developed) a connection to Hurwitz' $84g - 84$ Theorem of 1892 [19]. This was later established by Macbeath [22]. Mostow's rigidity theorem [26] and Selberg's Lemma [29] establish the connection between minimal co-volume lattices of hyperbolic n-space, $n \geq 3$, and maximal automorphism groups of hyperbolic manifolds, more generally. This is discussed in the case of three dimensions in [5], the sharp bounds being a consequence of the results established here.

The series of papers [12], [15], [13], [10] identified many universal geometric criteria satisfied by Kleinian groups in 3-dimensions and discussed connections with arithmeticity that underpin many of the results here.

Our main theorem, Theorem 1.1 below, identifies the two lattices of $\mathbb{H}^3$ of smallest co-volume. Surprisingly, both of these are generated by two elements of finite orders 2 and 3.

THEOREM 1.1. *Let Γ be a Kleinian group. Then either*

$$\mathrm{vol}(\mathbb{H}^3/\Gamma) = \mathrm{vol}(\mathbb{H}^3/\Gamma_0) = 275^{3/2}2^{-7}\pi^{-6}\zeta_k(2) = 0.0390\ldots \quad \text{and} \quad \Gamma = \Gamma_0, \text{ or}$$
$$\mathrm{vol}(\mathbb{H}^3/\Gamma) = \mathrm{vol}(\mathbb{H}^3/\Gamma_1) = 283^{3/2}2^{-7}\pi^{-6}\zeta_{k'}(2) = 0.0408\ldots \quad \text{and} \quad \Gamma = \Gamma_1, \text{ or}$$
$$\mathrm{vol}(\mathbb{H}^3/\Gamma) > V_0 := 0.041.$$

(The equality here is up to conjugacy.) Here is a description of the two groups Γ_0 and Γ_1 and the associated arithmetic data:

- Γ_0 is a two-generator arithmetic Kleinian group obtained as a $\mathbb{Z}_2$-extension of the index 2 orientation preserving subgroup of the group generated by reflection in the faces of the 3-5-3-hyperbolic Coxeter tetrahedron, and ζ_k is the Dedekind zeta function of the underlying number field $\mathbb{Q}(\gamma_0)$, with γ_0 a complex root of $\gamma^4+6\gamma^3+12\gamma^2+9\gamma+1=0$, of discriminant -275. The associated quaternion algebra is unramified. This group has a discrete and faithful representation in $\mathrm{SL}(2,\mathbb{C})$, determined uniquely up to conjugacy, generated by two matrices A and B with $\mathrm{tr}^2(A) = 0$, $\mathrm{tr}^2(B) = 1$ and $\mathrm{tr}(ABA^{-1}B^{-1}) - 2 = \gamma_0$.
- Γ_1 is a two-generator arithmetic Kleinian group, and $\zeta_{k'}$ is the Dedekind zeta function of the underlying number field $\mathbb{Q}(\gamma_1)$, with γ_1 a complex root of $\gamma^4 + 5\gamma^3 + 7\gamma^2 + 3\gamma + 1 = 0$, of discriminant -283. The associated quaternion algebra is unramified. This group has a discrete and faithful representation in $\mathrm{SL}(2,\mathbb{C})$, determined uniquely up to conjugacy, generated by two matrices A and B with $\mathrm{tr}^2(A) = 0$, $\mathrm{tr}^2(B) = 1$ and $\mathrm{tr}(ABA^{-1}B^{-1}) - 2 = \gamma_1$.

In particular we solve Siegel's problem in dimension 3.

COROLLARY 1.1. $\mu_3 = 275^{3/2}2^{-7}\pi^{-6}\zeta_k(2) = 0.03905\ldots$.

The volume formulas here are found in Borel's important paper [2] giving co-volume bounds for maximal arithmetic hyperbolic lattices in three dimensions.

Notice that in both instances above A represents an element of order 2 and B an element of order 3. Indeed it would also appear quite likely that the next two smallest co-volume lattices contain groups generated by two elements of finite orders 2 and 3 with low index. The co-volumes of these lattices would be $31^{3/2}2^{-6}\pi^{-4}(N\mathcal{P}_3 - 1)\zeta_k(2) = 0.0659\ldots$, ζ_k is the Dedekind zeta function of the number field $\mathbb{Q}(\gamma_3)$, γ_3 a complex root of $\gamma^3 + 4\gamma^2 + 5\gamma + 3 = 0$ of discriminant -31 and the quaternion algebra ramified at the finite place $\mathcal{P}_3$. This group contains orbifold (3,0)-(3,0) Dehn surgery on the Whitehead link as a subgroup of index 8 and so a group generated by elements of order 2 and 3 of index 4. Next, $44^{3/2}2^{-6}\pi^{-4}(N\mathcal{P}_2 - 1)\zeta_k(2) = 0.0661\ldots$, ζ_k is the Dedekind zeta function of the number field $\mathbb{Q}(\gamma_4)$, γ_4 a complex root of $\gamma^3 + 4\gamma^2 + 4\gamma + 2 = 0$ of discriminant -44, and the quaternion algebra ramified at the finite place $\mathcal{P}_2$ contains a group generated by elements of order 2 and 3 also of index 4.

1.1. *Basic notation and strategy of proof.* Before we give the basic strategy of the proof of Theorem 1.1 we need to set up some notation and give a few definitions. We denote by $\rho(A, B)$ the (hyperbolic) distance between A and B, where each of A and B is either a point or a subset of $\mathbb{H}^3$. Typically A and B are both geodesics, and in this case, we also have a notion of *complex* distance.

Suppose ℓ_1 and ℓ_2 are two hyperparallel oriented geodesics in $\mathbb{H}^3$. We let ℓ be their common perpendicular, p_i the point of intersection of ℓ_i with ℓ and Π_i the halfplane with boundary geodesic ℓ that contains the ray along ℓ_i emanating from p_i in the direction given by the orientation of ℓ_i. These two halfplanes meet along ℓ, and we define the angle between ℓ_1 and ℓ_2 to be the angle from Π_1 to Π_2 measured anticlockwise, as determined by the orientation of ℓ from p_1 to p_2 and the right-hand rule. This angle is well defined (modulo 2π) and independent of the order in which the geodesics are given. If the geodesics are unoriented, then the angle between them is still defined modulo π. Observe that if ℓ_1 and ℓ_2 cross, then the angle between them is only defined modulo supplementation. We define the angle between geodesics that meet only on the boundary to be zero. The *complex distance* $\Delta(\ell_1, \ell_2)$ between ℓ_1 and ℓ_2 is now defined to be $\delta + i\phi$, where δ and ϕ are respectively the distance and the angle between ℓ_1 and ℓ_2. We are mostly concerned with the *real* distance $\rho(\ell_1, \ell_2)$ between geodesics.

Every loxodromic or elliptic element f in a Kleinian group Γ fixes two points of the Riemann sphere $\hat{\mathbb{C}} = \partial\mathbb{H}^3$, and the closed hyperbolic line in $\mathbb{H}^3$ joining these two points is called the *axis* of f, denoted $\mathrm{ax}(f)$. Such an f translates along this axis by an amount $\tau_f \geq 0$, the *translation length*, and rotates by an angle $\eta_f \in (-\pi, \pi]$, the *holonomy* or *rotation angle*, about this axis. If f is elliptic ($\tau_f = 0$), then it is a rotation of some finite period (the order) about its axis, the axis itself being exactly the fixed point set of f. In this case the holonomy is just the rotation angle and is only defined up to sign. If f is loxodromic ($\tau_f > 0$), then we can define η_f unambiguously (modulo 2π) by orienting $\mathrm{ax}(f)$ in the direction of translation and letting η_f be measured anticlockwise as determined by this orientation and the right-hand rule.

An elliptic f is called *simple* if for all $h \in \Gamma$,

$$h(\mathrm{ax}(f)) \cap \mathrm{ax}(f) = \emptyset \quad \text{or} \quad h(\mathrm{ax}(f)) = \mathrm{ax}(f). \tag{1}$$

Thus, for simple f, the translates of the axis of f will form a disjoint collection of hyperbolic lines. More generally, for any set $X \subset \mathbb{H}^3$, if for all $h \in \Gamma$,

$$h(X) \cap X = \emptyset \quad \text{or} \quad h(X) = X, \tag{2}$$

then X is called *precisely invariant* — a term introduced by Maskit. A *collar* of radius r about a subset $E \subset \mathbb{H}^3$ is

$$\mathcal{C}(E, r) = \{x \in \mathbb{H}^3 : \rho(x, E) \leq r\}.$$

The *collaring radius* of a nonparabolic $f \in \Gamma$ is the supremum of those numbers r for which $\mathrm{ax}(f)$ has a precisely invariant collar of radius r. Such an r always exists for the loxodromic of shortest translation in a Kleinian group since (roughly) it will project to the shortest geodesic and this geodesic will be embedded. Further, every lattice is geometrically finite and so, in particular, the spectrum of traces of elements in the groups will be discrete and there will be a loxodromic with shortest translation length (which, for brevity, we will refer to as a *shortest loxodromic*).

If g is an elliptic element that is not simple, then g lies in a triangle subgroup that is either spherical (finite) or euclidean should $\mathbb{H}^3/\Gamma$ not be compact. Accordingly any elliptic element of order $n \geq 7$ is simple, but it may lie in a dihedral subgroup.

In [12] we show that given elliptic elements f and g of order p and q generating a Kleinian group, the allowable (hyperbolic) distances $\delta_i(p, q)$ between their axes has an initially discrete spectrum and, crucially, the first several initial values of the spectrum (at least for $p, q \leq 6$) are uniquely attained for arithmetic lattices [10]. In [17] a "collar-volume" formula is established which gives co-volume estimates simply in terms of the collaring radius of an elliptic axis. This radius is bounded below by half the minimum possible distance between axes of order p. This minimum is $\delta_1(p, p)$. For $p \geq 7$, the value $\delta_1(p, p) =$

$1/2\sin(\pi/p)$ gives enough volume to exceed the co-volume of $\mathrm{PGL}(2,\mathcal{O}_{\sqrt{-3}})$. This group has been identified as the minimal co-volume noncompact lattice by Meyerhoff [25] and is also known to be the minimal co-volume lattice containing torsion of order ≥ 6, [17]. Combining these results gives

THEOREM 1.2. *If Γ is a Kleinian group containing either a parabolic element or an elliptic element of order $p \geq 6$, then*

$$\mathrm{vol}(\mathbb{H}^3/\Gamma) \geq \mathrm{vol}(\mathbb{H}^3/PGL(2,\mathcal{O}_{\sqrt{-3}})) = 0.0846\ldots\ .$$

This result is sharp.

These arguments need to be slightly refined to deal with simple torsion of order 4 and 5, and this is also done in [17] with the lower bound 0.041 given. (This is not sharp and probably far from what is true.)

If our Kleinian group is torsion-free, then the collar-volume formula together with the $(\log 3)/2$ theorem of Gabai, Meyerhoff and Thurston [9] easily give considerably larger volume estimates; see [17]. Gabai, Meyerhoff and Milley [8] have recently proved the sharp result: the minimum covolume in the torsion-free case is attained by the Weeks manifold of volume $0.94\ldots$.

If there is a nonsimple elliptic, or equivalently when there is a finite spherical subgroup, the arguments are of a different nature but still based on knowledge of the spectra of possible axial distances for elliptics of orders $3, 4$ and 5. A *spherical point* is a point stabilized by a spherical triangle subgroup of a Kleinian group — namely, the tetrahedral, octahedral and icosahedral groups A_4, S_4 and A_5. Geometric position arguments based around the axes emanating from a spherical point show the distances between spherical points to be uniformly bounded below. In [18] we identify the initial part of this spectra of distances, significantly extending earlier work of Derevnin and Mednykh, [6]. Again, a crucial point is establishing arithmeticity of the first few extrema. This enables us to use arithmetic criteria to eliminate small configurations from consideration. Once this is done, a ball of maximal radius about a suitable spherical point will be precisely invariant and provide volume bounds. There are of course additional complications, but in the end we obtain the main result of [18], which is our Theorem 1.1 in the case that Γ has a tetrahedral, octahedral or icosahedral subgroup.

The above discussion shows us that there are two remaining cases to deal with in order to identify the two smallest co-volume Kleinian groups Γ: the case where Γ contains a simple elliptic of order 3, and the case where all elliptics in Γ are order 2. Most of the work of proving Theorem 1.1 comes from the latter case. Here, two elements of order 2 can only meet at right angles and then they generate a Klein 4-group. Already in [17] it was shown that these Klein 4-groups appear in many extremal situations.

One of the first results we establish is the following universal constraint concerning discrete groups generated by two loxodromics whose axes meet orthogonally. Groups with orthogonal loxodromics appear naturally as they "wind up" Klein 4-groups when projected to the quotient orbifold. In the case of two dimensions, Beardon gives an account of this and related universal constraints; see [1, Ch. 11].

THEOREM 1.3. *Let f and g be loxodromic transformations generating a discrete group such that the axes of f and g meet at right angles. Let τ_f and τ_g be the respective translation lengths of f and g. Then*

$$\max\{\tau_f, \tau_g\} \geq \lambda_\perp, \tag{3}$$

where

$$\lambda_\perp = \operatorname{arccosh}\left(\frac{\sqrt{3}+1}{2}\right) = 0.831446\ldots. \tag{4}$$

Equality holds when $\langle f, g\rangle$ is the two-generator arithmetic torsion-free lattice with presentation

$$\Gamma = \langle f, g : fg^{-1}fgfgf^{-1}gfg = gfg^2fgf^{-1}g^2f^{-1} = 1\rangle.$$

This group is a four-fold cover of (4,0) *and* (2,0) *Dehn surgery on the* 2 *bridge link complement* 6^2_2 *of Rolfsen's tables,* [28]. $\mathbb{H}^3/\Gamma_0$ *has volume* 1.01494160 . . . , *Chern-Simons invariant* 0 *and homology* $\mathbb{Z}_3 + \mathbb{Z}_6$.

This result significantly refines a particular case of an earlier theorem [14] concerning groups generated by loxodromics with intersecting axes. We expect that Theorem 1.3 represents the extreme case independently of the angle at which the axes of loxodromics meet.

1.2. *Outline of the Proof for Theorem* 1.1. As noted above, we are reduced to proving this theorem in the cases where Γ has either a simple elliptic of order 3 or no elliptics of order greater than 2. Our first step is to prove Theorem 1.3 above (Sections 3–6). Next, we establish some collar-volume estimates (Sections 8 and 9), the main result being

LEMMA 1.1. *If g is a loxodromic in a Kleinian group Γ that has a common axis with an elliptic of order k and a collaring radius of at least c_k, where $c_1 = 0.345$, $c_3 = 0.294$ and $c_k = 0.4075$ for $k \neq 1, 3$, then* $\operatorname{vol}(\mathbb{H}^3/\Gamma) > V_0$.

Here and subsequently we interpret the statement that g has a common axis with an elliptic of order 1 to mean that g has a common axis with no elliptic.

From the elliptic collaring theorems and the identification of all the small extremals as arithmetic in [10], any loxodromic in a Kleinian group that shares its axis with an elliptic of order 3 has collaring radius at least 0.294, or else

Γ is arithmetic and is either one of the groups Γ_0, Γ_1 of Theorem 1.1, or has co-volume greater than V_0. Lemma 1.1 thus gives Theorem 1.1 in the case where Γ has 3-torsion.

In the remaining case — where there is only 2-torsion — we prove (Section 10) the following upper bound on the shortest translation length in the extremal case; Theorem 1.3 plays an important part in establishing this.

LEMMA 1.2. *A Kleinian group* Γ *with no elliptics of order* 3 *or more and co-volume at most* V_0 *has a shortest loxodromic with translation length* $\tau \leq 0.497$. *This loxodromic does not share its axis with any elliptic element of* Γ.

After a preliminary collaring estimate (Section 11), we complete the proof of Theorem 1.1 by combining the above with the following result (Section 12).

LEMMA 1.3. *There is no Kleinian group* Γ *with co-volume less than* V_0, *no elliptics of order* 3 *or more, and shortest loxodromic with translation length* $\tau \leq 0.497$ *and collaring radius* $r \leq 0.345$.

Remark on computer assistance. The proof of our results will be computer assisted in a number of places, though this assistance occurs only in a mild way. We rigorously search certain regions to show there are no discrete groups within, somewhat akin to the search of [9]. Here however, we shall use new polynomial trace identities to provide inequalities analagous to Jørgensen's well-known result [20]. Most of our searches are broken down into lesser searches that take anything from a few minutes to 60 hours on a Macintosh G5, using (very simple) code written in Mathematica. Most of the computation is in the proof of Lemma 1.3.

All our computations take the same general form: we are given a domain $U \subseteq \mathbb{C}^2$ and a collection $\mathcal{Q} = \{q_i \mid i \in I\}$ of real valued functions defined on $\mathbb{C}^2$, and we must show that, at each point of U, at least one of the functions is positive. In each of the proofs involving a machine computation (Theorem 1.3, Lemmas 1.1 and 1.2) we will simply find the domain U and the family $\mathcal{Q}$. We relegate details of the computation (mostly finding Lipschitz bounds) to Appendix 1.

We point out that it is entirely possible that minor modifications to the Gabai-Meyerhoff-Thurston [9] search may reproduce most of the results here. This modification would be to the way certain regions are eliminated. In particular, one would need to identify those subregions where a contradiction to a choice of shortest loxodromic was achieved by identifying an elliptic of order 2 or 3. Presumably this is a very small part of the parameter space searched. Further, the collar-volume estimate shows one does not have to search all the way out to $\frac{1}{2}\log 3$. This would save a considerable amount of computation. However we did not take this approach. Firstly, despite some attempts we were not able to rewrite and run the code successfully. Second, our aim here was to produce results that could be verified on a reasonable machine in a reasonable amount of time. Thus we used geometry to reduce the size of search spaces as much as possible. Another of our aims was the hope that the simplified search procedure presented here using polynomial trace identities and generalisations of Jørgensen's inequality will be implemented to eventually give an alternative and simpler proof of the important $\frac{1}{2}\log 3$ theorem of [9].

2. Discrete groups and polynomial trace identities

For each $f \in \mathrm{Isom}^+(\mathbb{H}^3)$, we define the *trace* of f, which we denote by $\mathrm{tr}(f)$, by choosing a matrix representative for f in $\mathrm{PSL}(2,\mathbb{C})$,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{for} \quad f|\mathbb{C} = \frac{az+b}{cz+d},$$

and letting $\mathrm{tr}(f) = \mathrm{tr}(A)$; thus $\mathrm{tr}(f)$ is defined up to sign. We then define, for $f, g \in \mathrm{Isom}^+(\mathbb{H}^3)$,

$$\beta(f) = \mathrm{tr}^2(f) - 4 \quad \text{and} \quad \gamma(f,g) = \mathrm{tr}[f,g] - 2.$$

($\mathrm{tr}[f,g]$ is defined unambiguously by representing $[f,g]$ by a matrix of the form $[A,B]$.)

The parameters defined above conveniently encode various other geometric quantities. For instance, $\gamma(f,g) = 0$ if and only if f and g share a common fixed point in $\hat{\mathbb{C}}$, an elementary fact proved in [1], which we will often use in what follows. Other examples are the following. If f and g are each elliptic or loxodromic, with translation lengths τ_f and τ_g respectively and holonomies η_f and η_g respectively, then we have [16]

$$\beta(f) = 4\sinh^2\left(\frac{\tau_f + i\eta_f}{2}\right) = 2\left(\cosh(\tau_f + i\eta_f) - 1\right), \tag{5}$$

$$\beta(g) = 4\sinh^2\left(\frac{\tau_g + i\eta_g}{2}\right) = 2\left(\cosh(\tau_g + i\eta_g) - 1\right), \tag{6}$$

$$\gamma(f,g) = \frac{\beta(f)\beta(g)}{4}\sinh^2[\Delta(\mathrm{ax}(f),\mathrm{ax}(g))]. \tag{7}$$

Recall from Section 1.1 that the angle ϕ in the complex distance $\delta + i\phi := \Delta(\mathrm{ax}(f),\mathrm{ax}(g))$ is defined modulo π, so that the right-hand side of (7) is well defined.

From (5)–(7), we derive

$$\cosh(\tau_f) = \frac{|\beta(f)+4| + |\beta(f)|}{4}, \tag{8}$$

$$\cos(\eta_f) = \frac{|\beta(f)+4| - |\beta(f)|}{4}, \tag{9}$$

$$\cosh(2\delta) = \left|\frac{4\gamma(f,g)}{\beta(f)\beta(g)} + 1\right| + \left|\frac{4\gamma(f,g)}{\beta(f)\beta(g)}\right|, \tag{10}$$

$$\cos(2\phi) = \left|\frac{4\gamma(f,g)}{\beta(f)\beta(g)} + 1\right| - \left|\frac{4\gamma(f,g)}{\beta(f)\beta(g)}\right|. \tag{11}$$

We are often concerned with the case where one of the isometries, say g, is of order 2, in which case $\beta(g) = -4$, and (10) and (11) take the simpler form

$$\cosh(2\delta) = |1 - \gamma(f,g)/\beta(f)| + |\gamma(f,g)/\beta(f)|, \tag{12}$$

$$\cos(2\phi) = |1 - \gamma(f,g)/\beta(f)| - |\gamma(f,g)/\beta(f)|. \tag{13}$$

We view the space of all two-generator Kleinian groups modulo conjugacy as a subset of the three complex dimensional space $\mathbb{C}^3$ via the map

$$\langle f, g\rangle \to (\gamma, \beta, \beta'), \tag{14}$$

where here $\beta = \beta(f)$, $\beta' = \beta(g)$ and $\gamma = \gamma(f,g)$.

A fundamental problem is to find the points in $\mathbb{C}^3$ that correspond to discrete two-generator groups. This space has very complicated structure, but those points corresponding to two-generator lattices will be isolated. In order to identify these groups we will have to develop tests — such as Jørgensen's inequality — that identify nondiscrete groups. Our computer searches analyze particularly important parts of $\mathbb{C}^3$; specifically, certain slices of co-dimension one or two. These slices arise as there is a projection from the three complex dimensional space of discrete groups to the two complex dimensional slice $\beta' = -4$ that preserves discreteness and which comes about roughly because we can assume one generator has order 2. This is only a property of two-generator groups [3].

LEMMA 2.1. *Suppose (γ, β, β') are the parameters of a discrete group. Then so are $(\gamma, \beta, -4)$:*

$$(\gamma, \beta, \beta') \text{ discrete} \Rightarrow (\gamma, \beta, -4) \text{ discrete.} \tag{15}$$

Further, away from a small finite set of exceptional parameters this projection also preserves the property of being nonelementary.

This exceptional set will not concern us greatly here as it occurs only in the presence of finite spherical triangle subgroups (other than the Klein 4-group). The proof of Lemma 2.1 consists in examining the parameters for the $\mathbb{Z}_2$ extensions of the discrete subgroup $\langle f, gfg^{-1}\rangle$, generated by two elements with the same trace; see [14].

We now discuss a very interesting family of polynomial trace identities that will be used to obtain geometric information about Kleinian groups [14].

Let $\langle a, b\rangle$ be the free group on the two letters a and b. We say that a word $w \in \langle a, b\rangle$ is a *good word* if w can be written as

$$w = b^{s_1} a^{r_1} b^{s_2} a^{r_2} \cdots b^{s_{m-1}} a^{r_{m-1}} b^{s_m}, \tag{16}$$

where $s_1 \in \{\pm 1\}$, $s_j = (-1)^{j+1} s_1$ and $r_j \neq 0$ but are otherwise unconstrained. The good words start with b and end in $b^{\pm 1}$ depending on whether m is even or odd; the exponents of b alternate in sign. The following theorem is a key tool used in the study of the parameter spaces of discrete groups.

THEOREM 2.1 ([14, Th. 7.13]). *Let $w = w(a, b) \in \langle a, b\rangle$ be a good word, $\beta = \beta(f)$ and $\gamma = \gamma(f, g)$. Then there is a monic polynomial p_w of two complex variables, having integer coefficients such that*

$$\gamma(f, w(f, g)) = p_w(\gamma, \beta) \tag{17}$$

with $p_w(0, \beta) = 0$.

There are three things to note. The first is that if we assume that $b^2 = 1$, then the alternating sign condition is redundant and every $w(a, b)$, where w takes the form (16), is good. The second thing is that there is a natural semigroup operation on the good words. If $w_1 = w_1(a, b)$ and $w_2 = w_2(a, b)$ are good words, then so is

$$w_1 * w_2 = w_1(a, w_2(a, b)). \tag{18}$$

That is, we replace every instance of b in w_1 with $w_2(a, b)$. So, for example,

$$(bab^{-1}ab) * (bab^{-1}) = bab^{-1}a(bab^{-1})^{-1}abab^{-1} = bab^{-1}aba^{-1}b^{-1}abab^{-1}.$$

It is not too difficult to see that

$$p_{w_1 * w_2}(\gamma, \beta) = p_{w_1}(p_{w_2}(\gamma, \beta), \beta), \tag{19}$$

which corresponds to polynomial composition in the first slot. For a two generator Kleinian group $\langle f, g\rangle$, we make the assignment

$$f \mapsto a, \qquad g \mapsto b$$

and correspondingly call $w(f, g)$ a good word. Notice the obvious fact that $\langle f, g\rangle$ Kleinian implies $\langle f, w(f, g)\rangle$ discrete. Finally note that for any word $w = w(f, g)$ and $m, n \in \mathbb{Z}$, $\gamma(f, f^m w f^n) = \gamma(f, w)$ so that the requirement that the word start and end in a nontrivial power of b is simply to avoid some obvious redundancy.

Let us give two simple examples of word polynomials and how they generate inequalities. This might seem an aside to our task of studying small co-volume lattices, but our computer searches amount in large part to mechanising the following arguments. The direct calculation of the polynomial p_w from w by hand can be a little tricky but can be done for some short words w; see [14].

First is the classical example:

$$w = bab^{-1}, \qquad p_w(z) = z(z-\beta). \tag{20}$$

Here $z = \gamma(f,g)$ and $\beta = \beta(f)$. We suppress β and treat z as a variable. The next two examples come from the good words

$$w = bab^{-1}ab, \qquad p_w(z) = z(1+\beta-z)^2,$$
$$w = bab^{-1}a^{-1}b, \qquad p_w(z) = z(1-2\beta+2z-\beta z+z^2).$$

Let us indicate how these words are used to describe parts of the parameter space for two-generator Kleinian groups. We take for granted the well-known fact that the space of discrete nonelementary groups is closed. Jørgensen gave a proof of this as a consequence of his inequality, and it is a very general fact concerning groups of isometries of negative curvature. Let us use the first word $w = bab^{-1}$ and its associated polynomial to recover Jørgensen's trace inequality from the fact that the space is closed. Consider

$$J = \min\{|\gamma| + |\beta| : (\gamma,\beta,\beta') \text{ are the parameters of a Kleinian group}\}.$$

This minimum is attained by some Kleinian group $\Gamma = \langle f, g\rangle$. If $\Gamma' = \langle f, gfg^{-1}\rangle$ is Kleinian (it is certainly discrete), then, by minimality,

$$|\gamma| + |\beta| \le |\gamma(\gamma-\beta)| + |\beta|, \tag{21}$$

$$1 \le |\gamma-\beta|, \tag{22}$$

as $\gamma \neq 0$. (As noted above, $\gamma = 0$ implies that f and g share a fixed point on $\hat{\mathbb{C}}$, so Γ could not be Kleinian.) If f has order $2,3,4$ or 6, $|\beta| \ge 1$, so $J \ge 1$, and there is nothing more to prove. In all other cases, it follows from the classification theorem [1] of elementary discrete groups that Γ' is elementary implies f and gfg^{-1} have the same fixed point set. Thus, g fixes or permutes the fixed point set of f, implying that Γ is elementary, a contradiction.

We have shown that at the minimum we must have either $|\beta| \ge 1$ or $|\gamma-\beta| \ge 1$ and so $J = |\gamma| + |\beta| \ge 1$. This is Jørgensen's inequality. This inequality is attained with equality (the argument shows that in fact (21) holds with equality) for representations of the $(2,3,p)$-triangle groups.

As far as our search for Kleinian group parameters in $\mathbb{C}^3$ is concerned, Jørgensen's inequality tell us that the region

$$\{(\gamma,\beta,\beta') : |\gamma| + |\beta| < 1 \;\text{ or }\; |\gamma| + |\beta'| < 1\}$$

contains no parameters for Kleinian groups. We will want to extend this region significantly, to identify the places where Kleinian groups actually are. Following the argument that produces Jørgensen's inequality, we will consider other polynomial trace identities to get further inequalities. A point to observe here is that in more general situations — that is for other polynomials — we must examine and eliminate, for some geometric reason, the zero locus of p_w. (In the above example the variety $\{\gamma = \beta\}$ was eliminated through the classification of the elementary groups; we will consistently ignore the locus $\{\gamma = 0\}$ when the groups in question are Kleinian.) When $p_w \neq 0$, basically $\Gamma' = \langle f, w(f,g)\rangle$ is Kleinian and so a candidate for a minimisation procedure.

So for instance, if we minimize $|\gamma|+|1+\beta|$ and use the second polynomial $z(1+\beta-z)^2$, we see that at the minimum

$$|\gamma| + |1+\beta| \leq |\gamma(1+\beta-\gamma)^2| + |1+\beta|,$$
$$1 \leq |1+\beta-\gamma|$$

and so $|\gamma| + |1+\beta| \geq 1$. The zero locus we need to consider this time is the set $\{\gamma = 1+\beta\}$; these groups are Nielsen equivalent to groups generated by elliptics of order 2 and 3 [14]. In particular, this gives us the following inequality which we will use later.

LEMMA 2.2. *Let $\langle f, g\rangle$ be a Kleinian group. Then*

$$|\gamma| + |1+\beta| \geq 1$$

unless $\gamma = 1+\beta$ (*in which case it follows that fg or fg^{-1} is elliptic of order* 3).

As a consequence, if f has order 6, then $\beta = -1$ and we have

COROLLARY 2.1. *If $\langle f, g\rangle$ is a Kleinian group and f is elliptic of order* 6, *then*

$$|\gamma(f,g)| \geq 1. \tag{23}$$

This result is entirely analogous to the Shimizu-Leutbecher inequality for groups with one generator parabolic [1]. More generally, each good word gives rise to some such inequality in an analogous fashion.

There are a couple of further points we wish to make here. Suppose that we have eliminated a certain region from the possible values for $(\gamma, \beta, -4)$ among Kleinian groups. Then $(p_w(\gamma,\beta), \beta, -4)$ also cannot lie in this region. Thus, for instance, from Jørgensen's inequality we have

LEMMA 2.3. *Let $\langle f, g\rangle$ be a Kleinian group. Then*

$$|p_w(\gamma,\beta)| + |\beta| \geq 1$$

unless $p_w(\gamma,\beta) = 0$.

A useful special case (using (20)) is

LEMMA 2.4. *Let $\langle f, g\rangle$ be a Kleinian group. Then*

$$|\gamma(\gamma-\beta)| + |\beta| \geq 1 \tag{24}$$

unless $\gamma = \beta$.

The restrictive condition $p_w(\gamma,\beta)=0$ of Lemma 2.3 implies that f and $w=w(f,g)$ share a fixed point on $\hat{\mathbb{C}}$. We may use discreteness and geometry to find implications of this equation that can be used to eliminate it. For instance, if f is loxodromic, discreteness implies that w is not parabolic (see [1]) and that w and f have the same axis. Thus $[f,w]=\text{Identity}$, identifying a relation in the group. If f has some other special property such as being primitive, then w is either elliptic or a power of f, giving further relations. These additional properties of w in turn place greater constraints on γ and β.

The parameters of the commutator $[f,g]$ are easily derived from those of f and g. Specifically,

$$\beta([f,g]) = \operatorname{tr}^2[f,g]-4 = (\gamma+2)^2-4 = \gamma(\gamma+4) \tag{25}$$

and

$$\gamma([f,g],f) = \gamma(f, gf^{-1}g^{-1}) = \gamma(f, gfg^{-1}) = \gamma(\gamma-\beta), \tag{26}$$

using the fact that $[a,ab]$, $[a,b]$ and $([a,b^{-1}])^{-1}$ are all conjugate, together with (20) for the last equation.

3. Orthogonal axes

The aim of this and the next three sections is to prove Theorem 1.3. We first prove the special case that f and g have the same trace. In that case we can say somewhat more.

THEOREM 3.1. *Let f and g be loxodromic transformations whose axes meet at right angles and for which $\beta(f)=\beta(g)$. If $\langle f,g\rangle$ is discrete, then their translation lengths satisfy*

$$\tau_f = \tau_g \geq 1.06 \tag{27}$$

unless

$$\tau_f \in \{\lambda_\perp, 0.8538\ldots, 0.8812\ldots, 1.0098\ldots, 1.045\ldots, 1.0594\ldots\}.$$

Each of the above values is attained in a two-generator arithmetic lattice.

Proof. Let ϕ be an elliptic element of order 4 whose axis is orthogonal to the hyperbolic plane spanned by the axes of f and g, and that passes through the point of intersection of these two axes. Then

$$\phi f\phi^{-1} = g^{\pm 1}, \quad \phi^2 f\phi^{-2} = f^{-1}, \quad \phi^2 g\phi^{-2} = g^{-1},$$

and thus the group $\langle f,g,\phi\rangle = \langle f,\phi\rangle$ contains the group $\langle f,g\rangle$ with index at most 4. Hence $\langle f,\phi\rangle$ is discrete if and only if $\langle f,g\rangle$ is discrete.

Now $\psi = f\phi f^{-1}$ and ϕ are two elliptic elements of order 4 in a discrete group. The axis of f forms the common perpendicular between the axes of ϕ and ψ. Clearly $\tau_f = \rho(\psi,\phi)$. We are now in a position to apply our knowledge of the initial part of the spectra of distances between the axes of elliptics of order 4 generating a discrete group (see [12]), and this is the list presented above. Furthermore, the arithmeticity of the groups in question is decided in [10]. The smallest value is obtained in the arithmetic lattice generated by two elements φ of order 2 and ψ of order 4 with $\gamma(\varphi, \psi) = (-1 + i\sqrt{3})/2$. □

Actually, there is another relevant result here whose proof is more or less the same, once one adds the additional four-fold symmetries through the axes. One needs then to observe that two octahedral subgroups that contain a common element of order 4 are at distance at least $1.0594\ldots$; see [18] for this and related results for the other spherical triangle groups.

THEOREM 3.2. *Let f, g and h be loxodromic transformations whose axes all meet at right angles and for which $\beta(f) = \beta(g) = \beta(h)$. If $\langle f, g, h\rangle$ is discrete, then*

$$\tau_f = \tau_g = \tau_h \geq 1.0594\ldots. \tag{28}$$

This value is sharp and occurs when $\langle f, g\rangle$ (and hence $\langle f, g, h\rangle$) is a specific arithmetic lattice.

If z is a complex root of $z^3 + z^2 - z + 1$, then the number $1.0594\ldots$ is the real part of $2\operatorname{arcsinh}\left(\sqrt{z/2}\right)$.

In the same vein we have the following

LEMMA 3.1. *Let f and g be loxodromic transformations whose axes are distinct and meet (possibly on the boundary), and for which $\beta(f) = \beta(g)$. If there are elliptics ϕ and χ of order 2 that have the same axes as f and g respectively, and $\langle f, g, \phi, \chi\rangle$ is discrete, then*

$$\tau_f = \tau_g \geq \lambda_\perp. \tag{29}$$

Proof. By discreteness, the axes can only meet at an angle of π/n ($n = 2, 3, \ldots$). If ρ is rotation through an angle of π/n around the common perpendicular of $\operatorname{ax}(f)$ and $\operatorname{ax}(g)$, then $\langle f, g, \phi, \chi, \rho\rangle$ is still discrete, and the intersection point of $\operatorname{ax}(f)$ and $\operatorname{ax}(g)$ is on the axis of an elliptic of order $2n$ in this group. Two such axes must be at distance at least $\lambda_\perp$ [12], which gives (29). □

The previous three arguments have used the special symmetry of the situation in which we have equal traces. Next we give another intriguing consequence of the hypotheses of Theorem 3.1, which points the way to the more general result since it does not use the inherent symmetry of the situation in quite the same manner.

THEOREM 3.3. *Let f and g be loxodromic transformations whose axes meet at right angles and for which $\beta(f)=\beta(g)$. If $\langle f,g\rangle$ is discrete, then for all $m\in\mathbb{Z}\setminus\{0\}$,*

$$|\beta(f^m)|\geq 1 \quad\text{and}\quad |4+\beta(f^m)|\geq 1. \tag{30}$$

Proof. Since f and g have axes that meet at right angles, we have

$$\gamma(f,g) = -\frac{\beta(f)\beta(g)}{4} = -\frac{\beta^2}{4}, \tag{31}$$

where $\beta=\beta(f)$. Define the quadratic polynomial

$$p_\beta(z) = z(z-\beta).$$

If $|\beta|<1$, then 0 is an attracting fixed point for the iteration of this polynomial, its only preimages being $0,\beta$. The critical point for this polynomial is $\beta/2$, and its forward image is the point $p_\beta(\beta/2)=\beta/2(\beta/2-\beta)=-\beta^2/4=\gamma(f,g)=\gamma$. Set

$$h_1 = gfg^{-1}, \quad\text{and}\quad h_{n+1}=h_nfh_n^{-1}.$$

Then we have the polynomial trace identities,

$$\gamma(f,h_n) = p_\beta^n(\gamma).$$

In particular, the sequence $\{2+p_\beta^n(\gamma)\}_{n=1}^\infty$ is a sequence of traces in the discrete group $\langle f,g\rangle$. Since p_β is quadratic, under iteration the critical point, and therefore its image γ, must converge to the fixed point (see [4])

$$p_\beta^n(\gamma)\to 0 \text{ as } n\to\infty.$$

This will eventually contradict Jørgensen's inequality (in fact this is easily seen to be a contradiction in itself), unless for some n we have $p_\beta^n(\gamma)=0$. If $n=1$, then $\gamma=\beta$ which implies $\beta=-4$ and f is elliptic of order 2, a contradiction. Otherwise we have $n\geq 2$ and $p_\beta^{n-1}(\gamma)=\beta$. Thus $\langle f,h_{n-1}\rangle$ is a discrete nonelementary group with parameters $\beta(f)=\beta(h_{n-1})=\beta$ and $\gamma(f,h_{n-1})=\beta$. There are many ways to proceed from here to a contradiction; see [14].

Next observe that if ϕ is an elliptic of order 2 sharing its axis with f and ψ is an elliptic of order 2 sharing its axis with g, then $\langle f,g,\phi,\psi\rangle$ contains $\langle f,g\rangle$ with index at most 4 and is therefore discrete. Then $q_1=f\phi$ and $g_1=g\psi$ are both loxodromic with perpendicular axes, and

$$\beta(q_1) = 4\sinh^2\left(\frac{\tau_f+i(\theta_f+\pi)}{2}\right) = -4\cosh^2\left(\frac{\tau_f+i\theta_f}{2}\right) = -4-\beta, \tag{32}$$

which, together with the first part, proves the result. $\square$

Actually the reader should see that there is considerably more here. If $|2+\beta| < 1$, then $1+\beta$ is an attracting fixed point of p_β. Groups with $\gamma = 1+\beta$ are Nielsen equivalent to groups generated by elliptics of orders 2 and 3. A sequence of γ values converging to $1+\beta$ is not possible. (Use the trace identity $\gamma(f, gfg^{-1}fg) = \gamma(1+\beta-\gamma)^2$ to find a sequence converging to 0.) The argument of Theorem 3.3 produces

THEOREM 3.4. *Let f and g be loxodromic transformations whose axes meet at right angles and for which $\beta(f) = \beta(g)$. If $\langle f, g\rangle$ is discrete and contains no elliptics of order* 3, *then for all $m \in \mathbb{Z}\setminus\{0\}$,*

$$|2+\beta(f^m)| \geq 1. \tag{33}$$

Notice that this inequality is symmetric under the involution $\beta \leftrightarrow -4-\beta$ that we used in the previous result.

Many more arguments of this type are possible; essentially one only needs that $\frac{\beta}{2} - \frac{\beta^2}{4}$ lies in the hyperbolic part of the Mandelbrot set, and so this iteration procedure converges to an attracting fixed point. Moreover, the argument only requires that $\gamma = -\beta(f)\beta(g)/4$ lies in the interior of the filled-in Julia set of the polynomial.

Unfortunately, while this approach is effective for certain cases, estimating the size of filled-in Julia sets is quite difficult, although it is this process that in many ways underpins our search. We shall instead seek polynomial trace identities that bound parameters for discrete groups and turn this information into estimates on the geometric quantities, such as translation length.

4. Proof of Theorem 1.3

The proof of Theorem 1.3 involves a simple computer search. The main computation is described in Section 6. Here we establish some preliminary results.

4.1. *An extension.* Let f and g be as in the statement of Theorem 1.3. We begin by observing that the symmetry of the situation allows a reduction of the possible holonomies of the loxodromics f and g, just as in Theorem 3.3.

Let ϕ be the elliptic of order 2 that shares an axis with f, and let χ be the elliptic of order 2 that shares its axis with g. Then, of course,

$$\phi f = f\phi, \quad \chi g = g\chi, \quad \phi g = g^{-1}\phi, \quad \chi f = f^{-1}\chi,$$

and these relations show that the group $\langle f, g, \chi, \phi\rangle$ contains $\langle f, g\rangle$ with index at most 4, whence both groups are discrete and nonelementary. We may therefore replace f by $f\phi$ and/or g by $g\phi$, keeping the hypotheses of the theorem intact, but with the better holonomy bounds

$$0 \leq \eta_f \leq \pi/2, \quad -\pi/2 \leq \eta_g \leq \pi/2, \tag{34}$$

where, by symmetry (more precisely, by conjugating the group, if necessary, by a reflection in a plane containing $ax(f)$), we have assumed that $\eta_f \geq 0$.

Indeed, somewhat more is true. Since $\tau_{f^k\phi^\varepsilon} = k\tau_f$, we may replace f by any element (not of order 2) in the group $\langle f, \phi \rangle$, and similarly for g. This is especially useful if the holonomy of f or g is large, but the translation length is small.

Next, if

$$f_1 \in \langle f, \phi \rangle, \qquad g_1 \in \langle g, \chi \rangle, \tag{35}$$

then (7) gives

$$\gamma(f_1, g_1) = -\frac{\beta(f_1)\beta(g_1)}{4}. \tag{36}$$

4.2. *Small translation length.* We now eliminate the case where either of the translation lengths is small.

LEMMA 4.1. *Let f and g be loxodromic transformations generating a discrete group such that the axes of f and g meet at right angles. If* $\min\{\tau_f, \tau_g\} \leq 0.215$, *then*

$$\max\{\tau_f, \tau_g\} \geq \lambda_\perp. \tag{37}$$

Proof. We choose f_1, g_1 as in (35) (that is, one element from the stabiliser of each axis) and consider $h = g_1 f_1 g_1^{-1}$. Then we set

$$\gamma' = \gamma(f_1, h) = \gamma(\gamma - \beta), \quad \text{where} \quad \gamma = \gamma(f_1, g_1), \quad \beta = \beta(f_1).$$

Moreover, (5) gives

$$\beta(g_1^2) = \beta(g_1)(\beta(g_1) + 4). \tag{38}$$

The group $\langle f_1, h \rangle$ will be nonelementary, and so Jørgensen's inequality tells us that $|\gamma'| + |\beta| \geq 1$, whence using (36) and (38) we deduce

$$\begin{aligned} \left|\frac{\beta(f_1)\beta(g_1)}{4}\right| \left|-\frac{\beta(f_1)\beta(g_1)}{4} - \beta(f_1)\right| + |\beta(f_1)| &\geq 1, \\ |\beta(f_1)| \left(|\beta(f_1)| \left|\frac{\beta(g_1)(\beta(g_1)+4)}{4}\right| + 4\right) &\geq 4, \\ |\beta(f_1)| \left(|\beta(f_1)| \left|\frac{\beta(g_1^2)}{4}\right| + 4\right) &\geq 4. \end{aligned} \tag{39}$$

We now consider what happens for various choices of g_1. First, we put $g_1 = g$. If we assume $\tau_g \leq \lambda_\perp$, we have

$$\frac{|\beta(g_1^2)|}{4} \leq \cosh^2(\tau_g) \leq 1.8661$$

so that from (39),

$$|\beta(f_1)| \geq 0.7426. \tag{40}$$

This estimate is akin to the estimate $|\beta| > 1$, which we easily obtained for the case of equal traces. Of course by symmetry, (40) also holds with g_1 replacing f_1. If we assume that $\tau_f \leq 0.215$ and put

- $f_1 = f$ if $0 \leq \eta_f \leq 0.86$,
- $f_1 = f^3\phi$ if $0.86 \leq \eta_f \leq 1.2$, and
- $f_1 = f^2\phi$ if $1.2 \leq \eta_f \leq \pi/2$,

then for each such choice we obtain $|\beta(f_1)| < 0.7426$, which contradicts (40). □

5. **The extremal group**

The statement of Theorem 1.3 contains a claim regarding sharpness, which is attained by the arithmetic torsion-free lattice $\langle f, g\rangle$ with complex parameters

$$\beta(f) = \beta(g) = -3 + i\sqrt{3}, \quad \gamma(f,g) = \frac{3}{2}\left(-1 + i\sqrt{3}\right) = -\frac{\beta(f)^2}{4}, \tag{41}$$

for which $\tau_f = \tau_g = \lambda_\perp$ (so that this group also gives sharpness in the special case of equal traces dealt with in Theorem 3.1).

Our method of proof for Theorem 1.3 will be to search through the values of (β_f, β_g) corresponding to groups generated by two loxodromics, f and g, with orthogonal axes, both with translation length less than $\lambda_\perp$, and show that none of these groups is discrete. The sharp example is on the boundary of this region. The next result isolates this extremal group in the space of discrete groups so that we do not have to search near it.

Theorem 5.1. *Let $\langle f, g\rangle$ be a Kleinian group generated by two loxodromics with orthogonal axes, each with translation length less than $\lambda_\perp$. Then*

$$\begin{aligned} &\max\left\{|\beta_f - (-3 + i\sqrt{3})|, |\beta_g - (-3 + i\sqrt{3})|\right\} \geq 0.34 \ \text{ and} \\ &\max\left\{|\beta_f - (-3 - i\sqrt{3})|, |\beta_g - (-3 - i\sqrt{3})|\right\} \geq 0.34. \end{aligned} \tag{42}$$

If, moreover, f and g each share their axes with order 2 elliptics, then

$$\begin{aligned} &\max\left\{|\beta_f - (-1 + i\sqrt{3})|, |\beta_g - (-1 + i\sqrt{3})|\right\} \geq 0.34 \ \text{ and} \\ &\max\left\{|\beta_f - (-1 - i\sqrt{3})|, |\beta_g - (-1 - i\sqrt{3})|\right\} \geq 0.34. \end{aligned} \tag{43}$$

Proof. We suppose that the + sign is chosen in (42). (The other case follows by an obvious symmetry.) We let f and g satisfy the hypotheses of the theorem and suppose that β_f and β_g both lie in the disk D with center

$-3+i\sqrt{3}$ and radius 0.34. By (8), both are also in the interior of $2(E_{\lambda_\perp}-1)$, where, for $s>0$, E_s is the region bounded by the ellipse

$$E_s = \{z : |z-1|+|z+1| \le 2\cosh(s)\}. \tag{44}$$

The point $-3+i\sqrt{3}$ is on the boundary of $2(E_{\lambda_\perp}-1)$, so that β_f and β_g lie beneath the tangent line to the ellipse at this point; a simple calculation shows that this line has slope $2-\sqrt{3}=\tan(\pi/12)$. We set

$$\beta_f = 4(z-1), \quad \beta_g = 4(w-1), \quad \gamma(f,g) = -\frac{\beta_f\beta_g}{4},$$

where $\gamma(f,g)$ is given by (36); thus z and w lie in the open half-disk H with center $c_0=(1+i\sqrt{3})/4$ and radius 0.085, (strictly) beneath the line through c_0 with slope $\tan(\pi/12)$.

Let a,b be a pair of matrices in $\mathrm{PSL}(2,\mathbb{C})$ representing f and g respectively. Then, up to conjugacy, we have

$$a = \begin{pmatrix} \sqrt{z} & \sqrt{z-1} \\ \sqrt{z-1} & \sqrt{z} \end{pmatrix}, \quad b = \begin{pmatrix} \sqrt{w}+\sqrt{w-1} & 0 \\ 0 & \sqrt{w}-\sqrt{w-1} \end{pmatrix}.$$

Motivated by the relators in our extremal group (see the statement of Theorem 1.3), we let h and k be the transformations with respective matrix representations

$$W = ab^{-1}ababa^{-1}bab, \tag{45}$$
$$V = bab^2aba^{-1}b^2a^{-1}, \tag{46}$$

whence

$$\begin{aligned}\beta(h) &= 4wz(5-8z+4w(z(7-4w)-2+4z^2(w-1)))^2-4,\\ \beta(k) &= 16(w-1)(1+(4w-2)z)^2\\ &\qquad \times (w+4(w-1)(2w-1)z+4(1-2w)^2(w-1)z^2)\end{aligned}$$

and $\gamma(h,k)=-16(z-1)(w-1)[1+(4w-2)z]^2p(z,w)$, where

$$\begin{aligned}p(z,w) = w &+ 4(1-5w+10w^2-8w^3)z\\ &+16w(-4+18w-25w^2+12w^3)z^2\\ &+64w(5-35w+84w^2-86w^3+32w^4)z^3\\ &+256(w-1)^2w(-1+11w-27w^2+20w^3)z^4\\ &+1024(w-1)^3(2w-1)^2w^2z^5.\end{aligned}$$

Similarly, we define $\varepsilon(h,k)$ to be the sum of the elements of the subsidiary diagonal of $[V,W]$. We have $\varepsilon(h,k)=16(w-1)(1+(4w-2)z)\sqrt{z(z-1)}q(z,w)$, where

$$\begin{aligned} q(z,w) = {} & 1-3w+4w^2-2(4-21w+64w^2-104w^3+64w^4)z \\ & +8(1+17w-126w^2+280w^3-264w^4+96w^5)z^2 \\ & +32w(-26+238w-777w^2+1196w^3-888w^4+256w^5)z^3 \\ & +128w(9-117w+520w^2-1114w^3+1254w^4-712w^5+160w^6)z^4 \\ & +512w(w-1)^2(-1+19w-87w^2+162w^3-128w^4+32w^5)z^5 \\ & -2048w^2(1-3w+2w^2)^3z^6. \end{aligned}$$

We will show that whenever $z, w \in H$, some subgroup of $\langle f, g\rangle$ is not discrete. We first suppose that $\gamma(h,k) = \varepsilon(h,k) = 0$, whence either $z = 1$, $w = 1$, $z = 0$, $1 + (4w-2)z = 0$ or $p(z,w)$ and $q(z,w)$ vanish simultaneously. The first three cases are impossible, since $0, 1 \notin H$. Eliminating z from the equations $p(z,w) = 0$ and $q(z,w) = 0$ gives

$$(w-1)^2(1-2w+4w^2)^2(48w^3-88w^2+49w-8) = 0,$$

whose solutions again lie outside H; thus $1 + (4w-2)z = 0$.

In this case we consider the group $\langle h, g\rangle$. We calculate

$$\beta(h) = -\frac{2(4w-1)(4w^2-5w+2)(4w^2-2w+1)^2}{(2w-1)^5},$$

$$\gamma(h,g) = \frac{2(w-1)(4w-1)(4w^2-2w+1)^2}{(2w-1)^5}.$$

Now $\gamma(h,g)$ has no roots in H, and a simple calculation gives $|\beta(h)|+|\gamma(h,g)| < 1$ when $|w-c_0| = 0.085$; thus, by the maximum modulus principle, applied to the subharmonic function $|\beta|+|\gamma|$, this inequality remains true when $|w-c_0| \le 0.085$. Thus, by Jørgensen's inequality, $\langle h, g\rangle$ is not discrete. Note that in the case $z = w = c_0$, we get the extremal group (41), for which h and k both reduce to the identity.

We next suppose that $\gamma(h,k) = 0$, but $\varepsilon(h,k) \neq 0$, whence $p(z,w) = 0$. In this case the axes of h and k (if neither of these are parabolic) are distinct, and by (7), meet on the sphere at infinity. In particular, if $\langle h, k\rangle$ is discrete, then either at least one of h and k is parabolic, or by the classification of the elementary groups [1], both are elliptic of order $2, 3, 4$ or 6; that is,

$$\beta(h) = 0 \quad\text{or}\quad \beta(k) = 0 \quad\text{or}\quad \beta(k) \in \{0, -1, -2, -3, -4\}. \tag{47}$$

An easy computation shows that the pairs of equations $p(z,w) = \beta(h) = 0$, $p(z,w) = \beta(k) = 0$ and each pair of equations $p(z,w) = 0$ and $\beta(k) = -n$, for $n \in \{0, 1, 2, 3, 4\}$, have no solution in $(z,w) \in H^2$; thus $\langle h, k\rangle$ cannot be discrete in this case.

Since we have now dealt with the case $\gamma(h,k)=0$, it now suffices, by Jørgensen's inequality, to show that

$$|\beta(h)|+|\gamma(h,k)|<1 \tag{48}$$

when $(z,w)\in H\times H$. Again, by the maximum modulus principle, we need only prove this inequality on the product of the boundaries, which is a straightforward maximization problem in two real variables. This completes the proof of (42).

To prove (43), if f and g share their axes with order 2 elliptics ϕ and χ, respectively, then we apply (42) to the loxodromics $f\phi$ and $g\chi$, and we use $\beta(f)+\beta(f\phi)=\beta(g)+\beta(g\chi)=-4$. □

Remark. We need to use the half-disk H rather than the disk only when proving (48); even here a whole disk can be used but with a slightly smaller radius (0.3 rather than 0.34).

6. More tests

So far, in the above we have only used Jørgensen's inequality to generate constraints on parameters. However, to get to what we want we must develop more subtle discreteness tests. The first is a sharp generalisation [11], [3] of another result of Jørgensen [21], who found universal constraints on discrete groups generated by elements with the same trace.

LEMMA 6.1. *Let f and g generate a discrete nonelementary group. Let $\gamma=\gamma(f,g)$ and $\beta\in\{\beta_f,\beta_g\}$. If $\beta\neq -4$ and $\beta\neq\gamma$, then*

$$|\gamma(\gamma-\beta)|\geq 2-2\cos(\pi/7)=0.198\ldots. \tag{49}$$

This inequality is sharp and uniquely attained in the (2,3,7)*-triangle group.*

Notice that if f and g have axes meeting at right angles, then (36) identifies γ, and if we put $\beta'=\beta_g$, the inequality reads as

$$|\beta^2||\beta'(\beta'+4)|=|\beta(f)|^2|\beta(g^2)|\geq 16\times 0.198=3.168, \tag{50}$$

which has obvious implications.

We need a few more tools to prove Theorem 1.3. First we identify the search space we will use in the proof.

LEMMA 6.2. *If f and g are loxodromics whose axes meet at right angles that generate a discrete nonelementary, and if* $\max\{\tau_f,\tau_g\}<\lambda_\perp$, *then there* *exist such loxodromics for which each $\beta \in \{\beta_f, \beta_g\}$ satisfies each of*

$$\begin{gathered}\beta \in 2(E_{\lambda_\perp} - 1), \;\; \beta \notin 2(E_{0.215} - 1), \;\; \Re e(\beta) \geq -2, \; |\beta| \geq 0.7426, \\ \max\{|\beta_f + 1 + \sqrt{3}i|, |\beta_g + 1 + \sqrt{3}i|\} \geq 0.34 \;\; \text{and} \\ \max\{|\beta_f + 1 - \sqrt{3}i|, |\beta_g + 1 - \sqrt{3}i|\} \geq 0.34, \end{gathered} \tag{51}$$

where E_s is defined at (44).

Proof. The first inequality follows from the hypothesis $\max\{\tau_f, \tau_g\} < \lambda_\perp$ and (5), and the second similarly from Lemma 37. By (34), we may assume that holonomies of f and g do not exceed $\pi/2$ in absolute value, whence (5) again gives the third inequality. The remaining inequalities follow from (40) and (43). □

Since the space of discrete nonelementary groups is closed, it follows that the space of discrete groups we wish to discuss is compact. Hence there are loxodromics f and g with axes at right angles, such that the group $\langle f, g\rangle$ is discrete and the sum $\tau_f + \tau_g$ is minimized. We will show that, when β_f, β_g satisfy (51), it is always possible to find another such pair of loxodromics, whose translation lengths have a smaller sum than $\tau_f + \tau_g$, a contradiction. Lemma 6.5 below shows how this contradiction can be established computationally.

The argument depends on constructing loxodromics as a product of two order-2 elliptics. The first lemma below shows how we can append such elliptics to a group, while preserving discreteness.

LEMMA 6.3. *If Γ is discrete group, $f \in \Gamma$ is not parabolic or the identity, $h \in \Gamma$, $\tilde{f} = hfh^{-1}$ and ψ, $\tilde{\psi}$ are the two elliptics of order 2 which interchange* $\mathrm{ax}(f)$ *and* $\mathrm{ax}(\tilde{f})$*, then the group Γ_1 generated by ψ and $\tilde{\psi}$ together with the stabilizer of* $\mathrm{ax}(f)$ *in Γ is discrete. If* $\mathrm{ax}(\tilde{f})$ *is the closest translate of* $\mathrm{ax}(f)$ *in Γ, then it is also closest in Γ_1.*

Proof. Let k be in the stabilizer of $\mathrm{ax}(f)$ in Γ. We have $\psi k = (\psi k \psi^{-1})\psi = (hkh^{-1})^{\pm 1}\psi$. Similarly, $\tilde{\psi} k = (hkh^{-1})^{\mp 1}\tilde{\psi}$. Thus, by advancing ψ and $\tilde{\psi}$ to the right of each word, every isometry in Γ_1 can be written α, $\alpha\psi$, $\alpha\tilde{\psi}$ or $\alpha\psi\tilde{\psi}$ (ψ and $\tilde{\psi}$ commute), where α is the group generated by the stabilizers of $\mathrm{ax}(f)$ and $\mathrm{ax}(\tilde{f})$. Hence Γ_1 is discrete and the translates of $\mathrm{ax}(f)$ in Γ_1 are also translates of $\mathrm{ax}(f)$ in Γ ($\psi(\mathrm{ax}(f)) = \tilde{\psi}(\mathrm{ax}(f)) = \mathrm{ax}(\tilde{f})$). □

LEMMA 6.4. *Let $f, g \in \mathrm{Isom}^+(\mathbb{H}^3)$, with f neither the identity or parabolic. Then*

$$\cosh(\rho(\mathrm{ax}(f), \mathrm{ax}(gfg^{-1})) = \left|\frac{\gamma(f,g)}{\beta(f)} - 1\right| + \left|\frac{\gamma(f,g)}{\beta(f)}\right|. \tag{52}$$

Proof. Let ψ be an order-2 elliptic that interchanges $\mathrm{ax}(f)$ and $\mathrm{ax}(gfg^{-1}) = g(\mathrm{ax}(f))$, chosen so that $\psi f^{-1}\psi = gf^{-1}g^{-1}$, whence $\gamma(f,\psi) = \gamma(f,g)$. Since $\rho(\mathrm{ax}(f), \mathrm{ax}(gfg^{-1}) = 2\rho(\mathrm{ax}(f), \mathrm{ax}(\psi))$, the lemma now follows from (12). $\square$

Now suppose that f and g are loxodromics whose axes meet at right angles, which generate a discrete nonelementary group with $\max\{\tau_f, \tau_g\} < \lambda_\perp$, and suppose that $\tau_f + \tau_g$ is minimized subject to these conditions.

As in Section 4.1, we may extend the group $\langle f, g\rangle$, retaining discreteness, by the addition of two elements of order 2, ϕ and χ sharing their axes with f and g respectively. We note that

$$\phi f = f\phi \quad \text{and} \quad \chi g = g\chi.$$

If $s \in \langle f, g\rangle$, then $\mathrm{ax}(sfs^{-1}) = \mathrm{ax}(s\phi s^{-1}) = s(\mathrm{ax}(f))$. Since $\mathrm{ax}(f)$ and $\mathrm{ax}(sfs^{-1})$ are also the axes of the order-2 elliptics, they must either be hyperparallel (i.e., $\rho(\mathrm{ax}(f), \mathrm{ax}(sfs^{-1})) > 0$) or coincide, by Lemma 3.1.

Suppose $\mathrm{ax}(f)$ and $\mathrm{ax}(sfs^{-1})$ are hyperparallel. Then, as in Lemma 6.3, we extend $\langle f, \phi, sfs^{-1}, s\phi s^{-1}\rangle$, by adding the order-2 elliptics ψ and $\tilde{\psi}$ that interchange $\mathrm{ax}(f)$ and $\mathrm{ax}(sfs^{-1})$, obtaining the discrete group

$$\langle f, \phi, sfs^{-1}, s\phi s^{-1}, \psi, \tilde{\psi}\rangle.$$

In this group ϕ and ψ are elliptics of order 2 with hyperparallel axes, so their product $g' = \psi\phi$ is loxodromic with translation length

$$\tau_{g'} = 2\rho(\mathrm{ax}(\phi), \mathrm{ax}(\psi)) = \rho(\mathrm{ax}(f), \mathrm{ax}(sfs^{-1})).$$

Moreover, $\mathrm{ax}(g')$ is perpendicular to $\mathrm{ax}(f)$, so we have a contradiction to minimality of $\tau_f + \tau_g$ if $\tau_{g'} < \tau_g$. Since $\mathrm{ax}(f)$ and $\mathrm{ax}(g)$ are perpendicular, we have $\tau_g = \rho(\mathrm{ax}(f), \mathrm{ax}(gfg^{-1}))$, and so by Lemma 6.4, this contradiction is equivalent to

$$\left|\frac{\gamma(f,g')}{\beta(f)} - 1\right| + \left|\frac{\gamma(f,g')}{\beta(f)}\right| < \left|\frac{\gamma(f,g)}{\beta(f)} - 1\right| + \left|\frac{\gamma(f,g)}{\beta(f)}\right|.$$

Setting $\beta_f = \beta(f)$, $\gamma' = \gamma(f, g')$, this is in turn equivalent to

$$|\gamma' - \beta_f| + |\gamma'| < |\gamma - \beta_f| + |\gamma|. \tag{53}$$

Now, $s \in \langle f, g\rangle$, and if $s = w(f,g)$, for some good word w, then $\gamma(f,w)$ is a polynomial with integer coefficients in the variables γ and β, say

$$\gamma' = p_w(\gamma, \beta),$$

in which case the contradiction (53) can be written

$$|p_w(\gamma, \beta_f) - \beta_f| + |p_w(\gamma, \beta_f)| < |\gamma - \beta_f| + |\gamma|. \tag{54}$$

Since f and g have axes at right angles, we know that $\gamma(f,g) = -\beta_f\beta_g/4$. Moreover, by Theorem 2.1, we may write $p_w(\gamma,\beta) = \gamma q_w(\gamma,\beta)$ so that (54) becomes

$$\left|\frac{\beta_f\beta_g}{4}q_w(\gamma,\beta_f)+\beta_f\right|+\left|\frac{\beta_f\beta_g}{4}q_w(\gamma,\beta_f)\right|<\left|\frac{\beta_f\beta_g}{4}+\beta_f\right|+\left|\frac{\beta_f\beta_g}{4}\right|. \tag{55}$$

Thence, eliminating common factors (neither f nor g is parabolic), we obtain

$$|\beta_g\, q_w(\gamma,\beta_f)+4|+|\beta_g\, q_w(\gamma,\beta_f)|<|\beta_g+4|+|\beta_g|\,. \tag{56}$$

This is a testable inequality that is equivalent to $\tau_{g'} < \tau_g$ when $s = w(f,g)$. Thus, should it hold, we would have a contradiction to our choice of minimality, provided $\mathrm{ax}(sfs^{-1}) \neq \mathrm{ax}(f)$, which holds when $q_w(\gamma,\beta_f) \neq 0$.

Thus we have

Lemma 6.5. *Let $\langle f,g\rangle$ be a discrete group generated by two loxodromics with perpendicular axes, and let $\tau_f+\tau_g$ be minimal. Set $\gamma=\gamma(f,g)$. Then, for every good word polynomial p_w,*

$$q_w(\gamma,\beta_f)=0 \text{ or } |\beta_g\, q_w(\gamma,\beta_f)+4|+|\beta_g\, q_w(\gamma,\beta_f)|\geq|\beta_g+4|+|\beta_g|\,. \tag{57}$$

Here

$$q_w(\gamma,\beta)=p_w(\gamma,\beta)/\gamma.$$

We set $t_w(z_1,z_2)$ to be the smaller value in the set

$$\left\{|z_2+4|+|z_2|-\left|z_2q_w\left(-\frac{z_1z_2}{4},z_1\right)+4\right|\right.$$
$$\left.-\left|z_2q_w\left(-\frac{z_1z_2}{4},z_1\right)\right|,\left|q_w\left(-\frac{z_1z_2}{4},z_1\right)\right|\right\}.$$

Of course, $t_w(\beta_f,\beta_g)\leq 0$ is equivalent to (57).

We note that, for f, g as in Lemma 6.5, (57) must also hold when f is replaced by ϕf and/or g is replaced by χg. By (5), this means replacing β_f and β_g by $-4-\beta_f$ and $-4-\beta_g$ respectively. Of course, by symmetry (57) must also hold with f and g interchanged. Thus to prove Theorem 1.3, it now suffices to show that each (β_f,β_g) that satisfies (51) also satisfies the inequality

$$\max\{t_w(\beta_f,\beta_g),t_w(\beta_g,\beta_f),t_w(-4-\beta_f,\beta_g),t_w(\beta_g,-4-\beta_f),$$
$$t_w(\beta_f,-4-\beta_g),t_w(-4-\beta_g,\beta_f)\}>0, \tag{58}$$

some good word w.

In fact, we have found that (58) holds for each such point (β_f,β_g) for at least one of the twenty-one words w of the form (16) with $s_1=1$ and

$(r_1, r_2, \ldots r_{m-1})$ in the set

$$\{(1), (-3,1), (-1,-1), (-1,1), (-1,-1,-1), (-3,1,-3,1), (-1,-1,-1,-1),\\ (-1,-1,-1,1), (-1,1,-1,-1,1), (-1,-1,1,1,1,-1,-1), (-1,1,1,-1,1,1,-1),\\ (-2,1,1), (-1,-1,1), (-1,-1,-1,1,1,1,-1), (-1,-1,-1,1,1,1),\\ (-1,-1,1,-1,-1,1,1,1,0), (-1,-1,1,-1,-1,-1,0), (-1,-1,1,1),\\ (-1,1,1,-1), (-2,-1,1,1,1,-1), (-1,-1,1,-1,-1,1)\}.$$

This is a computation of the sort described in Section 1. Some technical details are given in Appendix 1.

7. Distances between geodesics

In this section we collect some useful formulas for the complex distance between two geodesics. We first define some notation.

If γ is a geodesic and ϕ is an isometry of $\mathbb{H}^3$, then $\phi(\gamma)$ will denote the image of γ under ϕ; if $\phi(z) = \lambda z$ ($\lambda \in \mathbb{C} \setminus \{0\}$), then we abbreviate this to $\lambda\gamma$. (Here and elsewhere we tacitly identify a Möbius transformation with its Poincaré extension.) If γ has endpoints z_1, z_2, and is oriented from z_1 to z_2, then we assume that $\phi(\gamma)$ is oriented from $\phi(z_1)$ to $\phi(z_2)$. We let I denote the geodesic with endpoints 0 and ∞, and abbreviate the complex distance $\Delta(\gamma, I)$ to $\Delta(\gamma)$. If ϕ is an orientation preserving isometry, note that $\Delta(\phi(\gamma_1), \phi(\gamma_2)) = \Delta(\gamma_1, \gamma_2)$ and, in particular,

$$\Delta(\lambda\gamma) = \Delta(\gamma).$$

THEOREM 7.1 ([24]). *If γ and γ' are geodesics in $\mathbb{H}^3$, oriented respectively from endpoint z_1 to endpoint z_2 and from endpoint w_1 to endpoint w_2, then*

$$\sinh^2\left[\frac{1}{2}\Delta(\gamma, \gamma')\right] = \frac{(z_1 - w_1)(w_2 - z_2)}{(w_1 - w_2)(z_1 - z_2)}. \tag{59}$$

In particular,

$$\sinh^2[\Delta(\gamma, I)] = \frac{4z_1 z_2}{(z_1 - z_2)^2}, \tag{60}$$

$$\sinh^2\left[\frac{1}{2}\Delta(\gamma, e^{\beta}\gamma)\right] = \frac{-4z_1 z_2}{(z_1 - z_2)^2}\sinh^2(\beta/2) \tag{61}$$

and so, combining these,

$$\sinh\left[\frac{1}{2}\Delta(\gamma, e^{\beta}\gamma)\right] = \pm i \sinh(z)\sinh(\beta/2), \tag{62}$$

where $z = \Delta(\gamma, I) = \Delta(e^{\beta}\gamma, I)$.

(Note that the orientation of I above is immaterial.)

8. **Volume estimates from collars**

The following gives an initial estimate of co-volume from collaring radius. It appeared, in a more general form, in [23] and it is the "collar-volume" estimate we referred to in the introduction. For $k \in \mathbb{N}$, and nonimaginary $w \in \mathbb{C}$, we let $\Lambda_k(w)$ denote the lattice in $\mathbb{C}$ generated by w and $2\pi i/k$. We set, for $\delta > 0$,

$$\mathcal{A}(\delta + i\theta) = \left\{ 2\text{arcsinh}\left(\frac{i \sinh\left(\frac{\tau + i\psi}{2}\right)}{\sinh(\delta + i\theta)} \right) : \tau \in [0, \delta],\ \psi \in [0, 4\pi] \right\}. \tag{63}$$

If $\Delta(L, I) = \delta + i\theta$, then, by (62), $\rho(L, e^{\beta}L) \le \delta$ if and only if $\beta \in \mathcal{A}(\delta + i\theta) + 2\pi i k$ $(k \in \mathbb{Z})$.

THEOREM 8.1. *Let Γ be a Kleinian group, ℓ_0, a geodesic whose stabilizer in Γ is generated by a loxodromic with translation length τ, possibly an elliptic of order k that fixes the endpoints of ℓ_0 (in the absence of such an elliptic we set $k = 1$) and possibly an elliptic of order 2 that interchanges the endpoints of ℓ_0. If ℓ is a closest translate of ℓ_0, with $\Delta(\ell_0, \ell) = \delta + i\theta$, with $\delta > 0$, and $\mathcal{E} \subseteq \mathcal{A}(\delta + i\theta)$ is convex and centrally balanced, then*

$$\tau \ge \frac{k}{8\pi} \text{Area } (\mathcal{E}),$$

and hence

$$\text{vol}(\mathbb{H}^3/\Gamma) \ge \frac{1}{16} \sinh^2(\delta/2) \text{Area } (\mathcal{E}).$$

Proof. We may assume that $\ell_0 = I$ so that $\text{Stab}(I) \supseteq \langle g_1, g_2 \rangle$, where $g_1(u) = e^{\tau + i\eta}u$, $g_2(u) = e^{2\pi i/k}u$, and $\eta \in (-\pi i/k, \pi i/k]$. Since the distance between any distinct translates of ℓ is at least δ, there is no nonzero point of $\Lambda_k(\tau + i\eta)$ in the interior of $\mathcal{E}$. Minkowski's theorem now gives

$$\text{Area } (\mathcal{E}) \le 4\text{Area}(\mathbb{C}/\Lambda_k(\tau + i\eta)) = 8\pi\tau/k,$$

and the theorem follows. □

The next two results enable us to apply this theorem.

LEMMA 8.1. *For all real θ and $\delta \ge \frac{1}{2}\text{arccosh}(11/5) = 0.712708\ldots$, the region $\mathcal{A}(\delta + i\theta)$ is convex. This value is sharp.*

Proof. $\mathcal{A}(\delta + i\theta)$ is bounded by the curve

$$f(\psi) = \text{arcsinh}\left(\frac{i \sinh\left(\frac{\delta + i\psi}{2}\right)}{\sinh(\delta + i\theta)} \right) \quad (\psi \in [0, 4\pi]),$$

hence it is convex if and only if $\Im m(f''(\psi)/f'(\psi)) \geq 0$ for all ψ. Since

$$f''(\psi)/f'(\psi) = \frac{i\cosh^2(\delta+i\theta)\sinh\left(\frac{\delta+i\psi}{2}\right)}{2\left[\cosh^2(\delta+i\theta)-\cosh^2\left(\frac{\delta+i\psi}{2}\right)\right]\cosh\left(\frac{\delta+i\psi}{2}\right)},$$

$\mathcal{A}(\delta+i\theta)$ is convex when, for all ψ, the real part of

$$(\cosh^2(\delta-i\theta)-\cosh^2((\delta-i\psi)/2))\cosh((\delta-i\psi)/2)\cosh^2(\delta+i\theta)\sinh((\delta+i\psi)/2)$$

is nonnegative. This real part is equal to $\sinh(\delta)X(\psi)/16$, where

$$\begin{aligned}X(\psi) = &-2\sin^2(\psi) - 2\cos(\psi)\cosh(\delta) + \cos(4\theta)\\ &+\cosh(4\delta) + 2\cosh(2\delta)\cos(\psi-2\theta)(\cos(\psi)-\cosh(\delta))\\ &+6\cosh(\delta)\sin(\psi)\sin(2\theta) + \sin(2\psi)\sin(2\theta).\end{aligned}$$

We have

$$\begin{aligned}\frac{\partial X}{\partial\psi} &= 2[\cosh(\delta)-2\cos(\psi)]\\ &\quad\times[\sin(\psi)+\cosh(2\delta)\sin(\psi-2\theta)-\cos(\psi)\sin(2\theta)-2\sin(2\theta)\cosh(\delta)],\end{aligned}$$

which is zero when $\cos(\psi) = \cosh(\delta)/2$, and also when both

(64)
$$\cos(\psi) = \frac{-\cosh^3(\delta)\sin^2(2\theta) \pm (1+\cos(2\theta)\cosh(2\delta))(\cosh^2(\delta)-\sin^2(\theta))/2}{(\sin^2(\theta)+\cosh^2(\delta))^2 - 4\sin^4(\theta)\cosh^2(\delta)}$$

and

(65)
$$\sin(\psi) = \frac{\cosh(\delta)\sin(2\theta)(1\pm\cosh(\delta))}{\sin^2(\theta)+\cosh^2(\delta)\pm 2\sin^2(\theta)\cosh(\delta)}.$$

Substituting (64) and (65) into X gives

$$\frac{(\cosh(\delta)\mp\cos(2\theta))(2\cosh(\delta)\pm 1)(\cos(2\theta)+\cosh(2\delta))^2}{\sin^2(\theta)+\cosh^2(\delta)\pm 2\sin^2(\theta)\cosh(\delta)},$$

which is positive for all θ and all $\delta > 0$. Thus $X(\psi) \geq 0$ for all ψ when $\cosh(\delta) \geq 2$, and the same inequality is true when $\cosh(\delta) < 2$ if it holds for $\cos(\psi) = \cosh(\delta)/2$.

The product $X(\psi)X(-\psi)$ is a polynomial in $\cos(\psi)$. Substituting $\cos(\psi) = \cosh(\delta)/2$ into this gives a polynomial $p(v,w)$ in $v = \cosh^2(\delta)$ and $w = \cos(2\theta)$:

$$\begin{aligned}p(v,w) = \frac{1}{4}\Big[&16w^4 + (8v-16v^2)w^3 + (-32+120v-63v^2+44v^3)w^2\\ &+(-8v-18v^2+100v^3-64v^4)w + 16 - 120v + 353v^2 - 592v^3\\ &+260v^4\Big].\end{aligned}$$

This polynomial is quartic in w and has discriminant

$$\Delta = -108v^6(1-4v)^2(v-4)^6(512-1856v+4440v^2-2687v^3+320v^4),$$

which has roots at $v = 8/5$, $v = 4$ and none in $(8/5, 4)$.

One easily checks that $p(4,w) > 0$ for all $w \in [-1,1]$ and $p(v,\pm 1) > 0$ for all $v \in (8/5, 4]$, whence by continuity, $p(v,w) > 0$ for all $w \in [-1,1]$ and $v \in (8/5, 4]$. It follows that, for all ψ and θ, $X(\psi) > 0$ when $v = \cosh^2(\delta) > 8/5$, whence $X(\psi) \geq 0$ when $\cosh^2(\delta) \geq 8/5$. For sharpness, observe that

$$p(v, 1/5) = (3/625)(5v-8)(2575v^3 - 1820v^2 + 700v - 96)$$

is negative for v in some interval $(a, 8/5)$. □

LEMMA 8.2. Area $(\mathcal{A}(\delta + i\theta)) \geq \dfrac{4\pi \sinh(\delta)}{e^{\delta}|\sinh(\delta + i\theta)|}$.

Proof. The region $\Omega = \frac{1}{2}\mathcal{A}(\delta + i\theta)$ is parametrized by

$$\left\{ \operatorname{arcsinh}\left(\frac{i \sinh\left(\frac{\tau + i\psi}{2}\right)}{\sinh(\delta + i\theta)} \right) : \tau \in [0, \delta],\ \psi \in [0, 4\pi] \right\}. \tag{66}$$

Thus, if $f(z) = \operatorname{arcsinh}\left(\frac{i\sinh(z)}{\sinh(\delta+i\theta)}\right)$ and A is the region $0 \leq x \leq \delta/2$, $0 \leq y \leq 2\pi$ $(z = x + iy)$, then

$$\begin{aligned} \operatorname{Area}(\Omega) &= \int_A |f'(z)|^2 \, dz = \int_A \frac{|\cosh^2(x+iy)|}{|\sinh^2(\delta+i\theta) - \sinh^2(x+iy)|} \, dx \, dy \\ &\geq \left| \int_A \frac{\sinh^2(x) + \cos^2(y)}{\sinh^2(\delta+i\theta) - \sinh^2(x+iy)} \, dx \, dy \right|. \end{aligned}$$

By the substitution $u = e^{iy}$, we obtain

$$\begin{aligned} &\int_0^{2\pi} \frac{\sinh^2(x) + \cos^2(y)}{\sinh^2(\delta+i\theta) - \sinh^2(x+iy)} dy \\ &\qquad = \int_C \frac{i(u^4 + 2u^2\cosh(2x) + 1)}{u(au^4 + bu^2 + c)} \, du, \end{aligned} \tag{67}$$

where C is the unit circle, $a = e^{2x}$, $b = -2[2\sinh^2(\delta + i\theta) + 1]$ and $c = e^{-2x}$. The denominator of the integrand factorizes as

$$au(u^2 - \exp[-2x + 2(\delta + i\theta)])(u^2 - \exp[-2x - 2(\delta + i\theta)]).$$

Thus the integrand has poles within C at $u = 0$ and $u = \pm \exp[-x - (\delta + i\theta)]$ and residues of ie^{2x} at $u = 0$ and

$$-i((e^{-2(x+\delta+i\theta)} + 2\cosh(2x) + e^{2(x+\delta+i\theta)}/(4\sinh(2z)))$$

at each of the other poles. By the residue theorem, the integral at (67) is

$$\frac{2\pi\cosh(2x)}{e^{\delta+i\theta}\sinh(\delta+i\theta)},$$

and the required inequality follows after another integration and taking absolute values. □

Theorem 8.1, together with Lemmas 8.1 and 8.2, now gives

LEMMA 8.3. *A Kleinian group with an axis with collaring radius*

$$r \geq \frac{1}{4}\operatorname{arccosh}(11/5) = 0.356354\ldots$$

has co-volume at least

$$\frac{\pi \sinh(2r)\sinh^2(r)}{4e^{2r}\cosh(2r)}.$$

This estimate is an increasing function of r. A simple calculation gives

COROLLARY 8.1. *The collaring radius of any axis of a Kleinian group whose co-volume is no more than V_0 is less than* 0.4075.

This gives Lemma 1.1 for $k \neq 1, 3$. In the next section we complete the proof by refining the collaring estimate of Lemma 8.3 in these remaining cases.

9. **Proof of Lemma 1.1**

Proof of Lemma 1.1. For a contradiction we will suppose that Γ has co-volume $\leq V_0$ and a loxodromic f with collaring radius $r_f \geq c_k$, translation length τ and holonomy η which we may assume to be in $(-\pi i/k, \pi i/k]$. In view of Corollary 8.1, we may assume that $k \in \{1, 3\}$ and that in these cases $r_f \in [c_k, 0.4075]$. We normalize so that $\operatorname{ax}(f) = I$, whence $\operatorname{Stab}(I) \supseteq \langle g_1, g_2 \rangle$, where $g_1(u) = e^{\tau + i\eta}u$ and $g_2(u) = e^{2\pi i/k}u$. Let $\tilde{f}$ be a conjugate of f in Γ with axis ℓ, chosen so that $\delta = \rho(I, \ell) = 2r_f$, and let $z = \Delta(I, \ell) = \delta + i\theta$, where, by symmetry, we may assume $\theta \in [0, \pi/2]$. By a further conjugation (using a loxodromic that fixes 0 and ∞), we may assume that the endpoints z_1 and z_2 of ℓ are mutually reciprocal, whence, using (60) and reindexing if necessary, we have $z_1 = \tanh(z/2)$, $z_2 = \coth(z/2)$.

The hypothesis that $\operatorname{co-vol}(\Gamma) \leq V_0$ gives the inequality

$$(0.91)^{-1}\frac{\pi\tau}{2k}\sinh^2(\delta/2) \leq V_0, \tag{68}$$

where here we use Przeworski's [27] upper bound of 0.91 for the packing density of hyperbolic tubes (or cylinders).

Now let $f(z) = e^w z$, and let Γ_1 be defined from Γ, f and $\tilde{f}$ as in Lemma 6.3, according to which the collaring radius of I in Γ_1 is still $\delta/2$. Explicitly,

$$\Gamma_1 \supseteq \langle g_1, g_2, \psi, \tilde{\psi} \rangle,$$

where $g_1(u) = e^{\tau + i\eta}u$, $g_2(u) = e^{2\pi i/k}u$,

$$\psi(u) = \frac{u - \tanh(z/2)}{\tanh(z/2)u - 1}, \quad \tilde{\psi}(u) = \frac{u - \coth(z/2)}{\coth(z/2)u - 1}.$$

We obtain the required contradiction by showing that there are two distinct translates of I in Γ_1 that are at distance less than δ from each other, hence a translate of I that is distance less than δ from I, contrary to the assumption that ℓ is a closest translate. For this purpose, we define for each lattice point $\alpha \in \Lambda_k(w)$ translates $\ell' = \ell'(z,\alpha)$ and $\ell^* = \ell^*(z,\alpha)$ by

$$\ell' = \psi(e^{\alpha}\ell) \quad \text{and} \quad \ell^* = \tilde{\psi}(e^{\alpha}\ell). \tag{69}$$

The endpoints of ℓ' are $\psi(e^{\alpha}\tanh(z/2))$ and $\psi(e^{\alpha}\coth(z/2))$, which simplify to

$$\frac{\sinh(\alpha/2)\sinh(z)}{\sinh(\alpha/2)\cosh(z)-\cosh(\alpha/2)}, \qquad \frac{\sinh(\alpha/2)\cosh(z)+\cosh(\alpha/2)}{\sinh(\alpha/2)\sinh(z)}. \tag{70}$$

Hence by (60), we have

$$\sinh^2[\Delta(\ell', I)] = 4\sinh^2(\alpha/2)\sinh^2(z)[\sinh^2(\alpha/2)\sinh^2(z)-1]. \tag{71}$$

The endpoints of ℓ^* are just the reciprocals of those of ℓ' (since $\tilde{\psi}\psi(z) = 1/z$), and so

$$\ell^* = \left(\frac{\cosh(z)\sinh(\alpha/2)-\cosh(\alpha/2)}{\cosh(z)\sinh(\alpha/2)+\cosh(\alpha/2)}\right)\ell'. \tag{72}$$

Now, for $\alpha\beta \in \Lambda_k(w)$, let

$$\begin{aligned}
E_1(z,\beta) &= i\sinh\left(\frac{1}{2}\Delta[\ell, e^{\beta}\ell]\right) = \pm\sinh(z)\sinh(\beta/2),\\
E_2(z,\alpha,\beta) &= i\sinh\left(\frac{1}{2}\Delta[\ell', e^{-\beta}\ell^*]\right)\\
&= \pm 2\sinh(z)\sinh(\alpha/2)[\sinh(\alpha/2)\sinh(\beta/2)\cosh(z)\\
&\qquad + \cosh(\alpha/2)\cosh(\beta/2)],\\
E_3(z,\alpha,\beta) &= \cosh(\Delta[\ell', e^{\beta}\ell'])\\
&= 1 + 8\sinh^2(z)\sinh^2(\alpha/2)\sinh^2(\beta/2)(1-\sinh^2(z)\sinh^2(\alpha/2)).
\end{aligned}$$

These are calculated using (62), (71) and (72). From these, we derive the real distance functions

$$\begin{aligned}
\mathcal{E}_1(z,\beta) &= \cosh\left(\frac{1}{2}\rho[\ell, e^{\beta}\ell]\right) = (|E_1(z,\beta)+1| + |E_1(z,\beta)-1|)/2,\\
\mathcal{E}_2(\alpha,\beta,z) &= \cosh\left(\frac{1}{2}\rho[\ell', e^{-\beta}\ell^*]\right) = (|E_2(\alpha,\beta,z)+1| + |E_2(\alpha,\beta,z)-1|)/2,\\
\mathcal{E}_3(\alpha,\beta,z) &= \cosh\left(\rho[\ell', e^{\beta}\ell']\right) = (|E_3(\alpha,\beta,z)+1| + |E_3(\alpha,\beta,z)-1|)/2.
\end{aligned}$$

We have the required contradiction if for each $z = \delta + i\theta$, $w = \tau + i\eta$, for which $2c_k \le \delta \le 0.815$, $0 \le \theta \le \pi/2$,

$$0 < \tau \le 2(0.91)kV_0/(\pi\sinh^2(\delta/2)) \quad \text{and} \quad -\pi i/k \le \eta \le \pi i/k, \tag{73}$$

there are some $\alpha, \beta \in \Lambda_k(w)$ such that $\alpha \neq 0$, and at least one of the following inequalities holds:

$$\text{with } e^{\beta} \neq 1, \qquad \cosh(\delta/2) - \mathcal{E}_1(z, \beta) > 0, \tag{74}$$

$$\min \left\{ \cosh(\delta/2) - \mathcal{E}_2(\alpha, \beta, z),\ |E_2(\alpha, \beta, z)| \right\} > 0, \tag{75}$$

$$\text{with } e^{\beta} \neq 1, \qquad \cosh(\delta) - \mathcal{E}_3(\alpha, \beta, z) > 0. \tag{76}$$

Each of these inequalities shows that the (real) distance between some two translates of ℓ is less than δ. The extra conditions ($E_2(\alpha, \beta, z) \neq 0$ in (75) and $e^{\beta} \neq 1$ in (74) and (76)) ensure that these translates are distinct.

For purposes of computation it is convenient to introduce the functions

$$\mathcal{L}_1(k, z, w, \mathbf{n}) = \mathcal{E}_1(z, 2\pi i n_1/k + n_2 w) \qquad (0 \leq n_1 < k),$$

and for $i = 2, 3$,

$$\mathcal{L}_i(k, z, w, \mathbf{n}) = \mathcal{E}_i(z, 2\pi i n_1/k + n_2 w, 2\pi i n_3/k + n_4 w) \quad (0 \leq n_1, n_3 < k),$$

where k is a positive integer, $\mathbf{n} \in \mathbb{Z}^2$ in $\mathcal{L}_1$ and $\mathbf{n} \in \mathbb{Z}^4$ in $\mathcal{L}_2$ and $\mathcal{L}_3$. These are essentially the same functions as the $\mathcal{E}_i$, except that we have now made the dependence explicit on the generators of the lattice $\Lambda_k(w)$ rather than the lattice points themselves. We complete the proof of Lemma 1.1 by showing that at each point of the search space (73) the function $\mathcal{L}_i(k, z, w, \mathbf{n})$ is positive for at least one of the following values of i and $\mathbf{n}$. For $k = 1$,

$$\begin{aligned}
&i = 1 : \mathbf{n} = (0, j),\ (i \leq j \leq 9),\\
&i = 2 : \mathbf{n} = (0, i, 0, j) \text{ for } (i, j) = (1, -3),\ (1, -2),\ (1, 1),\ (1, 2),\ (1, 3)\\
&\qquad\quad (2, -3),\ (2, -1),\ (2, 1),\ (2, 2), (3, -2),\ (3, -1),\ (3, 1),\ (3, 2),\\
&i = 3 : \mathbf{n} = (0, 1, 0, 1),\ (0, 1, 0, 3),\ (0, 2, 0, 2),\ (0, 2, 0, 3),\ (0, 3, 0, 3),
\end{aligned}$$

and for $k = 3$,

$$\begin{aligned}
&i = 1 : \mathbf{n} = (0, 1),\ (2, 1),\ (2,\ 2),\ (1, 0),\ (1, 2),\\
&i = 2 : \mathbf{n} = (0, 1, 1, 0),\ (0, 1, 1, -1),\ (0, 1, 1, -2),\ (0, 1, 2, 0),\ (0, 1, 2, -1),\\
&\qquad\quad (0, 1, 2, -2),\ (1, 0, 1, 2),\ (1, 0, 2, 2),\\
&i = 3 : \mathbf{n} = (0, 1, 0, 1),\ (1, 0, 0, 1),\ (1, 0, 1, 1),\ (1, -1, 0, 1),\ (0, -1, 1, -1),\\
&\qquad\quad (0, -1, 1, -1),\ (2, 1, 2, -1),\ (2, 1, 0, -2),\ (2, 1, 1, -2).
\end{aligned}$$

Appendix 1 gives some more details of this computation.

10. **Volume estimates**

In this section we obtain some lower bounds for volume, in the absence of elliptics of order ≥ 3. We finish the proof of Lemma 1.2.

LEMMA 10.1. *Let Γ be a discrete group with no torsion of order $p \geq 3$, and shortest translation length τ. If $K \subseteq \mathbb{H}^3$ is convex, and if the group H generated by the elliptics in Γ whose axes meet the interior of K*

1. *leaves K invariant;*
2. *contains only the identity and elliptics (that is trivial, cyclic or the Klein 4-group),*

and, if K/H has diameter at most τ, then $\mathrm{Vol}(K/H) \leq \mathrm{Vol}(\mathbb{H}^3/\Gamma)$.

Proof. Let x, y be points of the interior of K that are in the same Γ-orbit. There is a y' in the H-orbit of y that is distance less than τ from x. If $y' \neq x$, then there must be an elliptic $\varphi \in \Gamma$ mapping y' to x and, since K is convex and $\mathrm{ax}(\varphi)$ passes through the midpoint of x and y', $\varphi \in H$. Thus x and y are in the same H-orbit, and the lemma follows. □

LEMMA 10.2. *Let Γ be a discrete group with no torsion of order $p \geq 3$, and shortest translation length τ. Let ℓ be the axis of both an order-2 elliptic φ and a loxodromic with translation length σ, p a point on ℓ midway between two adjacent Klein-4 fixed points (or anywhere on ℓ if there are no such points) and B the solid of revolution of length* $\min\{\sigma/2, \tau\}$ *about ℓ whose radius r_x at displacement x along ℓ from p is given by* $\min\{\tau/2, \operatorname{arccosh}\sqrt{\cosh(\tau)/\cosh(2x)}\}$ $(|x| \leq \min\{\sigma/4, \tau/2\})$. *Then B is convex, its interior meets no elliptic axes other than ℓ, $B/\langle\varphi\rangle$ has diameter at most τ and*

$$\mathrm{vol}(\mathbb{H}^3/\Gamma) \geq \mathrm{vol}(B)/2 = \pi \int_0^a \min\left\{\sinh^2(\tau/2), \frac{\cosh^2(\tau)}{\cosh^2(2x)} - 1\right\} \, \mathrm{d}x, \tag{77}$$

where $a = \min\{\sigma/4,\ \tau/2\}$.

Proof. The convexity of B is easy to prove (consider the analogous region in the plane), and we omit details.

The distance between two adjacent transverse elliptic axes through ℓ is at least $\sigma/2$, and other elliptic axes must be distance at least $\tau/2$ from ℓ. Thus the interior of B meets no elliptic axes other than ℓ. Let Π_x be the intersection of B with the plane that meets ℓ perpendicularly at the point displaced x from p. If $0 \leq x, y \leq \sigma/4$, $u \in \Pi_x$, $v \in \Pi_{-y}$ and d is the distance from u to v in $B/\langle\varphi\rangle$, then by Pythagoras,

$$\cosh(d) \leq \cosh(x+y)\cosh(r_x)\cosh(r_y) \leq \frac{\cosh(x+y)\cosh(\tau)}{\sqrt{\cosh(2x)\cosh(2y)}} \leq \cosh(\tau),$$

using the logarithmic convexity of $\cosh(x)$. The volume estimate (77) then follows from Lemma 10.1. □

Remark. When ℓ has no transverse elliptics, (77) applies with the upper limit for the integral being $a = \tau/2$.

LEMMA 10.3. *An elliptic element of order 2 in a co-compact Kleinian group Γ either shares its axis with a loxodromic of translation length $\geq \lambda_\perp$ or has a collaring radius $\geq \lambda_\perp/2$.*

Proof. Let $\chi \in \Gamma$ be an elliptic of order 2. Since Γ is co-compact, χ has positive collaring radius and shares its axis with some loxodromic f. Let $\tilde{f}$ be the nearest translate of f, and let Γ_1 be defined as in Lemma 6.3. Then $\psi\chi$ is a loxodromic in Γ_1 whose axis is perpendicular to $ax(f)$. By Theorem 1.3, either f or $\psi\chi$ has translation length $\geq \lambda_\perp$, and, in the latter case, $\rho(\mathrm{ax}(f), \mathrm{ax}(\psi))$, which is the collaring radius of f, is at least $\lambda_\perp/2$. □

We now apply these results to small co-volume groups. First, the previous lemma, Theorem 1.2, the fact that $\lambda_\perp/2 > 0.4075$ and Lemma 1.1 immediately give

LEMMA 10.4. *Every order-2 elliptic in a Kleinian group Γ with co-volume $\leq V_0$ shares its axis with a loxodromic of translation length $\geq \lambda_\perp$.*

Proof of Lemma 1.2. The group Γ must have some elliptic of order 2 (otherwise we are in the manifold case and the co-volume is well over V_0) and, by Lemma 10.4, this elliptic must share its axis with a loxodromic of translation length $\geq \lambda_\perp$. Now we apply Lemma 10.2 (with $\sigma = \lambda_\perp$). The estimate (77) (which is clearly an increasing function of both σ and τ) gives co-vol$(\Gamma) > V_0$ if $\tau > 0.497$. The last part of the lemma now follows from Lemma 10.4. □

11. **A collaring estimate**

The following distance formula appeared in [24, Lemma 5.1]. The angle between two disjoint rays is as defined in Section 1.1.

LEMMA 11.1. *Let g be a geodesic in $\mathbb{H}^3$, r_1 and r_2 geodesics perpendicular to g, p_i the point of intersection between r_i and g, a_i a point on r_i at distance x_i from g $(i = 1, 2)$, θ the angle between the two rays emanating from p_i through a_i and ℓ the distance between p_1 and p_2, then*

$$\cosh(\rho(a_1, a_2)) = \cosh(x_1)\cosh(x_2)\cosh(\ell) - \sinh(x_1)\sinh(x_2)\cos(\theta). \tag{78}$$

LEMMA 11.2. *Let Γ be a Kleinian group with no parabolics and no torsion of order $p \geq 3$ and with shortest translation length τ, and suppose that f is a loxodromic with this translation length. If the complex distance between* $\mathrm{ax}(f)$ *and its nearest translate* $\mathrm{ax}(\tilde{f})$ *is* $2(r + i\theta)$ $(0 \leq \theta \leq \pi/2)$, *then*

$$\sinh(r) \geq \min\{\sin(\theta), \cos(\theta)\}\tanh(\tau/2). \tag{79}$$

Proof. Let x and y be the points where $\mathrm{ax}(f)$ and $\mathrm{ax}(\tilde{f})$ respectively meet their common perpendicular. There is an isometry g that maps $\mathrm{ax}(f)$ to $\mathrm{ax}(\tilde{f})$. By composing, if necessary, with a translation along $\mathrm{ax}(\tilde{f})$, we may assume that

g is loxodromic and the distance between $x' := g(x)$ and y is at most τ. Let b be the point on $\mathrm{ax}(\tilde{f})$ midway between x' and y, and set $a = g^{-1}(b)$, whence $\rho(a,x) = \rho(b,x') = \rho(b,y) \le \tau/2$. Using (78),

$$\cosh(\tau) \le \cosh(\rho(a,b)) \le \cosh^2(\tau/2)\cosh(2r) + \sinh^2(\tau/2)|\cos(2\theta)|.$$

(The angle between the rays starting at x through a and starting at y through b may be 2θ or $2\theta+\pi$.) The inequality (79) follows after some further algebraic manipulation. □

A point to note here is that the collar estimate of Lemma 11.2 goes in the other direction to the usual collaring theorems, which typically bound r from below when τ is small.

12. **Proof of Lemma 1.3**

We suppose Γ is a Kleinian group that contradicts Lemma 1.3; that is, Γ has co-volume less than V_0, no torsion of order 3 or more, and the shortest loxodromic in Γ has translation length at most 0.497 and collaring radius at most 0.345.

Let f be this shortest loxodromic, $\tilde{f}$ a nearest conjugate of f (which is chosen to minimize the distance between $\mathrm{ax}(f)$ and $\mathrm{ax}(\tilde{f})$), and let ψ be an order-2 elliptic that interchanges these axes. Let τ and η be respectively the translation length and holonomy of f, and let $r + i\theta$ be the complex distance between $\mathrm{ax}(f)$ and $\mathrm{ax}(\psi)$. Thus

$$0 < \tau \le \tau_{\max} := 0.497, \quad -\pi < \eta \le \pi, \quad 0 < r \le 0.345,$$

and we may assume ψ chosen so that

$$0 \le \theta \le \pi/4.$$

Let $\beta = \beta(f)$, $\gamma = \gamma(f,\psi)$ and $\omega = \sinh^2(r+i\theta)$. By (7) and the fact that $\beta(\psi) = -4$, we have $\gamma = -\beta\omega$. In terms of ω, (10) gives

$$\cosh(2r) = |1+\omega| + |\omega| = (|(2\omega+1)+1| + |(2\omega+1)-1|)/2. \tag{80}$$

It will be convenient to use β and ω as variables in our computations. The bounds above give that $\beta \in R_1 := 2(E_{\tau_{\max}} - 1)$ (as defined at (44)) and $\omega \in R_2 := 0.5(Q_1 - 1)$, where Q_1 is the first quadrant of $E_{0.69}$.

We will show that on the search space $R_1 \times R_2$ at least one of the following is true:

(i) $\frac{\pi}{2}\tau\sinh^2(r) > (0.91)V_0$;
(ii) $|1 - \beta\omega(4-\beta\omega)/4| + |\beta\omega(4-\beta\omega)/4| < |1+\beta/4| + |\beta/4|$ and $\beta\omega \ne 2$;
(iii) $|\beta|(|\omega|+1) < 1$;
(iv) $|1+\beta| + |\beta\omega| < 1$;
(v) $\tau < 0.0979$;

(vi) $\sinh(r) < \min\{\sin(\theta), \cos(\theta)\} \tanh(\tau/2)$;
(vii) $|(\beta+4)\beta|(1+|\omega|) < 1$;
(viii) $|(\beta+3)^2\beta|(1+|\omega|) < 1$;

or, for some good polynomial p, with $\omega' = -p(-\beta\omega, \beta)/\beta$,

(ix) $\omega' \neq -1, 0$ and $|\omega'+1| + |\omega'| < |\omega+1| + |\omega|$;
(x) $\omega' \neq -1, 0$ and $|\beta^2\omega'(\omega'+1)| < 0.198$;
(xi) $\omega' \neq -1, 0$ and $|\beta\omega'(\beta\omega'-4)| + |\beta^2\omega'(\omega'+1)| < 1$;
(xii) $\omega' \neq -1, 0$ and $|\beta^3(\beta+4)\omega'^2(\omega'+1)| < 0.198$;
(xiii) $\omega' \neq -1, 0$ and $|\beta^3(\beta+4)\omega'^2(\omega'+1)| + |\beta\omega'(\beta\omega'-4)| < 1$.

Each of these conditions entails a contradiction.

(i) implies that co-vol$(\Gamma) > V_0$, by Lemma 1.2 and (68) with $k = 1$;

(ii) contradicts the minimality of τ, using (8) and (25) (since $[f, \psi] = f\tilde{f}^{\pm 1} \in \Gamma$). The condition $\beta\omega \neq 2$ insures that $[f, \psi]$ is not an elliptic of order 2;

(iii) and (iv) contradict the discreteness of $\langle f, \psi\rangle$ (hence, by Lemma 6.3, of Γ), using Jørgensen's inequality and Lemma 2.2, respectively;

(v) gives a contradiction, by the result in [7] that $\tau < 0.0979$ gives a collaring radius > 0.345. (This result, which was originally stated in the torsion-free case, still applies when elliptics are present);

(vi) by Lemma 11.2;

(vii) and (viii) by Jørgensen's inequality (applied to the groups $\langle f^2, \psi\rangle$ and $\langle f^3, \psi\rangle$ respectively).

The remaining conditions (ix)–(xiii) all use Theorem 2.1, by which β and $\gamma' = p(\gamma, \beta) = -\beta\omega'$ are parameters of a subgroup $\langle f, g\rangle$ of $\langle f, \psi\rangle$. Since $\omega' \neq -1, 0$, we have $\gamma' \neq 0, \beta$ and so, in particular, this subgroup is nonelementary. Using Lemmas 6.3 and 6.4, we see that (ix) contradicts the minimality of r. The remaining conditions all contradict discreteness; this follows from Lemma 6.1 for (x) and Jørgensen's inequality, Lemma 6.1 and Lemma 2.4, all applied to the group $\langle f, [f, g]\rangle$, for (xi), (xii) and (xiii) respectively, using (25) and (26).

Each of the thirteen conditions above is easily written as an assertion that some function is positive. Our computations in fact show that at each point of the search space, either one of the conditions (i)–(viii) hold, or one of (ix)–(xiii) holds for at least one of the 99 words w of the form (16) where $s_1 = 1$ and $(r_1, r_2, \dots r_{m-1})$ is one of the vectors listed in Appendix 2. We give more details of the calculation in the next section.

13. **Appendix 1: Computations**

We have used machine computation in three proofs (Theorem 1.3, Lemmas 1.1 and 1.3). Recall that, in each case, we have identified a family of functions $\mathcal{Q} = \{q_i \mid i \in I\}$ defined on a region $U \subseteq \mathbb{C}^2$. We must show that at least one of the functions is positive at each point of U.

It will always be possible to find Lipschitz constants for these functions, so we proceed as follows. Define the *box*, $B(z, w, d) \subseteq \mathbb{C}^2$ to be the Cartesian product of the two squares with centers z and w, each with edge length d. We can then find functions $\tilde{q}_i(z, w, d)$ such that

$$\tilde{q}_i(z, w, d) \leq q_i(u, v) \qquad \forall (u, v) \in B(z, w, d).$$

Find a grid $\mathcal{G}$ with mesh size d so that the boxes $B(z, w, d)$ with $(z, w) \in \mathcal{G}$ cover U. Now define a function $\mathrm{Test}(z, w, d)$ recursively as follows. If $B(z, w, d)$ lies wholly out of U, set $\mathrm{Test}(z, w, d) = 1$. If $\max_i\{\tilde{q}_i(z, w, d)\} > 0$, then $\mathrm{Test}(z, w, d)$ takes this value. Otherwise set $\mathrm{Test}(z, w, d)$ to be the minimum of $\mathrm{Test}(z \pm d/4 \pm id/4, w \pm d/4 \pm id/4, d/2)$ (where all four choices of sign are made independently). That is, test each of the sixteen boxes obtained by subdividing the original one. The process continues up to a specified maximum number of subdivisions. We have the required result if $\mathrm{Test}(z, w, d) > 0$ for each $(z, w) \in \mathcal{G}$. As a precaution against roundoff error, we will actually require that $\mathrm{Test}(z, w, d) > \varepsilon = 0.0001$.

For analytic functions we derive Lipschitz bounds using Taylor's approximation. More precisely, let $q(z, w)$ be analytic, $(z, w) \in U$, $d \geq 0$, $m = 0, 1, 2, \ldots$ and $(z', w') \in B(z, w, d)$. We bound the mth term in the Taylor series centered at (z, w) for $q(z', w')$ by

$$q_m(z, w, d) = \frac{1}{m!} \sum_{i=0}^{m} \binom{m}{i} \left| \frac{\partial^m q}{\partial^i z \partial^{m-i} w}(z, w) \right| \left(\frac{d}{\sqrt{2}} \right)^m.$$

Thus, if q is a polynomial and $(z', w') \in B(z, w, d)$, then for any $N \geq 0$,

$$|q(z, w) - q(z', w')| \leq \sum_{m=1}^{N} q_m(z, w, d) + |q_{N+1}|(|z| + d/\sqrt{2}, |w| + d/\sqrt{2}, d), \tag{81}$$

where the bars around the q_{N+1} indicate that the absolute values are taken of each coefficient.

When (as occurs in Lemma 1.1) q is a polynomial in hyperbolic functions, the same bound applies with an obvious modification of the remainder term $|q_{N+1}|$; in this case absolute values are taken of all coefficients, imaginary parts of arguments are dropped and sinh is replaced by cosh throughout.

It is now straightforward to bound all the functions that we use, in the box $B(z, w, d)$. In several places we encounter functions of the form $h(f(z))$, where $h(z) = |z| \pm |1 + z|$. Clearly

$$|h(f(z)) - h(f(z'))| \leq 2|f(z) - f(z')| \leq 2M|z - z'|, \tag{82}$$

where M is a Lipschitz constant for f. The following estimate refines this.

LEMMA 13.1. *For $z \in C$, $d \geq 0$, let $h(z) = |z+1| \pm |z|$ and*

$$\Delta(z,d) = d\sqrt{2\left(1 \pm \frac{|z|^2 + \Re e(z)}{|z||z+1|} + \frac{2d}{|z+1|(|z+1|-d)}\right)}; \tag{83}$$

then

$$|h(z) - h(z')| \leq \Delta(z,d) \quad \textit{when } |z-z'| \leq d. \tag{84}$$

Proof. Let

$$\begin{aligned} w = h(z) &= (|(2z+1)+1| \pm |(2z+1)-1|)/2 \\ &= \cosh(\Re e(\operatorname{arccosh}(2z+1))) && \text{(with + sign)} \\ \text{or} \quad &= \cos(\Im m(\operatorname{arccosh}(2z+1))) && \text{(with − sign).} \end{aligned}$$

We have

$$\frac{\partial}{\partial z} \operatorname{arccosh}(2z+1) = \frac{2}{\sqrt{(2z+1)^2 - 1}},$$

so

$$\begin{aligned} |\mathrm{d}w| &\leq \frac{2|\mathrm{d}z|}{\sqrt{|(2z+1)^2 - 1|}} \sqrt{|1 - h(z)^2|} \\ &= \frac{2|1 - (|z| \pm |z+1|)^2|^{1/2} |\mathrm{d}z|}{\sqrt{|(2z+1)^2 - 1|}} = \sqrt{2\left(1 \pm \frac{|z|^2 + \Re e(z)}{|z||z+1|}\right)} |\mathrm{d}z|. \end{aligned}$$

Since also

$$\left| \frac{|z|^2 + \Re e(z)}{|z||z+1|} - \frac{|z'|^2 + \Re e(z')}{|z'||z'+1|} \right| \leq \frac{2|z-z'|}{|z+1|(|z+1| - |z-z'|)},$$

(84) follows. □

Usually, but not always (for example near $z = -1$), (84) is an improvement on (82); we use whichever is better. In the proof of Lemma 1.3 we use these estimates to find upper and lower bounds for $\cos(2\theta) = |1+\omega| - |\omega|$, $\cosh(2r) = |1+\omega| + |\omega|$ and $\cosh(\tau) = |\beta/4| + |1 + \beta/4|$. In the proof of Lemma 1.1, (82) alone suffices.

It remains to specify the grid points used in the calculation. We will do this in each case by specifying the step size and the extreme values of each of the four (real) variables involved. The unions of the boxes centered at the grid points will be a box that contains the search space. Since in each case we have a search space U with curved boundary being covered by a rectangular grid, some boxes will not lie in U. We therefore initially test each box to determine whether or not it meets U, and we discard it if it does not.

13.1. *Theorem* 1.3. The variables are $\beta_1 = x_1 + iy_1$ and $\beta_2 = x_2 + iy_2$. We use step size $d = 0.1$ and the grid $\mathcal{G}$ with points (x_1, x_2, y_1, y_2), where $-1.95 \le x_1, x_2 \le 0.75$, $0 \le y_1 \le 1.85$, $-1.85 \le y_1 \le 1.85$. We use at most four subdivisions.

13.2. *Lemma* 1.1. The variables are $z = \delta + i\theta$, $w = \tau + i\eta$. We use step size $d = 0.02$. For $k = 1$, we use the grid $\mathcal{G}$ with points $(\delta, \theta, \tau, \eta)$, with $\delta = 0.70(d)0.82$ (that is, δ takes the values $0.70, 0.70 + d, 0.70 + 2d \ldots 0.82$), $\theta = 0.001(d)1.561$, $\tau = 0.002(d)0.182$, $\eta = -3.14(d)3.14$. The boxes with edge length d, centered at these grid points, cover the search space (73). Similarly, for $k = 3$, we use grid points $(\delta, \theta, \tau, \eta)$, with $\delta = 0.598(d)0.818$, $\theta = 0.001(d)1.561$, $\tau = 0.00(d)0.80$, $\eta = -1.04(d)1.04$. The calculation uses at most three subdivisions for $k = 1$, and at most one for $k = 3$.

13.3. *Lemma* 1.3. The variables are $\beta = \tau + i\eta$ and $\omega = x + iy$. For convenience, we divide the domain of β, $R_1 := 2(E_{\tau_{\max}} - 1)$, into six subregions, E_1, E_2, E_3, E_4, E_5 and E_6 corresponding to τ intervals $[-2\cos(\tau_{\max}) - 2, -4]$, $[-4, -3.5]$, $[-3.5, -3]$, $[-3, -2]$, $[-2, -1]$ and $[-1, 0]$. (Jørgensen's inequality alone easily eliminates the piece of R_1 where $\tau > 0$.) We use step size $d = 0.02$ in E_1 and E_3, $d = 0.025$ in E_2, $d = 0.02$ in E_4 and E_5 and $d = 0.1$ in E_6. The grid points are (x, y, τ, η), where x begins at $x = -0.5 + d/2$ and increments in steps of d while $x \le 0.15$; $y = 0(d)0.4$; in the τ interval $[a, b]$, $\tau = a + d/2\ (d)\ b - d/2$, except for E_1, where τ is initialized at -4.27; and $\eta = 0(d)\eta_{\max}$, where $\eta_{\max}$ is 0.48, 0.8, 0.94, 1.05, 1.05 and 1.0 for the regions $E_1 \ldots E_6$, respectively. At most four subdivisions are used in any of these calculations.

14. Appendix 2: Table of polynomials

The words used in the proof of Lemma 1.3 are of the form (16) where $(r_1, r_2, \ldots r_{m-1})$ is one of the 99 vectors in the following list:

$\{(1), (3), (-3,-1), (-3,1), (-3,2), (-3,3), (-2,-1), (-2,1), (-1,-1), (-1,1),$
$(-3,1,-3), (-3,1,-2), (-3,1,1), (-3,2,-3), (-3,2,-1), (-2,-2,1), (-2,-1,-1),$
$(-2,1,-2), (-2,2,-2), (-2,3,-2), (-1,-1,-1), (-1,3,-1), (-3,-1,-3,2),$
$(-3,-1,-1,-3), (-3,1,-2,3), (-3,2,-2,-1), (-3,2,-2,1), (-3,2,-1,2),$
$(-2,-3,1,-2), (-2,-1,-1,-2), (-2,1,-2,-1), (-2,1,-2,1), (-2,1,-2,2),$
$(-2,1,-1,2), (-2,1,1,-2), (-2,1,2,-2), (-2,2,-3,2), (-2,2,-1,-1),$
$(-2,2,2,-2), (-2,3,-1,1), (-1,-1,-1,-1), (-1,1,-3,1), (-1,1,-2,1),$
$(-3,-1,-2,-1,-3), (-3,-1,-1,1,-2), (-3,-1,1,-2,1), (-3,1,-2,-3,1),$
$(-3,1,-1,-1,-1), (-3,1,-1,-1,2), (-3,1,1,-2,1), (-3,2,-3,1,1),$
$(-3,2,-1,3,-2), (-2,-1,-2,-1,-2), (-2,-1,-1,-2,-1), (-2,-1,-1,-1,1),$
$(-2,-1,1,-2,2), (-2,-1,1,-1,-2), (-2,-1,1,1,1), (-2,1,-3,1,-2),$

$(-2,1,-2,-1,-1),(-2,1,-2,1,-2),(-2,1,1,-1,1),(-2,1,1,1,-1),$
$(-2,2,-1,2,-2),(-2,2,-1,2,1),(-2,3,-2,2,-1),(-1,-2,2,-2,-1),$
$(-1,-1,-2,1,-1),(-1,-1,1,-1,-1),(-1,-1,2,-1,-1),(-1,2,-2,2,-1),$
$(-2,-1,-1,-1,-1,1),(-2,-1,-1,1,-1,-1),(-2,-1,1,-2,2,-2),(-2,-1,1,-2,2,-1),$
$(-2,1,-1,-1,1,-2),(-2,2,-2,-1,-1,1),(-2,2,-1,-1,-2,2),(-2,2,-1,-1,2,-2),$
$(-2,2,-1,1,2,-1),(-1,-2,-1,-1,-2,-1),(-1,-2,1,1,1,-1),$
$(-1,-2,2,-1,1,-1),(-1,-2,2,-1,1,1),(-1,-1,-2,-2,-1,-1),$
$(-1,-1,-1,-1,1,-1),(-1,-1,-1,1,-1,-1)(-1,-1,-1,2,-2,1),,$
$(-1,-1,1,-2,1,-1),(-1,1,-2,-2,1,-1),(-1,1,1,1,1,-1),$
$(-1,1,2,-2,2,-1),(-1,-1,-1,1,-1,-1,-1),(-1,-1,1,-1,-1,-1,1),$
$(-1,-1,1,1,1,1,-1),(-1,1,1,1,1,1,-1),(-1,-1,-1,1,-1,-1,-1,1),$
$(-1,-1,-1,1,1,-1,-1,-1),(-1,-1,1,-1,-1,-1,1,1)\}.$

Not all of these words are used in all the calculations. Below are lists for $1 \leq i \leq 6$ of the positions in the above vector of the words used for the calculation with $\beta \in E_i$:

$E_1:\ \{4,8,9,11,12,17,18,21,24,30,31,32,33,47,53,54,60,68,85,98\},$

$E_2:\ \{1,2,3,4,5,8,9,11,12,15,16,17,18,21,22,24,26,28,30,32,33,$
$35,41,42,43,37,49,55,56,57,64,72,74,80,83,86,87,88,90,96\},$

$E_3:\ \{1,2,4,6,8,9,12,14,15,19,21,23,25,26,27,29,32,35,36,37,39,$
$41,42,44,46,48,50,52,57,59,64,66,67,75,77,81,84,89,91,92,95\},$

$E_4:\ \{1,2,3,4,5,6,7,9,10,11,12,13,14,15,17,18,19,20,21,26,27,30,$
$32,34,35,38,39,40,41,42,43,45,46,48,50,55,57,59,61,63,64,65,66,$
$67,69,70,71,73,75,76,79,82,86,87,89,91,93,94,95,96,97,99\},$

$E_5:\ \{1,9,20,21,26,51,62,63,69,70,78,94,99\},$

$E_6:\ \{1,7,9,58,62\}.$

References

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

[2] A. Borel, Commensurability classes and volumes of hyperbolic 3-manifolds, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **8** (1981), 1–33. MR 0616899. Zbl 0473.57003. Available at http://www.numdam.org/item?id=ASNSP_1981_4_8_1_1_0.

[3] C. Cao, On three generator Möbius groups, *New Zealand J. Math.* **23** (1994), 111–120. MR 1313446. Zbl 0839.30048.

[4] L. Carleson and T. W. Gamelin, *Complex Dynamics*, *Universitext: Tracts in Math.*, Springer-Verlag, New York, 1993. MR 1230383. Zbl 0782.30022.

[5] M. Conder, G. Martin, and A. Torstensson, Maximal symmetry groups of hyperbolic 3-manifolds, *New Zealand J. Math.* **35** (2006), 37–62. MR 2222175. Zbl 1104.20035.

[6] D. A. Derevnin and A. D. Mednykh, Geometric properties of discrete groups acting with fixed points in a Lobachevskiĭ space, *Dokl. Akad. Nauk SSSR* **300** (1988), 27–30. MR 0948799. Zbl 0713.30044.

[7] D. Gabai, G. R. Meyerhoff, and P. Milley, Volumes of tubes in hyperbolic 3-manifolds, *J. Differential Geom.* **57** (2001), 23–46. MR 1871490. Zbl 1029.57014. Available at http://projecteuclid.org/euclid.jdg/1090348088.

[8] ———, Minimum volume cusped hyperbolic three-manifolds, *J. Amer. Math. Soc.* **22** (2009), 1157–1215. MR 2525782. Zbl 1204.57013. http://dx.doi.org/10.1090/S0894-0347-09-00639-0.

[9] D. Gabai, G. R. Meyerhoff, and N. Thurston, Homotopy hyperbolic 3-manifolds are hyperbolic, *Ann. of Math.* **157** (2003), 335–431. MR 1973051. Zbl 1052.57019. http://dx.doi.org/10.4007/annals.2003.157.335.

[10] F. W. Gehring, C. Maclachlan, G. J. Martin, and A. W. Reid, Arithmeticity, discreteness and volume, *Trans. Amer. Math. Soc.* **349** (1997), 3611–3643. MR 1433117. Zbl 0889.30031. http://dx.doi.org/10.1090/S0002-9947-97-01989-2.

[11] F. W. Gehring and G. J. Martin, Some universal constraints for discrete Möbius groups, in *Paul Halmos*, Springer-Verlag, New York, 1991, pp. 205–220. MR 1113275. Zbl 0784.30038.

[12] ———, Axial distances in discrete Möbius groups, *Proc. Nat. Acad. Sci. U.S.A.* **89** (1992), 1999–2001. MR 1154948. Zbl 0753.30030. http://dx.doi.org/10.1073/pnas.89.6.1999.

[13] ———, 6-torsion and hyperbolic volume, *Proc. Amer. Math. Soc.* **117** (1993), 727–735. MR 1116260. Zbl 0790.30032. http://dx.doi.org/10.2307/2159135.

[14] ———, Commutators, collars and the geometry of Möbius groups, *J. Anal. Math.* **63** (1994), 175–219. MR 1269219. Zbl 0799.30034. http://dx.doi.org/10.1007/BF03008423.

[15] ———, On the minimal volume hyperbolic 3-orbifold, *Math. Res. Lett.* **1** (1994), 107–114. MR 1258496. Zbl 0834.30029.

[16] ———, On the Margulis constant for Kleinian groups. I, *Ann. Acad. Sci. Fenn. Math.* **21** (1996), 439–462. MR 1404096. Zbl 0854.30031. Available at http://www.acadsci.fi/mathematica/Vol21/gehring.html.

[17] ———, Precisely invariant collars and the volume of hyperbolic 3-folds, *J. Differential Geom.* **49** (1998), 411–435. MR 1669657. Zbl 0989.57010. Available at http://projecteuclid.org/euclid.jdg/1214461106.

[18] ———, Minimal co-volume hyperbolic lattices. I. The spherical points of a Kleinian group, *Ann. of Math.* **170** (2009), 123–161. MR 2521113. Zbl 1171.30014. http://dx.doi.org/10.4007/annals.2009.170.123.

[19] A. Hurwitz, Ueber algebraische Gebilde mit eindeutigen Transformationen in sich, *Math. Ann.* **41** (1892), 403–442. MR 1510753. JFM 24.0380.02. http://dx.doi.org/10.1007/BF01443420.

[20] T. Jørgensen, On discrete groups of Möbius transformations, *Amer. J. Math.* **98** (1976), 739–749. MR 0427627. Zbl 0336.30007. http://dx.doi.org/10.2307/2373814.

[21] ______, Commutators in $\mathrm{SL}(2, \mathbf{C})$, in *Riemann Surfaces and Related Topics: Proceedings of the* 1978 *Stony Brook Conference* (State Univ. New York, Stony Brook, N.Y., 1978), *Ann. of Math. Stud.* **97**, Princeton Univ. Press, Princeton, N.J., 1981, pp. 301–303. MR 0624821. Zbl 0459.20044.

[22] A. M. Macbeath, On a theorem of Hurwitz, *Proc. Glasgow Math. Assoc.* **5** (1961), 90–96 (1961). MR 0146724. Zbl 0134.16603.

[23] T. H. Marshall and G. J. Martin, Volumes of hyperbolic 3-manifolds. Notes on a paper of D. Gabai, G. Meyerhoff, and P. Milley, "Volumes of tubes in hyperbolic 3-manifolds" [J. Differential Geom. **57** (2001), no. 1, 23–46; MR1871490], *Conform. Geom. Dyn.* **7** (2003), 34–48. MR 1992036. Zbl 1065.57017. http://dx.doi.org/10.1090/S1088-4173-03-00081-X.

[24] ______, Cylinder and horoball packing in hyperbolic space, *Ann. Acad. Sci. Fenn. Math.* **30** (2005), 3–48. MR 2140297. Zbl 1079.52011. Available at http://www.acadsci.fi/mathematica/Vol30/martin.html.

[25] R. Meyerhoff, Sphere-packing and volume in hyperbolic 3-space, *Comment. Math. Helv.* **61** (1986), 271–278. MR 0856090. Zbl 0611.57010. http://dx.doi.org/10.1007/BF02621915.

[26] G. D. Mostow, Quasi-conformal mappings in n-space and the rigidity of hyperbolic space forms, *Inst. Hautes Études Sci. Publ. Math.* **34** (1968), 53–104. MR 0236383. Zbl 0189.09402. http://dx.doi.org/10.1007/BF02684590.

[27] A. Przeworski, A universal upper bound on density of tube packings in hyperbolic space, *J. Differential Geom.* **72** (2006), 113–127. MR 2215457. Zbl 1098.52005. Available at http://projecteuclid.org/euclid.jdg/1143593127.

[28] D. Rolfsen, *Knots and Links*, *Math. Lecture Series* **7**, Publish or Perish, Berkeley, CA, 1976. MR 0515288. Zbl 0339.55004.

[29] A. Selberg, On discontinuous groups in higher-dimensional symmetric spaces, in *Contributions to Function Theory* (Internat. Colloq. Function Theory, Bombay, 1960), Tata Institute of Fundamental Research, Bombay, 1960, pp. 147–164. MR 0130324. Zbl 0201.36603.

[30] C. L. Siegel, Discontinuous groups, *Ann. of Math.* **44** (1943), 674–689. MR 0009959. Zbl 0061.04504. http://dx.doi.org/10.2307/1969104.

(Received: September 10, 2007)
(Revised: January 16, 2012)

American University of Sharjah, Sharjah, UAE
E-mail: tmarshall@aus.edu

Massey University, Auckland, New Zealand
E-mail: G.J.Martin@massey.ac.nz