Enumerating Kleinian Groups

David Gabai, Robert Meyerhoff, Nathaniel Thurston, and Andrew Yarmola

Abstract. In this manuscript, we give an overview of the tools and techniques needed for successfully classifying "low-complexity" Kleinian groups. In particular, we focus on extracting topological and geometric properties of discrete Kleinian groups, such as bounds on tube radii, cusp geometry, volume, relators in group presentation, and similar quantities. These methods provide an effective avenue for studying and classifying hyperbolic 3-manifolds that satisfy some geometric or topological constraints.

Contents

1. Introduction

The geometrization theorem shows that closed oriented irreducible 3-manifolds are made up of pieces that each admit a geometric structure arising from one of eight Lie groups. The holonomy of these structures gives rise to lattices in the corresponding Lie groups. Of the eight types, Kleinian groups – corresponding for hyperbolic 3-manifolds – have garnered a great deal of interest. At the same time, computational methods have played an important role in low-dimensional topology in the past several decades by providing both examples and computationally-assisted verified proofs. Our main goal with this manuscript is to provide an avenue, via examples and methods, for finding geometric constraints on discrete groups and extracting topological, algebraic, and geometric information that may be used to prove further results. As detailed in Section 2, these methods have had a number of notable successes and we believe that they have great potential for future development. Our examples here will be limited to Kleinian groups, that is discrete subgroups of $\mathrm{PSL}(2,\mathbb{C})$, as studied in [**21, 19**], but we expect these methods to have broader application, especially in low-rank where Margulis Super-Rigidity and Arithmeticity do not apply.

Let G be a Lie group and $\Gamma \leq G$ a discrete group. One basic approach to study Γ is to focus on finitely generated subgroups that have specific properties. For example, if one is interested in finding maximal tubes about a short geodesic in a hyperbolic manifold, the subgroup of interest would be generated by an element representing a shortest geodesic and a conjugating element that corresponds to a self-tangency of the tube. A similar approach can be taken for maximal cusps, or for analyzing Margulis numbers. Most of our examples focus 2 or 3 generator subgroups as their parameter spaces tend to be small enough for computational tools to work, yet too large for by-hand analysis. The goal of the methods we describe here is to classify which part of the parameter space contains all discrete groups and to extract useful information about such groups. This information includes provable relators in the group, bounds on the generators, and other geometric and algebraic output. In Section 2, we give examples of results that can be directly extracted for such searches and some relevant historical background.

The basic idea is to eliminate places in the parameter space where some geometric measurement violates discreteness or other assumptions. Starting with Section 4, the text is organized as follows. We first introduce the notion of *marking* and give examples of how to choose one. Next, we discuss how to setup the parameter space itself. The hard work begins with trying to understand how to cut away "bad" portions of the parameter space, that is, remove places where discrete groups cannot exist (or some other assumptions about the marking fail). Here, words in the generators play a key role. Outside of helping one eliminate parameters, words can be used to find guaranteed relators for the discrete groups whose parameters appear in the parameter space. We explain this dichotomy when discussing *killer*, *quasi-relator* and *variety* words. In Section 7, we get down to the details of how to computationally partition the parameter space into boxes in an organized way as to make sure that the problem of finding the aforementioned words has a chance of being feasible. In Section 8, we discuss how to build word-finding algorithms and their pitfalls. Finally, in Section 9, we give an overview of our preferred choice of arithmetic and in Section 10, we discuss methods for sanity checking progress.

REMARKS 1.1. It is worth pointing out that most of this paper deals with "large" parameter spaces. In particular, we aren't concerned with studying a specific manifold most of the time. However, it is often possible to get information for specific manifolds as a result. For example, we can use relators found for the Whitehead link to show the maximal cusp volume for that specific manifold decreases under Dehn filling of one of the cusps. Similarly, it should be possible to compute more specific rigorous information for hyperbolic manifolds, such as length spectra, upper and lower bounds on volume, and canonical triangulations.

It is important to note that for studying specific manifolds, results along these lines are spectacularly provided by SnapPea [**40**] and its descendants such as Snap [**24, 13**] and SnapPy [**15**]. Its current incarnation SnapPy is indispensable to workers in the field and can produce verified results for several computations, such as volume and isometry validation.

2. Applications and historical context

In the context of closed hyperbolic 3-manifolds, the existence of sufficiently thick tubes about short geodesics is crucial to many results in the subject. One such result is the $\log(3)/2$-*theorem* of [**21**], which was proven using techniques discussed in this manuscript. The sharp form of that result **[23]** asserts that with the exception of the manifold known as Vol3, any closed hyperbolic 3-manifold X has a $\log(3)/2$ tube about some geodesic; furthermore, with five exceptions this geodesic can be taken to be any shortest one. Results that crucially use this work include the Smale conjecture for hyperbolic 3-manifolds **[18]** and the important inequality (slightly updated to reflect **[23]**), that if $X \neq$ Vol3 is a closed hyperbolic 3-manifold, then it is obtained by filling a 1-cusped hyperbolic 3-manifold Y such that $\mathrm{vol}(Y) < 3.0177\,\mathrm{vol}(X)$ **[6]**. This inequality is based on earlier works of Agol **[4]** and Agol-Dunfield **[7]** and uses Perelman's work on Ricci flow **[35]**, **[36]**. It is needed to show that the Weeks manifold is the unique closed hyperbolic 3-manifold of least volume, **[22]**.

Recall that in the 1960's Gregory Margulis proved that each end of a complete finite-volume hyperbolic 3-manifold Y is homeomorphic to $T^2 \times (1, \infty)$. These ends are called *cusps*. Note, the term cusp in this paper always refers to a *rank-two* cusp, even if the manifold in question is not finite-volume. Each cusp contains a properly embedded *horocusp*, which is a region isometric to a quotient of H_∞, where $H_\infty = \{(x, y, t) : t > 1\} \subset \mathbb{H}^3$ in the upper half-space model and we quotient by a group of translations of (x, y) isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$, see **[38]**. Further, after isometrically changing coordinates, the end can be maximally enlarged to the *maximal horocusp* κ. Here, κ is embedded in Y and $\partial\kappa$ has finitely many self-tangencies. A basic fact is that $\mathrm{vol}(\kappa) = \mathrm{area}(\partial\kappa)/2$, called the *cusp volume*.

The works of Jørgensen and Thurston from the 1970's demonstrate the close relation between thick tubes about geodesics and maximal horocusps of complete manifolds **[37]**. Indeed, Jørgensen showed that after passing to a subsequence, any infinite sequence Y_i of distinct complete hyperbolic 3-manifolds of uniformly bounded volume limits geometrically to a complete manifold Y_∞, where thicker and thicker tubes of the Y_i's limit to cusps of Y_∞. Conversely, Thurston showed that given a cusp of a complete finite-volume 3-manifold Y and $\epsilon > 0$, then all but finitely many fillings on that cusp produce hyperbolic manifolds geometrically ϵ-close to Y. The papers, **[34]**, **[28]**, **[29]**,**[17]** have made this connection more explicit and quantitative.

Let κ be a maximal horocusp of the complete hyperbolic 3-manifold $Y = \mathbb{H}^3/\Gamma$ and normalized so that H_∞ and H_0 both cover κ and are conjugate under Γ, where H_0 is a horoball centered at 0 and tangent to H_∞. Let $\mathfrak{B} = \langle m, n, g\rangle \leq \Gamma$, where m, n are generators of $\{\gamma \in \Gamma : \gamma \cdot H_\infty = H_\infty\}$ and g is the conjugating element $g(H_\infty) = H_0$. By studying the parameter space of such 3-generator groups, the authors in **[19]** prove the following two results.

Theorem 2.1. *The figure-8 knot complement and its sister are the complete hyperbolic 3-manifolds with a maximal horocusp of minimal volume. This volume is* $\sqrt{3}$.

Remarks 2.1.

(i) These manifolds minimize cusp volume among all complete hyperbolic 3-manifold with torus cusps, even among multi-cusped manifolds and complete manifolds of infinite volume.
(ii) Results on this question go back to the 1980's. A lower bound of $\sqrt{3}/4$ for the volume of a maximal cusp was obtained in **[32]**. Adams improved this bound by a factor of 2 in **[2]**. There were no further improvements until Cao and Meyerhoff **[12]** gave a bound of 3.35/2 for the volume of a maximal cusp in a hyperbolic 3-manifold, which is within 0.0571 of the actual value of $\sqrt{3}$.

Theorem 2.2. *If Y is a complete hyperbolic 3-manifold with a maximal horocusp of volume ≤ 2.62, then $\pi_1(Y) = \mathfrak{B} = \langle m, n, g\rangle$ and $\mathrm{vol}(Y) < \infty$. Furthermore, there exists a reduced relation $w(m, n, g)$ where g and g^{-1} appear at most 7 and at least 4 times in total.*

This gives an effective analogue of a theorem of Agol **[5]**, which shows that some relator always exists whenever $\text{vol}(\kappa) < \pi$.

By applying Theorem 2.2, the authors of **[19]** are able to give finite list of manifolds whose Dehn fillings give all manifolds with maximal cusp volume at most 2.62. This also shows that all such manifolds are tunnel number one. Enumerating small volume manifolds is also possible using this result. By a packing argument (see **[33]**), if Y has a maximal horocusp of volume V, then $\text{vol}(Y) \geq (2v_3/\sqrt{3})V \approx (1.17195...)V$, where v_3 is the volume of the regular ideal tetrahedron in $\mathbb{H}^3$. It follows that either $\text{vol}(Y) \geq 2.62 \cdot (1.17195...) \approx 3.0705...$ or Y is obtained by filling one of the manifolds listed in **[19]**. This further allows for the classification the three smallest volume hyperbolic 3-manifolds.

In his revolutionary work in the 1970's on the geometry of the figure-8 knot complement, Bill Thurston showed that it has exactly 10 non-hyperbolic Dehn fillings. Thurston's result spawned a tremendous amount of work towards understanding hyperbolic and non-hyperbolic fillings of hyperbolic 3-manifolds, e.g. see the surveys in **[25]**, **[10]**, **[30]**. In particular, there are infinitely many 1-cusped hyperbolic manifolds with six non-hyperbolic fillings. By 1998, there were (resp. eight, two, zero, one)) known 1-cusped manifolds with (resp. seven, eight, nine, ten) exceptional surgeries. This led Cameron Gordon to conjecture in **[25]** that if M be a hyperbolic 3-manifold with boundary a torus, then M has at most 8 non-hyperbolic fillings unless M is the figure eight knot exterior.

The Gromov-Thurston 2π-theorem showed that if a manifold Y' is obtained by filling the cusped manifold Y along a curve $\gamma \subset \partial\kappa$ of length $> 2\pi$ then Y' has a complete metric of negative sectional curvature (hence a hyperbolic structure by Perelman). The shortest curve on $\partial\kappa$ has length ≥ 1 and Thurston used this fact to show that the number of non-hyperbolic fillings is bounded by 48. But coupling a length lower bound with a maximal horocusp area lower bound can lead to an improvement, and this was carried out by Bleiler and Hodgson ([BH92]) — using Adams's area lower bound — to get a bound of 24 on the number of non-hyperbolic fillings. Had the Cao-Meyerhoff area bound been available at that time, the non-hyperbolic-filling bound would have been reduced to 14.

In 2000, Agol and Lackenby independently improved 2π to 6 with the weaker conclusion that Y' has non-elementary and word-hyperbolic fundamental group (again Perelman implies that Y is hyperbolic). Coupling the 6-theorem with the Cao-Meyerhoff bound led to a non-hyperbolic filling bound of at most 12 tantalizingly close to the figure-eight case bound of 10. As the 6-theorem is sharp, attention was focused on area bound improvements. However, a new area bound of 3.7 in the "non-Mom" case in **[22]** did not improve upon the non-hyperbolic filling bound of 12. The **[22]** area result seemed quite strong at the time, so potential for progress on the non-hyperbolic filling bound seemed bleak. But Lackenby and Meyerhoff were able to prove in **[30]** the 10 bound without using the area bound improvements by generalizing the 6-theorem and exploiting Mom technology.

In **[3]** and **[5]**, Agol proved that if Y is a 1-cusped hyperbolic 3-manifold with 9 or more non-hyperbolic fillings, then the maximal cusp of Y has volume $> 18/7 = 2.57....$ Since $18/7 < 2.62$ all the 1-cusped hyperbolic 3-manifolds with a maximal cusp of volume $\leq 18/7$ are Dehn filling of those given in **[19]**, which after another Dehn filling argument, also given by Tom Crawford **[14]**, proves Gordon's conjecture.

3. Background

Our examples will focus on hyperbolic 3-manifolds. We will use the upper-half-space model $\mathbb{H}^3 = \{(z,t) \in \mathbb{C} \times \mathbb{R} : t > 0\}$ of hyperbolic 3-space. Consider $q \in \mathbb{H}^3$ in quaternion notation $q = x + \mathfrak{i}y + \mathfrak{j}t$, where $\mathfrak{i}^2 = \mathfrak{j}^2 = \mathfrak{k}^2 = \mathfrak{i}\mathfrak{j}\mathfrak{k} = -1$. Identifying $\text{Isom}^+(\mathbb{H}^3) \cong \text{PSL}(2,\mathbb{C})$, the action on $q = x + \mathfrak{i}y + \mathfrak{j}t$ is given by

$$\pm\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot q = (aq+b)(cq+d)^{-1}.$$

Notice that for $z \in \mathbb{C}$, we have that $z\,\mathfrak{j} = \mathfrak{j}\,\overline{z}$, so one computes

$$\pm\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot t\,\mathfrak{j} = (a\,t\,\mathfrak{j}+b)(c\,t\,\mathfrak{j}+d)^{-1} = \frac{t^2\,a\,\overline{c}+b\,\overline{d}}{t^2|c|^2+|d|^2} + \frac{t\,\mathfrak{j}}{t^2|c|^2+|d|^2}. \tag{3.1}$$

In particular, $\mathfrak{j}$ is mapped to a point of Euclidean height $1/(|c|^2+|d|^2)$, which we will use later.

The group $\mathrm{Isom}^+(\mathbb{H}^3)$ acts transitively on $\mathbb{H}^3$ with stabilizer $\mathrm{SO}(3)$. On the boundary at infinity, thought of as the Riemann sphere $\hat{\mathbb{C}} = \partial_\infty\mathbb{H}^3$, the action is transitive on triples of points. Elements of $\mathrm{Isom}^+(\mathbb{H}^3)$ fall into three classes based on their fixed points in $\overline{\mathbb{H}}^3$: *loxodromic* elements fix two points on $\hat{\mathbb{C}}$, *parabolic* elements fix one point on $\hat{\mathbb{C}}$, and *elliptic* elements fix points on the interior $\mathbb{H}^3$. Discrete groups $\Gamma \leq \mathrm{Isom}^+(\mathbb{H}^3) \cong \mathrm{PSL}(2,\mathbb{C})$ are called *Kleinian groups*. If they are torsion-free, then they contain no elliptics. Any subgroup of Γ made up of only parabolics has rank at most 2. Further, if $\mathbb{H}^3/\Gamma$ is compact, then Γ contains no parabolics. The action of a loxodromic preserves a complete geodesic in $\mathbb{H}^3$, called its axis. This action about the axis is a composition of a translation and a rotation, giving loxodromics the notion of complex lengths, where the real part is the translation and the imaginary is the rotation. For a parabolic, if one conjugates its fixed point to $\infty \in \hat{\mathbb{C}}$, its action on $\hat{\mathbb{C}}$ looks like $z \mapsto z+c$ for $c \in \mathbb{C}$. On the interior, this action preserves *horoballs at infinity*, which are sets of the form $\mathbb{H}_\infty = \{(z,t) : t > a\}$ for some real $a \in \mathbb{R}^+$. The image of H_∞ under an element not fixing ∞ looks like a Euclidean ball tangent to $\mathbb{C}$ at a point, called its *center*.

Geodesics of $Y = \mathbb{H}^3/\Gamma$ correspond to conjugacy classes in Γ, and *tubes* about geodesics in Y corresponds to regular neighborhoods of axes of loxodromics in $\mathbb{H}^3$, also often called tubes. A maximal tube of Y lifts to a collection of tangent tubes upstairs. We often only focus on one tangency at a time and this corresponds to an element $w \in \pi_1(Y) = \Gamma$ that maps one tube in a tangent pair to the other. When Y has finite volume, its ends are of the form $T^2 \times (1,\infty)$ and contain embedded neighborhoods isometric to H_∞/Π, where Π is a group of parabolics fixing infinity and isomorphic to $\mathbb{Z}\oplus\mathbb{Z}$. Such neighborhoods are called *horocusps*. Just like for tubes, tangencies of maximal horocusps correspond to elements of $\pi_1(Y)$.

4. Markings

In this section, we introduce the notion of a *marking*, which plays a central role in our techniques for analyzing parameter spaces of subgroups of $\mathrm{PSL}(2,\mathbb{C})$. We will focus on the mathematical setup and examples before discussing the computational aspects in later sections. The general ideal is to look at groups generated by a *marked* set of generators. Such markings, defined below, will then allow us to (1) parametrize all possible generating sets and (2) efficiently eliminate ranges of generating sets from consideration. This will leave us with some (possibly large) collection of pieces of the original parameter space that includes all (correctly marked) discrete groups of interest, allowing us to finally extract topological, geometric, combinatorial and algebraic information.

4.1. Markings and examples.

Fix a Lie group G. For us, this will almost always be $\mathrm{PSL}(2,\mathbb{C})$.

Definition 4.1. *A* marking *in a Lie group G is the tuple $(S, U, \mathfrak{Dicta})$, where $S = \{g_1,\dots,g_k\} \subset G$ is an ordered set, U is a subset of a homogenous space of G, and $\mathfrak{Dicta}$ is a collection of desirable properties expressed in terms of S and U. We will let Γ_S denote the (possibly indiscrete) group $\langle g_1,\dots,g_l\rangle$. $\mathfrak{Dicta}$ should be chosen such that if all properties hold for Γ_S, then it is a group of interest. Further, if some property in $\mathfrak{Dicta}$* fails *for a particular Γ_S, it should be possible to verify this in finite time.*

Examples. While the definition of marking is rather informal, we hope the following examples will make this notion relatively clear. Specifically, $\mathfrak{Dicta}$ must be carefully designed to work in the relevant context and be precise enough to pinpoint the groups of interest.

Tubes. Suppose we would like to find the largest radius of a tube that is guaranteed to be embedded around a shortest geodesic in any closed hyperbolic 3-manifold Y. Let γ be a shortest geodesic in Y and consider a maximal embedded (open) tube R around γ, so by maximality its boundary has a self-tangency. In the universal cover $\mathbb{H}^3$ of Y, this picture corresponds to a tube U around a complete geodesic (i.e lifts of R and γ, respectively) and a tube tangent to U of the form $w \cdot U$ for some $w \in \pi_1(Y)$. From this, we build our marking as follows. Take $G = \mathrm{PSL}(2, \mathbb{C})$, $S = \{f, w\}$, and U a tube around a complete geodesic in $\mathbb{H}^3$. We will want to think of f as corresponding to our shortest geodesic γ in Y. So, we take $\mathfrak{Dicta}$ to be the following set of assumptions:

(1) f, w are non-commuting loxodromics,
(2) the core geodesic of U is the axis of f,
(3) U and $w \cdot U$ are tangent,
(4) f has smallest real translation length amongst $\Gamma_S \setminus \{id\}$,
(5) Γ_S is torsion-free, and
(6) for any $h \in \Gamma_S \setminus \langle f \rangle$, the interiors of U and $h \cdot U$ are disjoint.

Conditions (5) and (6) imply that Γ_S is a torsion-free Kleinian group and $R = U/\Gamma_S$ is our maximal embedded tube around γ in $Y = \mathbb{H}^3/\Gamma_S$. Note, conditions such as (1)-(3) can often be encoded in the choice of f, w and U when we consider setting up a parameter space of markings. This marking was used in **[21]** to prove that outside of a few exceptions, the tube radius can be taken to be at least $\log(3)/2$.

Cusps. Similarly, suppose you want to study maximal cusps neighborhood and their geometry in a finite-volume hyperbolic 3-manifold Y. Any such maximal cusp neighborhood has a horoball lift in $\mathbb{H}^3$. This lift is then preserved by some pair of parabolics in $\pi_1(Y)$ and, by maximality, is moved to a tangent horoball by some element of in $\pi_1(Y)$. Taking this into account, we let $S = \{m, n, g\}$ and U a horoball in $\mathbb{H}^3$. Then $\mathfrak{Dicta}$ are the conditions:

(1) m, n are commuting non-parallel parabolics,
(2) m, n preserve U,
(3) m has shortest translation length on ∂U amongst $\langle m, n \rangle \setminus \{id\}$,
(4) $g \cdot U$ is tangent to U,
(5) Γ_S is torsion-free, and
(6) the interiors $w \cdot U$ and U are disjoint for all $w \in \Gamma_S \setminus \langle m, n \rangle$.

Notice that if any of these conditions fail, they can be computationally verified to fail. This marking was used in **[19]** to classify the infinite family of cusped manifolds with maximal cusp volume less than or equal to 2.62 and showing that the figure eight knot and its sister minimize maximal cusp volume. Forthcoming work **[20]** will give improved bounds on the distance between exceptional slopes in 1-cusped hyperbolic manifolds.

Margulis. Recall that the Margulis number of a Kleinian group Γ is the number

$$\mu_\Gamma = \sup\{d \in \mathbb{R}^+ : \text{ if } d_{\mathbb{H}^3}(p, x_i \cdot p) < d \text{ for some } x_1, x_2 \in \Gamma \text{ and } p \in \mathbb{H}^3, \text{ then } x_1, x_2 \text{ commute}\}$$

Say we want to find the smallest Margulis number amongst hyperbolic 3-manifolds, called the Margulis constant. To accomplish this, we take $S = \{x_1, x_2\}$, $U = p$ is a point, and $\mathfrak{Dicta}$ that assumes:

(1) x_1, x_2 do not commute,
(2) if $d_{\mathbb{H}^3}(p, w \cdot p) < \max_i d_{\mathbb{H}^3}(p, x_i \cdot p)$ for some $w \in \Gamma_S$, then w commutes with x_1 or x_2,

(3) $\max_i d_{\mathbb{H}^3}(q, w_i \cdot q) \leq \max_i d_{\mathbb{H}^3}(p, x_i \cdot p)$ for any non-commuting pair $w_1, w_2 \in \Gamma_S$ and $q \in \mathbb{H}^3$,
(4) Γ_S is torsion-free, and
(5) x_1 has real translation length at most that of x_2.

When Γ_S is discrete, (2) implies that $\mu_{\Gamma_S} \leq \max_i d_{\mathbb{H}^3}(p, x_i \cdot p)$. Together with (3), we get that $\mu_{\Gamma_S} = \max_i d_{\mathbb{H}^3}(p, x_i \cdot p)$, so x_1, x_2 and p realize the Margulis number of Γ_S. It is interesting to note that (3) may be dropped depending on the context. If the goal is to find the Margulis constant, removing (3) would result in a larger family of markings, but any marking where μ_{Γ_S} is the Margulis constant would still be present. We also note that (5) allows us to reduce the size of the parameter space of markings. Work on this parameter space with the goal of showing that the Weeks manifolds uniquely realizes the Margulis constant is currently in progress [**16**].

Sys. For hyperbolic surfaces, one may want to find the maximal injectivity radius of a surface with fixed topology. Here, one would take U to be empty and construct S in some systematic way from parameters on moduli space (e.g. Fenchel-Nielsen coordinates for some fixed pants decompositions). Then $\mathfrak{Dicta}$ simply assumes that the injectivity radius of Γ_S is at most some known number. As one finds coordinates for which Γ_S has larger and larger injectivity radii, the space of possible S can be further restricted until it is small enough to analyze by hand. As we shall see later, our approach and proof technique does indeed allow for such iterative methods, though this particular one is in early stages of implementation [**39**] unlike previous examples.

4.2. Geometric and incorrect markings. We call a marking *geometric* if it arises from a hyperbolic manifold (or a discrete group of interest). That is, if Γ_S is discrete, torsion-free, and $\mathfrak{Dicta}$ is satisfied for the given ordered generating set S and the set U. Our objective is to find bounds on S and U that come from geometric markings, or, even better, prove that geometric markings of a particular type must satisfy some collection of relators between elements of S. For example, in [**19**], it is show that any geometric marking in Cusps with cusp-volume at most 2.62 must have a (cyclically reduced) relator between m, n, g where the total occurrence of g or g^{-1} is at most 7.

Remark 4.2. The requirement of torsion-free appears here only because all of our examples are focused on hyperbolic 3-manifolds. One could just as easily drop this condition to look at orbifolds. In fact, it would be interesting to see if Marshall and Martin's proof that finds the two smallest volume hyperbolic-3 orbifolds [**31**] could be re-proven using techniques described here.

We now address the utility of markings. Consider Example Tubes, which looks for embedded tubes about shortest geodesics. Assume we start with some given $f, w \in \mathrm{PSL}(2, \mathbb{C})$ and tube U around the axis of f. We can then begin to search for possible contradictions to $\mathfrak{Dicta}$. In particular, if we find an element $g \in \Gamma_S \setminus \{id\}$ whose real translation length is provably less than that of f, then we know that the marking we started with cannot be geometric, and therefore can be ignored. Similarly, if we find an element $h \in \Gamma_S \setminus \langle f \rangle$ such that the interiors of U and $h \cdot U$ provably intersect, then we either violate discreteness, the fact that f must be primitive to be shortest, or the embeddedness of the tube corresponding to U, so again, the marking is not geometric. Both types of these contradictions are a bit subtle. It is possible that Γ_S is, in fact, discrete but simply our choice of f or U and w was incorrect. If we take care in building a large enough parameter space of markings, then we can assume a correct marking for this discrete Γ_S lies *elsewhere* in the parameter space. In particular, this marking can be ignored.

Definition 4.3. *A marking is* incorrect *if some condition in* $\mathfrak{Dicta}$ *is not satisfied by* S *and* U.

Another illustration of incorrect markings can be found in Example Margulis. Say that we find a word $w \in \Gamma_S \setminus \{id\}$ that moves our point $U = p$ less that x_1 or x_2, then even if Γ_S is discrete, this means that we chose the wrong pair x_1, x_2 to realize the Margulis number of Γ_S. Similarly, if we find a non-commuting pair $w_1, w_2 \in \Gamma_S$ and a point q for which w_1 and w_2 move q less than x_1 and x_2 move p, then our marking is incorrect and can be ignored. In particular, if Γ_S is discrete and w_1, w_2 actually realize its Margulis number, then a correct marking has $S = \{w_1, w_2\}$ and $U = q$. As remarked previously, in the actual search, we do not necessarily need to look for such w_1, w_2, q for the purposes of eliminating parameters if our goal is just to find the Margulis constant.

Before we construct parameters for markings, let us define the notion of equivalence for markings.

Definition 4.4. *Two markings* $(S_1, U_1, \mathfrak{Dicta}_1)$ *and* $(S_2, U_2, \mathfrak{Dicta}_2)$ *in* G *are said to be equivalent if there is an isomorphism* $\psi : \Gamma_{S_1} \to \Gamma_{S_2}$ *such that* $\psi(g_i) = h_i$ *for* $S_1 = \{g_1, \ldots, g_k\}$ *and* $S_2 = \{h_1, \ldots, h_k\}$ *and* $\mathfrak{Dicta}_1 = \mathfrak{Dicta}_2$ *under the replacing* g_i *by* h_i *and* U_1 *by* U_2.

Basic examples of this are conjugation in $\mathrm{PSL}(2, \mathbb{C})$ or replacing a generator in S by its inverse (as long as that doesn't break the validity of $\mathfrak{Dicta}$). The discussion in the next section shows how we can use the notion of equivalence to normalize our markings so that the space of parameters can be made small.

5. Parameter spaces

The next step is to build a parameter space $\mathcal{P} \subset \mathbb{C}^d$ that is large enough to include parameters representing each equivalence class of the geometric markings of interest. When picking $\mathcal{P}$, we want to makes sure that we can recover S and U (up to equivalence) from $\mathfrak{p} \in \mathcal{P}$ and, for computational purposes, that we try to make sure that each equivalence class of markings appears once or only a few times in $\mathcal{P}$.

For example, if we want to prove that the figure-eight knot complement and its sister are the two manifolds with smallest maximal cusp volume, then we need to consider markings of type Cusps that have maximal cusp volume at most $\sqrt{3}$. Our tasks will be (1) to show that this parameter space can be chosen to be compact up to equivalence of markings, (2) computationally eliminate large regions of this parameter space from consideration, and (3) show that the remaining regions only contain markings corresponding to the figure-eight knot complement and its sister. Tasks (2) and (3) will be addressed in Section 6. The goal of this section is to explain (1) in detail as this is a prototype for addressing a wide range of geometric problems.

5.1. Parameters for Cusps.

Here is how such a parameter space is constructed. Our generators are $\{m, n, g\}$ with m and n commuting non-parallel parabolics and U is a horoball centered at the fixed point of m and n. Under our $\mathfrak{Dicta}$ assumptions, we now do a sequence of transformations to find an equivalent marking that is nicely normalized. Working in the upper-half-space model of $\mathbb{H}^3$, we will conflate elements of $\mathrm{PSL}(2, \mathbb{C})$ with orientation preserving Möbius transformation and orientation-preserving isometries of $\mathbb{H}^3$. Note, there is some isometry ψ taking U to a horoball at infinity. Replacing m, n, g with their conjugates under ψ, we may assume that every element of $\langle m, n\rangle$ is of the form $z \mapsto z + c$ for some $c \in \mathbb{C}$. Let $|c|$ denote the length of such an element. To satisfy $\mathfrak{Dicta}$, we take $m(z) = z + \lambda$ be a shortest nontrivial element of $\langle m, n\rangle$. Conjugating by $z \mapsto \lambda^{-1} z$, we may assume that m is the map $z \mapsto z + 1$. Further, we can take $n(z) = z + \ell$ to be a shortest element of $\langle m, n\rangle$ linearly independent of m (i.e. *next-shortest*). Next, we have that $g \cdot U$ is a horoball tangent to U and centered at some point $c \in \mathbb{C}$. Conjugating by $z \mapsto z - c$, we may assume $g \cdot U$ is centered at 0. With these normalizations, we have $g^{-1}(z) = p + 1/(s^2 z)$ for some complex numbers p and s, where p is the center of the horoball $g^{-1}U$ and s controls its height and rotation.

To summarize, each equivalence class of markings of type CUSPS can be assigned a *parameter* $\mathfrak{p} = (p, s, \ell) \in \mathbb{C}^3$ with $s \neq 0$, were we can build a representative marking from this parameter by letting

$$g_{\mathfrak{p}}^{-1} = \pm \begin{pmatrix} ps\,\mathrm{i} & \mathrm{i}/s \\ s\,\mathrm{i} & 0 \end{pmatrix} \qquad g_{\mathfrak{p}} = \pm \begin{pmatrix} 0 & -\mathrm{i}/s \\ -s\,\mathrm{i} & ps\,\mathrm{i} \end{pmatrix}$$

$$m_{\mathfrak{p}} = \pm \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad n_{\mathfrak{p}} = \pm \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix}$$

and choosing $U_{\mathfrak{p}}$ as follows. Since we want $g_{\mathfrak{p}} \cdot U_{\mathfrak{p}}$ to be tangent to $U_{\mathfrak{p}}$, by equation (3.1), we must have

$$U_{\mathfrak{p}} = \{(z, t) \in \mathbb{H}^3 : t > 1/|s|\}.$$

In our parametrization, when the marking is geometric, we have $\mathrm{vol}(U_{\mathfrak{p}}/\Gamma_{S_{\mathfrak{p}}}) = |s^2\ \mathrm{im}(\ell)|/2$. Observe that in inverting this construction, we have guaranteed that some but not all of the $\mathfrak{Dicta}$ assumptions are satisfied. In particular, we know that m, n are commuting non-parallel parabolics and g is a loxodromic with $g \cdot U$ tangent to U. In the next few sections, we will use the rest of the $\mathfrak{Dicta}$ conditions to eliminate nearly all sets of parameters that give rise to incorrect markings.

We can now show that given a geometric marking of the type CUSPS with the restriction that the maximal cusp volume is $\leq \sqrt{3}$, then it is equivalent to a marking whose parameter lies in the following compact set $\mathcal{P}_{\text{CUSPS}} \subset \mathbb{C}^3$.

DEFINITION 5.1. *Let $\mathcal{P}_{\text{CUSPS}}$ be the subset of $\mathbb{C}^3$ defined by the following conditions:*

(0) $|s| \geq 1$
(1) $\mathrm{im}(s) \geq 0$, $\mathrm{im}(\ell) \geq 0$, $\mathrm{im}(p) \geq 0$, $\mathrm{re}(p) \geq 0$
(2) $-1/2 \leq \mathrm{re}(\ell) \leq 1/2$
(3) $|\ell| \geq 1$
(4) $\mathrm{im}(p) \leq \mathrm{im}(\ell)/2$
(5) $\mathrm{re}(p) \leq 1/2$
(6) $|s^2\ \mathrm{im}(\ell)| \leq 2\sqrt{3}$

PROPOSITION 5.2. *Suppose $(\{m, n, g\}, U, \mathfrak{Dicta})$ is a geometric marking of type* CUSPS *with a maximal cusp of volume at most $\sqrt{3}$. Then, there is $\mathfrak{p} = (p, s, \ell) \in \mathcal{P}_{\text{CUSPS}}$ such that $(\{m, n, g\}, U, \mathfrak{Dicta})$ is equivalent to $(\{m_{\mathfrak{p}}, n_{\mathfrak{p}}, g_{\mathfrak{p}}\}, U_{\mathfrak{p}}, \mathfrak{Dicta})$. Further, $\mathcal{P}_{\text{CUSPS}}$ is compact.*

PROOF. As above, we may assume U is centered at ∞ and $m(z) = z+1$ is a shortest-length generator and $n(z) = z + \ell$ is the next-shortest in $\langle m, n\rangle$. Since n is next-shortest, we have $|\ell| \geq 1$ and $-1/2 \leq \mathrm{re}(\ell) \leq 1/2$. If $\mathrm{im}(\ell) < 0$, then we replace n with $z \mapsto z - \ell$, which is also next-shortest and has $\mathrm{im}(-\ell) > 0$.

Next, let $p \in \mathbb{C}$ be the center at infinity of the horoball $g^{-1} \cdot U$. By post-composing g with elements of $\langle m, n\rangle$, we may choose g such that $-\mathrm{im}(\ell)/2 \leq \mathrm{im}(p) \leq \mathrm{im}(\ell)/2$ and $-1/2 \leq \mathrm{re}(\ell) \leq 1/2$. Reflecting across the x-axis creates isomorphic groups with isometric quotients. Thus we may assume $0 \leq \mathrm{im}(\ell) \leq \mathrm{im}(L)/2$. Reflecting across the y-axis allows us to assume likewise that $0 \leq \mathrm{re}(\ell) \leq 1/2$. Finally, changing s to $-s$ leaves f invariant, so we may assume $\mathrm{im}(s) \geq 0$. This determines p, s, and ℓ and a marking isomorphism. Properties (1), (2), (4), and (5) clearly hold and (3) holds since n is next-shortest. Further, since U/Γ_S is a maximal cusp neighborhood in $\mathbb{H}^3/\Gamma_S$, we know that U has Euclidean height ≤ 1. Thus, $1/|s| \leq 1$ and (0) holds.

Finally, the volume of U/Γ_S is $|s^2\,\mathrm{im}(\ell)|/2$, and this is at most $\sqrt{3}$, so (6) holds. Since all the inequalities are satisfied, $(p, s, \ell) \in \mathcal{P}$ as desired.

Note that the positive lower bounds for $|s|$ and $|\mathrm{im}(\ell)|$ impose upper bounds for $|s|$ and $|\mathrm{im}(\ell)|$ by using $|s^2\ \mathrm{im}(\ell)| \leq \sqrt{3}$. Thus $\mathcal{P}_{\text{CUSPS}}$ is compact. ☐

We hope that the above explicit example shows how to map markings to parameter spaces in general and what ideas to keep in mind. A good exercise is to work out a parameter space for TUBES. See **[21]** for details. Most importantly, we are now restricted to working with a bounded and explicitly described set of parameters $\mathcal{P}$. In the next section, we will discuss how to eliminate portions of this parameter space by looking for elements of Γ_S that violate the assumptions of $\mathfrak{Dicta}$.

REMARK 5.3. Our example looks at complex parameters because they are computationally more efficient in this setting. Using real parameters is possible, but requires computational considerations due to the choice of verified arithmetic used. This is one of several computational considerations that make the problems addressed in this paper computationally feasible. See Sections 7, 8 and 9 for more techniques.

6. Analyzing parameter spaces and elimination criteria

Naturally, we are interested in parameters $\mathfrak{p} \in \mathcal{P}$ where $\Gamma_{S_\mathfrak{p}}$ is geometric. To investigate these points it is convenient to introduce the following subsets of $\mathcal{P}$.

DEFINITION 6.1. *Let*

$$\mathcal{D} = \{\mathfrak{p} \in \mathcal{P} : \Gamma_{S_\mathfrak{p}} \textit{ is discrete, non-elementary, and torsion-free}\}.$$

and

$$\mathfrak{D} = \{\mathfrak{p} \in \mathcal{D} : (S_\mathfrak{p}, U_\mathfrak{p}, \mathfrak{Dicta}) \textit{ is equivalent to a geometric marking}\}.$$

As alluded to previously, the delicate distinction between $\mathfrak{D}$ and $\mathcal{D}$ is that, in general, even if $\Gamma_{S_\mathfrak{p}}$ is discrete, there could be an element $h \in \Gamma_{S_\mathfrak{p}}$ that violates some condition in $\mathfrak{Dicta}$. In Example CUSPS, recall that a violating $h \in \Gamma_{S_\mathfrak{p}}$ could be one where the interiors of the horoballs $h \cdot U_\mathfrak{p}$ and $U_\mathfrak{p}$ intersect, or h is parabolic and shorter than $m_\mathfrak{p}$. Thus, if we chose sufficiently many conditions for $\mathfrak{Dicta}$, we can expect that $\mathcal{D} \setminus \mathfrak{D}$ is comprised of *incorrect* markings. That is, we want to choose $\mathfrak{Dicta}$ flexible enough so that the set of correct markings is contained inside $\mathfrak{D}$. In particular, this would imply that for any $\mathfrak{p} \in \mathcal{D} \setminus \mathfrak{D}$ there is $\mathfrak{q} \in \mathfrak{D}$ such that $\Gamma_{S_\mathfrak{p}} = \Gamma_{S_\mathfrak{q}}$, allowing us to eliminate/ignore the parameter $\mathfrak{p}$.

Recall that $\mathfrak{Dicta}$ must also be chosen such that if a condition fails, then we can computationally verify this fact in finite time. As such, the conditions in $\mathfrak{Dicta}$ that cannot be encoded into the parametrization itself must be verifiable by evaluating some inequality. Regions where the inequality fails can then be eliminated from $\mathcal{P}$. Since we are working with groups, our $\mathfrak{Dicta}$ conditions can usually be falsified by finding violating elements $h \in \Gamma_{S_\mathfrak{p}}$, just like in the example above. Of course, one does not need to be restricted to only such criteria. For example, the bound on maximal cusp volume in $\mathcal{P}_{\text{CUSPS}}$ could be thought of as a $\mathfrak{Dicta}$ criterion. However, we will focus on word-related elimination criteria here as they are quite special to this type of parameter space analysis.

6.1. Killer words. Since $S_\mathfrak{p} = \{g_1, \ldots, g_k\}$ is an explicit set of generators $\Gamma_{S_\mathfrak{p}}$ in our Lie group G, it is best to consider words in the free group $\mathfrak{F}$ on symbols $\{t_1, \ldots, t_k\}$ and a presentation $\rho_\mathfrak{p} : \mathfrak{F} \to \Gamma_{S_\mathfrak{p}}$ where $\rho_\mathfrak{p}(t_i) = g_i$. We will often abuse notation and use g_i to denote t_i when the context is clear.

DEFINITION 6.2. *For a word $w \in \mathfrak{F}$ and $\mathfrak{p} \in \mathcal{P}$, let*

$$w_\mathfrak{p} = \rho_\mathfrak{p}(w) = \pm \begin{pmatrix} a_w(\mathfrak{p}) & b_w(\mathfrak{p}) \\ c_w(\mathfrak{p}) & d_w(\mathfrak{p}) \end{pmatrix} \in \mathrm{PSL}(2, \mathbb{C}).$$

We will words in $\mathfrak{F}$ to define computable open sets in $\mathcal{P}$ that are guaranteed to be disjoint from $\mathfrak{D}$. Essentially, we look for words that violate conditions specified in $\mathfrak{Dicta}$. If we can give a cover of $\mathcal{P}$ by such open sets, then $\mathcal{P}$ cannot cannot contain any geometric markings. If we are not so lucky, we can at least eliminate large chunks of $\mathcal{P}$ by cutting away these open sets and be left with a few regions of interest. Below are two examples, one of type TUBES and one for CUSPS.

LEMMA 6.3. *Consider markings of type* TUBES *with parameter space* $\mathcal{P}_{\text{TUBES}}$. *Given* $w \in \mathfrak{F}$, *define*

$$\mathcal{S}_w = \{\mathfrak{p} \in \mathcal{P}_{\text{TUBES}} : 0 < \operatorname{re}\operatorname{arccosh}(\operatorname{tr}(w_\mathfrak{p})/2) < \operatorname{re}\operatorname{arccosh}(\operatorname{tr}(f_\mathfrak{p}))/2)\}.$$

Then $\mathcal{S}_w \cap \mathfrak{D} = \emptyset$.

PROOF. Notice that $\mathcal{S}_w$ is the open set of parameters where the real translation length of $w_\mathfrak{p}$ is nonzero but shorter than that of $f_\mathfrak{p}$. Since the translation length of $w_\mathfrak{p}$ is positive, $w_\mathfrak{p}$ is a loxodromic and so cannot be the identity. If $\mathfrak{p} \in \mathfrak{D}$, this would violates the criterion that $f_\mathfrak{p}$ has shortest real translation length, so we conclude $\mathcal{S}_w \cap \mathfrak{D} = \emptyset$. □

LEMMA 6.4. *In the context of markings of type* CUSPS, *define*

$$\mathcal{Z}_w = \{\mathfrak{p} = (p, s, \ell) \in \mathcal{P}_{\text{CUSPS}} : |c_w(\mathfrak{p})/s| < 1\}$$

and

$$\mathcal{K}_w = \{\mathfrak{p} \in \mathcal{Z}_w : |c_w(\mathfrak{p})| > 0 \text{ or } a_w(\mathfrak{p}) \neq \pm 1 \text{ or } d_w(\mathfrak{p}) \neq \pm 1\}.$$

Then $\mathfrak{D} \cap \mathcal{K}_w = \emptyset$ *for all* $w \in \mathfrak{F}$.

PROOF. Pick $\mathfrak{p} \in \mathcal{K}_w \subset \mathcal{Z}_w$. Let $U_w(\mathfrak{p})$ be the image of the horoball $U_\mathfrak{p}$ under $w_\mathfrak{p}$. Recall that $U_\mathfrak{p}$ has Euclidean height $1/|s|$. Thus, by 3.1, $U_w(\mathfrak{p})$ has Euclidean height $|s|/|c_w(\mathfrak{p})|^2$. It follows that the interiors of $U_\mathfrak{p}$ and $U_w(\mathfrak{p})$ intersect precisely when $|s|/|c_w(\mathfrak{p})|^2 > 1/|s|$, or equivalently $|c_w(\mathfrak{p})/S| < 1$.

Assume also that $\mathfrak{p} \in \mathfrak{D}$. Then $\Gamma_{S_\mathfrak{p}}$ is geometric and therefore $U_\mathfrak{p}$ is maximal. Since the interiors of $U_w(\mathfrak{p})$ and $U_\mathfrak{p}$ intersect, we must have $U_\mathfrak{p} = U_w(\mathfrak{p})$ by discreteness and therefore $w_\mathfrak{p} \in \langle m_\mathfrak{p}, n_\mathfrak{p}\rangle$. However, since $\mathfrak{p} \in \mathcal{K}_w$, it is impossible for $w_\mathfrak{p}$ to be parabolic fixing ∞, a contradiction. Thus, $\mathfrak{D} \cap \mathcal{K}_w = \emptyset$ for all $w \in \mathfrak{F}$. □

Note, it is possible that $\mathcal{K}_w \cap \mathcal{D}$ is non-empty, but all such points would be incorrectly marked as discussed previously. The upshot of this lemma is that we can eliminate chunks of $\mathcal{P}_{\text{CUSPS}}$ by finding words for which $\mathcal{P}_{\text{CUSPS}} \cap \mathcal{K}_w$ is non-empty. Similarly for $\mathcal{S}_w$ and $\mathcal{P}_{\text{TUBES}}$. We call such words *killer words*. Notice that if $\bigcup_{w \in \mathfrak{F}} \mathcal{K}_w$ is a cover of $\mathcal{P}_{\text{CUSPS}}$, then it follows that $\mathfrak{D}$ is empty. In particular, if $\mathcal{P}_{\text{CUSPS}}$ is made a bit smaller by setting the upper bound of $\sqrt{3} - \epsilon$ for the maximal cusp volume, then such a cover exists and is given in [**19**]. However, it is often the case that we are not as lucky. For example, if the cutoff is set at $\sqrt{3} + \epsilon$, then the parameters corresponding to a marking of the figure eight knot complement and its sister will never be covered by $\mathcal{K}_w$ for any w, leaving us with two small regions that we cannot eliminate. One could stop here and have a result that every cusped hyperbolic 3-manifolds has cusp volume greater than $\sqrt{3} + \epsilon$ with the exception of manifolds that contain a marking whose parameter falls into one of these two small regions. Luckily, we can often do better. In the next section, we show how to get further restrictions for the location of the $\mathfrak{D}$ parameters.

6.2. Variety and quasi-relator words. Outside of pure violations of markings, we can make use of many rigidity results of Kleinian groups. The goal here is to extract a list of topological properties that must be satisfied by geometric markings. Our first example is that of the set $\mathcal{Z}_w$ defined for markings of type CUSPS in the previous section.

LEMMA 6.5. *For markings of type* CUSPS, *if* $\mathfrak{p} \in \mathcal{Z}_w \cap \mathfrak{D}$, *then* $m_\mathfrak{p}^a n_\mathfrak{p}^b w_\mathfrak{p} = id$ *for some* $a, b \in \mathbb{Z}$.

PROOF. In the proof of Lemma 6.4, we saw that $w_{\mathfrak{p}} \in \langle m_{\mathfrak{p}}, n_{\mathfrak{p}} \rangle$ by the fact that $|c_w(\mathfrak{p})/s| < 1$ forces the interiors of $U_{\mathfrak{p}}$ and $U_w(\mathfrak{p})$ to intersect and discreteness of $\Gamma_{S_{\mathfrak{p}}}$ forces them to be equal. Thus, the result follows. □

If we can now cover $\mathcal{P}_{\text{Cusps}} \setminus \bigcup_{w \in \mathfrak{F}} \mathcal{K}_w$ by a finite collection $\mathcal{Z}_{w_1}, \ldots, \mathcal{Z}_{w_k}$, then we have a lot of restrictions on the groups corresponding to parameters in $\mathfrak{D}$. Call such words *quasi-relator* words, since they imply the existence of relators of a given form. In **[19]**, such a cover is found for markings of hyperbolic 3-manifolds with a maximal cusp of volume at most 2.62 and by analyzing the quasi-relator words the authors prove that all such manifolds were Dehn fillings of a finite list of known manifolds.

Note, in the above example, some extra conditions could be used to force $a = b = 1$. For example, we could require that $w_{\mathfrak{p}}$ move the point j less than $m_{\mathfrak{p}}$, which would imply $w_{\mathfrak{p}} = id$. Thus, one could even find explicit relators using the above technique. Another useful trick that is similar in nature uses the rigidity provided by Jørgensen's and the Shimizu-Leutbecher inequalities.

DEFINITION 6.6. *Define the variety of* w *by* $\mathcal{V}_w = \{\mathfrak{p} \in \mathcal{P} : w_{\mathfrak{p}} = id\}$.

LEMMA 6.7. *For markings of type* CUSPS, *define the neighborhood* $\mathcal{N}_w$ *of* $\mathcal{V}_w$ *by*

$$\mathcal{N}_w = \{\mathfrak{p} \in \mathcal{P} : |c_w(\mathfrak{p})| < 1 \text{ and } |b_w(\mathfrak{p})| < 1\}.$$

For all nontrivial $w \in \mathfrak{F}$, *if* $\mathfrak{p} \in \mathcal{D} \cap \mathcal{N}_w$ *then* $\mathfrak{p} \in \mathcal{V}_w$. *In particular,* w *is a relator in* $\Gamma_{S_{\mathfrak{p}}}$.

PROOF. The Shimizu-Leutbecher Theorem states that if

$$m_{\mathfrak{p}} = \pm \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \pm \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

generate a discrete subgroup of $\mathrm{PSL}(2, \mathbb{C})$ and $|c| < 1$, then we must have $c = 0$. For $\mathfrak{p} \in \mathcal{D}$, $\Gamma_{S_{\mathfrak{p}}}$ is discrete and torsion-free, in particular, $m_{\mathfrak{p}}$ and $w_{\mathfrak{p}}$ generate a discrete subgroup of $\mathrm{PSL}(2, \mathbb{C})$. Suppose $\mathfrak{p} \in \mathcal{D} \cap \mathcal{N}_w$. Dy definition $|c_w(\mathfrak{p})| < 1$, but discreteness implies $c_w(\mathfrak{p}) = 0$, so $w_{\mathfrak{p}} \in \langle m_{\mathfrak{p}}, n_{\mathfrak{p}} \rangle$. Further, since $|b_w(\mathfrak{p})| < 1$ and $m_{\mathfrak{p}}$ is a shortest-length generator of $\langle m_{\mathfrak{p}}, n_{\mathfrak{p}} \rangle$, we must have $b_w(\mathfrak{p}) = 0$. Therefore, we conclude that $\mathfrak{p} \in \mathcal{V}_w$. □

If we find $w \in \mathfrak{F}$ such that $\mathcal{N}_w \cap \mathcal{P} \neq \emptyset$, then we automatically know that any geometric parameters in that intersection would have w as a relator (and have to lie on $\mathcal{V}_w$). We call such a w a *variety* word or *relator* word. Taking this a step further, if $\mathfrak{p} \in \mathcal{N}_{w_1} \cap \mathcal{N}_{w2}$, we know that either $\mathfrak{p}$ arises form geometric marking that have both w_1 and w_2 as relators, or it can be ignored. This technique can now allow us to pinpoint geometric markings exactly as follows. $\mathcal{V}_{w_i}$ corresponds to a variety in $\mathcal{P} \subset \mathbb{C}^d$ whose polynomial equations are encoded in w and use the parameters as variables. The set $\mathcal{V}_{w_1} \cap \mathcal{V}_{w_2} \cap \mathcal{P}$ can often be analyzed by using a computer algebra system, especially if the intersection is a discrete set of points. If we are lucky, we can hope for the following three things to happen (1) $\mathcal{V}_{w_1} \cap \mathcal{V}_{w_2} \cap \mathcal{P}$ is just one point, (2) the group with presentation $\langle t_1, \ldots, t_k \mid w_1, w_2 \rangle$ can be identified as the fundamental group of a manifold Y, and (3) Y contains a geometric marking that lands in $\mathcal{P}$. If this is the case, then the intersection point must arise form a geometric marking of Y. Often, this is enough to identify the manifolds on the nose, as is done in **[19]** for the figure eight knot and its sister to show that they have the smallest maximal cusp-volume.

REMARK 6.8. A lemma similar to 6.7 can be obtained from Jørgensen's inequality. Pick $g \in S$ and let

$$\mathcal{J}_w = \{\mathfrak{p} \in \mathcal{P} : |tr(w_{\mathfrak{p}})^2 - 4| + |tr(w_{\mathfrak{p}} g_{\mathfrak{p}} w_{\mathfrak{p}}^{-1} g_{\mathfrak{p}}^{-1} - 2| < 1 \text{ or } |tr(g_{\mathfrak{p}})^2 - 4| + |tr(g_{\mathfrak{p}} w_{\mathfrak{p}} g_{\mathfrak{p}}^{-1} w_{\mathfrak{p}}^{-1} - 2| < 1\}$$

Then for any $\mathfrak{p} \in \mathcal{J}_w$, we must have that $g_\mathfrak{p}$ and $w_\mathfrak{p}$ generate an elementary Kleinian group. In particular, if g is loxodromic, we get a relator of the form $wgw^{-1}g^{-1}$.

7. Computational setup

Now that we have explored the notion of markings, their parameter spaces, and various elimination criteria, our next goal is to describe the computational aspects involved in finding a provable cover of $\mathcal{P}$ by open sets of the forms $\mathcal{K}_w, \mathcal{Z}_w, \mathcal{N}_w$, and similar, depending on the context. The basic idea is to subdivide $\mathcal{P}$ into small boxes and to associate to each box a word w for which we can computationally verify that the entire small box lies in the specified the open set associated to w. Thus, to provide a proof of the existence of such a cover, we need to do three things (1) encode the locations of all the small boxes whose union contains $\mathcal{P}$, (2) find a word and associated open set that covers each box, and (3) to provide verified computational tools to check if a box lands in a given open set. As we shall see, one benefit of this approach will be that the proof can be stored separate from the computational tools and also does not require any complex logic that may have been used to find the words associated to the cover. In fact, as tools for verified computation become better and easier to use, the arithmetic used in **[21]** and **[19]** could easily be replaced while still using the original data. For completeness, we will give a brief discussion of this arithmetic in the last section. Here, we will focus on our methods for accomplishing for (1) and (2).

To encode our subdivision of $\mathcal{P}$ into small boxes, we use two tricks. First, we embed $\mathcal{P}$ into a large bounding box $\mathcal{B}$ that will be easier to subdivide. Then, for the subdivision, we will use the technique of Binary Space Partitioning, which was first used in computer graphics in the 1960's.

Again, we focus on the case of Cusps, with much of the overview below appearing in **[19]**. We place the compact parameters space $\mathcal{P}_{\text{Cusps}}$ into a large box

$$\mathcal{B} = \{(x_0, x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^6 : |x_i| \leq \texttt{scale} \cdot 2^{-i/6}\},$$

where each subsequent side of the box is $1/\sqrt[6]{2}$ times the size of the previous and `scale` is some scaling factor. For $\mathfrak{p} = (p, s, \ell) \in \mathcal{P}_{\text{Cusps}}$, the embedding is given by $\ell = x_3 + \mathrm{i}x_0$, $s = x_4 + \mathrm{i}x_1$, and $p = x_5 + \mathrm{i}x_2$. In **[19]**, `scale` is taken to be 8, which makes $\mathcal{B}$ is large enough to contain all of $\mathcal{P}_{\text{Cusps}}$ for manifolds of maximal cusp volume less than 2.62.

The constant side ratio of $1/\sqrt[6]{2}$ for $\mathcal{B}$ is chosen specifically such that if we cut along the 1^{st} dimension, then the two resulting boxes have the same constant side ratio as $\mathcal{B}$. Cutting those along the 2^{nd} dimension yields 4 boxes that still have the same constant side ratio. This is reminiscent of the $1/\sqrt{2}$ side ratio of A-series printer paper, except for 6-dimensional boxes instead of rectangles in the plane.

This behavior yields two advantages. First, the (sub-)boxes of $\mathcal{B}$ obtained by cutting in this manner stay relatively "round," making computational corrections for rounding error less dramatic. Second, this subdivision allows us to encode these (sub-)boxes in binary as *boxcodes*. For example, a boxcode 0 corresponds to the box $\mathcal{B}_0$ obtained by cutting $\mathcal{B}$ in half along the 1^{st} dimension and taking the resulting box on the "left." We fix some preferred orientation on $\mathbb{R}^6$ to define "right" and "left." The box $\mathcal{B}_{01}$ corresponds to cutting $\mathcal{B}_0$ in half along the 2^{nd} dimension and taking the piece on the "right." For deeper boxes, we keep cutting along the next dimension and after cutting along the 6^{th} dimension, we start again with the 1^{st}. For a boxcode $\mathfrak{b}$, we let $\mathcal{B}_\mathfrak{b}$ denote the corresponding box. See Figure 1 for a picture of this subdivision for a box in $\mathbb{R}^3$.

To analyze all of $\mathcal{B}$, we build a binary tree $\mathcal{T}$ corresponding to the boxcodes. The terminal nodes of the tree give a subdivision of $\mathcal{B}$ into boxes of different sizes. The reason for this approach is that the number of boxes needed ca be quite large. In **[19]**, the number of boxes was 1,394,524,064. If we were

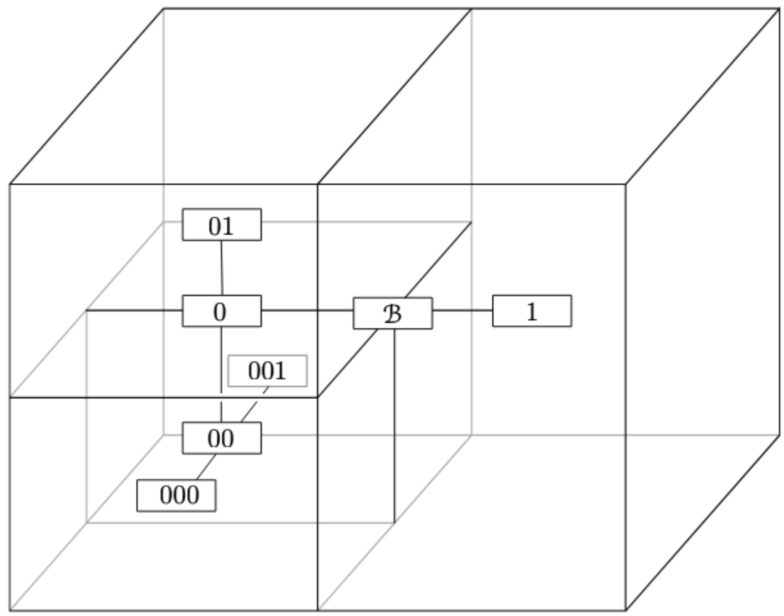

FIGURE 1. Boxcodes and subdivision in $\mathbb{R}^3$ with labels at the box centers from **[21]**.

to divide $\mathcal{B}$ into boxes all of the same size, our computation and verification time would exponentially increase.

Each terminal node $\mathfrak{b}$ of this tree has an associated killer word, quasi-relator word, variety word, boundary condition, or other desired elimination criterion. Here, a *boundary condition* is simply a test that implies that the box does not intersect the embedding of $\mathcal{P}$ in $\mathcal{B}$. In the $\mathcal{P}_{\text{Cusps}}$ example, this involves checking that one of the inequalities defining $\mathcal{P}_{\text{Cusps}}$ fails over the entire box. Similarly, if w is a killer word associated to $\mathfrak{b}$, then we can rigorously verify that $\mathcal{B}_{\mathfrak{b}} \subset \mathcal{K}_w$ by evaluating a series of inequalities encoded by w. These conditions could be checked in theory by hand given an unreasonably large amount of time.

The verification process simply checks conditions at all terminal nodes by traversing the binary tree from the root node (in depth-first order). A successful traversal of the tree guarantees that the terminal boxes form a cover of $\mathcal{B}$ and therefore $\mathcal{P}$. Since each terminal box either misses $\mathcal{P}$ or lies inside an associated open set $\mathcal{K}_w, \mathcal{Z}_w, \mathcal{N}_w$, or similar, we have a good over of $\mathcal{P}$ that provides information about $\mathfrak{D}$. Example code of a verification program can be found alongside **[21]** and **[19]**. Further, **[21]** goes through specific examples in the text that provide a good understanding of the verification process.

8. Finding useful words

Here, we outline our methods for successfully finding useful words that allow us to obtain a complete decomposition of $\mathcal{B}$ into (sub-)boxes with associated words.

A naive approach uses diagonal enumeration to combine breadth-first enumeration of the tree of boxes with breadth-first enumeration of the tree of words (with three edges from each word, one for each generator or inverse-generator that isn't the inverse of the last symbol in the word, pointing to the concatenation of the word with the generator). The naive algorithm has running time $O(3^L 2^D)$, where L is maximum word length and D is maximum box depth. This is much too slow.

To speed it up, there are three basic approaches, all of which are necessary:

(1) avoid considering most boxes by stopping once we have a solution
(2) reuse words that work on one box elsewhere
(3) use geometric heuristics to prefer words that are more likely to work.

The exposition will proceed in rough chronological order, in the hope that by describing some of the wrong turns, we'll help others avoid making the same mistakes. Note, we will mostly focus on killer words because they tend to also work as relator and variety words. In fact, you can see in the examples given, killer words are almost always relator words that just also prove the word is not a relator in the box. However, we often want our relator or variety words to be few and to have special properties, so killer words are still necessary.

The most obvious way of speeding up the search is to avoid the search entirely when feasible: a killer word works on a neighborhood of a region, and by testing killer words found for nearby boxes, most of the time the search is not necessary.

Still, some searches have used words as long as 44, and testing all of the roughly 3^{44} combinations would be prohibitive. Rather than blindly selecting words in first-in-first-out order, the algorithm can rank the words under consideration based on a heuristic estimate of the likelihood of their being useful. We note first that short words tend to be better than long words, as they have fewer steps and less error accumulation across computations. Second, we want words that we expect to violate conditions in $\mathfrak{Dicta}$. This usually involves some geometric measurement about how far a word moves the set U for parameters in a box. For example, the algorithm for the search in **[19]** for markings of type CUSPS was derived from code that would find all w (of bounded word length) for which the horoball $w\dot{U}$ has center in a fundamental domain of $\langle m, n\rangle$ and whose Euclidean height is greater than some cutoff δ. Often, this would quickly find a ball that would crash into U. The algorithm can be found as part of visualizing software at **[26]**, but is a fun exercise on its own. We describe this algorithm at the end of this section. In summary, the ranking of a word w for a box was a combination of the word length and how close the image of $w \cdot U$ was to U over the box. Such an algorithm, of course, heavily depends on the geometry of the marking. Further, if different types of words are required, mixing several geometric search algorithms could be considered.

Note, this approach can get stuck. If the box is close to a variety, you will often keep finding the same variety word over and over. In some cases, one variety word may be enough, but not in others. For example, the relator you keep finding is not of the form you need or you may need two or more relators to identify a specific group. To step around this, one needs a "diversity" heuristic, which asks for words of a different type, or just a list of "bad relators." For example, if a relator that has been found as a particular abelianization, then look for relators with different abelianizations to guarantee that they are independent.

One other important aspect about searching for and evaluating words: don't use rigorous arithmetic until you think you should. Whenever a word was evaluated, a kind of triage should be used to determine whether that word was likely to kill the box in question, likely to kill any of its n^{th} generation descendants, or unlikely to kill any descendants of the box. An answer to such a heuristic tells us whether to try a verified evaluation (with the error term included), defer further evaluation until the box had been subdivided n more times, or exclude that word from further consideration on any descendant of the box. As simple example of such a heuristic is to use the fastest arithmetic you have to just evaluate at the center of the box. Then, if a words works for the center, check the corners. Finally, if it works on all the corners, try using expensive verified arithmetic over the entire box. However, if the value at the center is too far from being successful for the size of the box (since words encode inequalities so there is natural measure of "too far"), then the word is probably not going to work for any descendants. And, of course, if the word seems close to working, it might be good to have it at the top of the list for evaluation on smallest sub-boxes. With these heuristics, the program for **[21]** wound up using expensive arithmetic, on average, about 10 of the roughly 13200 word-bank words per box.

Even with such heuristics in mind, it is important to have control of the word search to prevent it from from running forever. It should be should be temporarily abandoned after some number of steps, and re-done once the number of descendant boxes doubles or is sufficiently large by some other metric. This way, the search could run forever, but only if the subdivision process runs forever. Knowing how to profile your code to see where the computation spends most of its time is also quite useful. This merged process of alternately searching and subdividing we call the *decomposition algorithm.*

The decomposition algorithm can go through several revisions. One benefit of using the tree data structure and for recording successful elimination of boxes is that the search can restart at the boxes that are still open, or it can run over the tree again to make improvements or corrections.

Usually the first attempt used to determine the feasibility of the whole effort and computational correctness — should do the search in depth-first order on a collection of small sub-boxes of particular interest. These could be boxes around known discrete points or places where simple geometric arguments are hard to do by hand. Such an initial step has two benefits. First, it checks that your arithmetic and word search code works to provide expected results. Second, it seeds a bank of words that can now be used in other parts of the parameter space.

After some amount feasibility has been determined, a run that uses the existing word bank and the basic boundary conditions can now try to sculpt away the parts of $\mathcal{B}$ that are easy to eliminate. Essentially, leaving parts of $\mathcal{P}$ that aren't close to the killer neighborhoods of the words found in the initial targeted search. Note, this search should run in breadth-first order and do rather little word searching as it is a time consuming process.

In the third stage, one can try a more aggressive word search for the remaining trouble regions, possibly with human input. For example, software that can visualize the action of the group Γ_S on U (or other sets of interest) can be very helpful in finding or even guessing killer words or new conditions that can be added to kill boxes. The search heuristic can also be amended and updated to take into account new observations. Having tools that can provide information about the remaining boxes can also be quite helpful.

If the third stage successfully produces a complete tree, a final step could be do a run that tries to reduce the number of boxes. Essentially, this run attempts all found killer words in a large region (about a thousand boxes) on all the parent boxes in the region and does no new searching.

Finally, once a completed tree is produced, the last task is to produce a minimal program that reads and check the inequalities encoded in the tree using verified arithmetic. This is important to make sure no bugs in the search code have an effect on tree validation. Further, this allows for including clear and coherent documentation for each test. See **[41]** for an example of such a program.

8.1. Cusps word search algorithm. Recall that when searching for candidate killer or relator words in the Cusps setup, we are interested in finding all images of the horoball $U = H_\infty$ that have large height. Since we are interested in short words, we only care about such translates $w \cdot H_\infty$, where w uses few g's or G's. We call this number of g's and G's the g-length of w. In particular, walk the group and find all horoballs of height at least some h which are the images of words of g-length at most some K. We start with the horoballs $g \cdot H_\infty$ and $G \cdot H_\infty$. Next, we ask the question: how far can we translate them via $\langle m, n\rangle$ until we know their future images under g or G will be smaller than h? Indeed, such a bound must exists because the distance between $g \cdot H_\infty$ and $m^a n^b g H_\infty$ will then be the distance between H_∞ and $G m^a n^b g H_\infty$. But this latter distance controls the height of $G m^a n^b g H_\infty$. A similar distance argument applies to $G \cdot H_\infty$ and any translate. Since for each parameter $g \cdot H_\infty$ and $G \cdot H_\infty$ are fixed, they can compute a bound on the exponents a, b. Iterating this process allows us to find all horoballs with a given height cutoff which are images of words with given g-length. In a discrete group, this process will simply generate disjoint horoballs. However, as soon as one leaves the discrete locus, horoballs will start to intersect. Conjugating one of an intersecting pair to H_∞ guarantees the existence of a large horoball at an indiscrete parameter.

9. Arithmetic and computational considerations

The inequalities in the definitions of $\mathcal{K}_w$, $\mathcal{U}_w$, and $\mathcal{V}_w$ are straightforward to check after constructing the matrix corresponding to w at points of the box $\mathcal{B}_{\mathfrak{b}}$. To prove that these conditions hold over the entire box, we often use affine 1-jets with error and round-off error for computations, which was introduced in [**21**]. If one wanted to use interval arithmetic instead, we expect that further subdivision would be necessary. Using our arithmetic or similar should have significant benefits when implementing a version of the decomposition algorithm. The reason it generally works well is that affine 1-jets approximate a function over the box and not just upper and lower bounds of that function. So if the sum of two functions cancels out and is close to zero, our arithmetic can see this over a large box, while interval arithmetic would need rather fine subdivision.

9.1. 1-jets and error control. Here, we define affine 1-jets in the context of $\mathcal{P}_{\text{CUSPS}}$, which is a 3-complex dimensional parameter space. Let

$$\mathcal{A} = \{(z_0, z_1, z_2) \in \mathbb{C}^3 : |z_i| \leq 1\}$$

and consider a (holomorphic) function $g : \mathcal{A} \to \mathbb{C}$. An affine 1-jet approximation with error ϵ of g is a linear map $(z_0, z_1, z_2) \mapsto c + a_0 z_0 + a_1 z_1 + a_0 z_2$ such that

$$|g(z_0, z_1, z_2) - c + a_0 z_0 + a_1 z_1 + a_0 z_2)| \leq \epsilon.$$

Considering all maps approximated by a given 1-jet, one defines the jet-set

$$S(c, a_0, a_1, a_2; \epsilon) = \{g : \mathcal{A} \to \mathbb{C} : |g(z_0, z_1, z_2) - c + a_0 z_0 + a_1 z_1 + a_0 z_2)| \leq \epsilon\}.$$

In [**21**], the authors derive the arithmetic for jets. For example, given jet-sets S_1 and S_2, they show how to compute the parameters for a jet-set S_3 where $f \cdot g \in S_3$ for all $f \in S_1$ and $g \in S_2$. They do this for all basic arithmetic operations $\pm, \times, /$, and $\sqrt{\cdot}$. If one requires other (holomorphic) operations, new code needs to be written, as well as the accompanying proof of its validity. Note, non-holomorphic arithmetic will require additional parameters as the derivative of a function $f : \mathbb{R}^6 \to \mathbb{R}^2$ is 12 real-dimensional. Code for a version of this arithmetic should be available soon as it is necessary for MARGULIS [**16**].

For our computations in most parameter searches to date, we have used this exact same arithmetic. Note, to make this arithmetic fast, the parameters of $S(c, a_0, a_1, a_2; \epsilon)$ are all given as floating-point doubles. Recall that floating-point numbers are a finite subset of the reals represented via a list of bits on a computer. As such, any arithmetic operation between two floating-point numbers must make a choice about which floating-point result to give, as the real value of this result may not be representable as a floating-point number. On a machine conforming to the IEEE-754 standard [**1**], all operations performed with floating points numbers are guaranteed to round in a consistent way to a closest floating-point representative, as long as overflow and underflow have not occurred. Thus, code for all verification projects must include check for underflow and overflow during validation, see [**41**] for an example.

In the case of CUSPS, just as in Section 7 of [21], each boxcode corresponds to a (sub)box of our parameter space that has a floating-point center $(c_0, c_1, c_2, c_3, c_4, c_5)$ and a floating-point size $(s_0, s_1, s_2, s_3, s_4, s_5)$. Using the IEEE-754 standard, the floating-point size is large enough so that the floating-point version of the box, i.e. $\{(x_0, x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^6 : |x_i - c_i| \leq s_i\}$, contains the true box as a proper subset. Recall that over a box we associate the coordinate functions $\ell = x_3 + \mathrm{i}x_0$, $s = x_4 + \mathrm{i}x_1$, and $p = x_5 + \mathrm{i}x_2$ for our parameters $\mathfrak{p} = (p, s, \ell) \in \mathcal{P}_{\text{CUSPS}}$. We can replace these coordinate functions with linear maps $g_\ell, g_s, g_p : \mathcal{A} \to \mathbb{C}$ given by:

$$g_\ell(z_0, z_1, z_2) = c_3 + \mathrm{i}c_0 + (s_3 + \mathrm{i}s_0)z_0,$$

$$g_s(z_0, z_1, z_2) = c_4 + \mathrm{i}c_1 + (s_4 + \mathrm{i}s_1)z_1,$$

$$g_p(z_0, z_1, z_2) = c_5 + \mathrm{i}c_2 + (s_5 + \mathrm{i}s_2)z_2.$$

Notice that, by construction, the linear maps see all the ℓ, s, p values over the given box. This means that for every point $(p, s, \ell) \in \mathcal{B}_\mathfrak{b}$, there are z_0, z_1, z_2 such that $\ell = g_\ell(z_0, z_1, z_2)$, $s = g_s(z_0, z_1, z_2)$ and $p = g_p(z_0, z_1, z_2)$. In fact, these functions also see values that are a little outside of $\mathcal{B}_\mathfrak{b}$ due to the change from a box in $\mathbb{R}^6$ to a box in $\mathbb{C}^3$. We can now think of g_ℓ, g_s, g_p as living in jet-sets. For example, $g_p \in S(c_5 + \mathrm{i}c_2, 0, 0, s_5 + \mathrm{i}s_2; 0)$ and similarly for others. It follows that any evaluations with jet-arithmetic will include the true values over the box up to the error that accumulates. See [21], Sections 7 and 8 for concrete details. Lastly, it is important to remember that this verified arithmetic should be used sparingly. Specifically, only when the chances of a successful elimination of a box are high. This is especially important if the dimensionality of the parameter search increases and the cost of this arithmetic grows polynomially in dimension for most operations.

10. Sanity checking

One of the benefits of keeping the verification code separate from the search code is that the latter is prone to frequent changes which may introduce errors. Making sure that the validation is done by a concise, clean, and easily-readable piece of code is essential. For a deeper discussion on why one should be confident in such a `verify` program and its internal consistency, we point the reader to the introduction of [21, p 339-341] that covers many aspects of this question in detail. However, during the search process itself, it is also important to preform sanity checks and to build tools that allow by-hand analysis of found examples consistent with theory. To accomplish this, we use several strategies: (1) find boxcodes for as many known examples as one can and validate that the search behaves as expected at those boxcodes, (2) write visualization tools to analyze marking, (3) use other tools or hands-on methods to verify the output of the search, and (4) write and re-run tests of the arithmetic whenever changes to it are made.

In many of our contexts, the first task can be accomplished by using a census of known manifolds. For finite-volume hyperbolic 3-manifolds and knot exteriors, SnapPy [15] contains several large collections tabulated by Hodgson, Weeks, Callahan, Dean, Weeks, Champanerkar, Kofman, Patterson, Dunfield, and Christy. Extracting the relevant invariants from the census can be a challenge in its own right, but having concrete boxcodes to test against is essential. See [42] for an example.The search must correctly classify all known examples and discrepancies discovered must be thoroughly investigated. Such discrepancies could include things like mistakes in the census-to-boxcode mapping or making incorrect assumptions or generalizations when experimenting with elimination criteria. For example, a census parameter might land on the boundary of a box, so all boxes that contain this point on its boundary must be considered. We must also remark that a given census manifold may give rise to multiple points in the parameter space as the manifolds might contains several valid markings (in TUBES or CUSPS, there might be multiple self-tangencies of a tube or a cusp, giving different markings). In addition, several representatives of the same equivalence class may exist in the parameter space, especially on its boundary.

Another useful approach is the creation of visualization tools. There are two important types to consider. First, given a marking $(S, U, \mathfrak{Dicta})$, it is a good idea to write code that draws the orbit of U under Γ_S alongside any other relevant information. For Cusps, as studied in [**19**], examples of such code can be found at [**42**], where the cusp diagram of a given marking is drawn from its boxcode. Studying such a visual diagram often helps deal with troublesome boxcodes where the word search algorithm gets stuck, but human intervention can quickly find relator or killer words. Another visualization tool that can be useful is the parameter space itself. Plotting slices through the parameter space and looking for the location of census manifolds and troublesome boxes can give hits to the structure of the set of discrete groups in question.

It is also important to have a by-hand or an alternative methods of building confidence in the partial results as one implements the search. In the context of Kleinian groups, it is often possible to use a combination of software like `heegaard` [**9**] , `twister` [**8**], SnapPy [**15**] and Regina [**11**] to check that relator words found at census boxcodes do indeed recover the marking subgroup. This process goes from a (realizable) presentation to a Heegaard splitting, to a triangulation, and then checks isometry type in SnapPy. This pipeline has been assembled in a small program called `coover` by Haraway [**27**] with code contributions from Dunfield, Linton, Futer, Purcell, Schleimer, and Yarmola. While this tool is not a rigorous method of validation, as compared to the `verify` program, it has played an essential role in finding and correcting issues with more complex elimination criteria and sanity-checking many runs of the search algorithm.

11. Acknowledgments

The first author was partially supported by NSF grants DMS-1006553, DMS-1607374 and DMS-2003892. The forth author was partially supported as a Visiting Student Research Collaborator with DMS-1006553 and Postdoctoral Researcher with DMS-1607374. This work was partially supported by a grant from the Simons Foundation (#228084 to Robert Meyerhoff) and by a grant from the National Science Foundation (DMS-1308642 to Robert Meyerhoff). The authors also acknowledge the support and resources of the Polar Computing Cluster at the Mathematics Department and the Program in Applied and Computational Mathematics at Princeton University. The authors also thank David Futer for comments and corrections.

References

1. *IEEE standard for binary floating-point arithmetic*, ANSI/IEEE Std 754-1985 (1985), 1–20.
2. Colin C. Adams, *The noncompact hyperbolic 3-manifold of minimal volume*, Proc. Amer. Math. Soc. **100** (1987), no. 4, 601–606. MR 894423
3. Ian Agol, *Bounds on exceptional Dehn filling*, Geom. Topol. **4** (2000), 431–449. MR 1799796
4. ———, *Volume change under drilling*, Geom. Topol. **6** (2002), 905–916. MR 1943385
5. ———, *Bounds on exceptional Dehn filling II*, Geom. Topol. **14** (2010), no. 4, 1921–1940. MR 2680207
6. Ian Agol, Marc Culler, and Peter B. Shalen, *Dehn surgery, homology and hyperbolic volume*, Algebr. Geom. Topol. **6** (2006), 2297 2312. MR 2286027
7. Ian Agol, Peter A. Storm, and William P. Thurston, *Lower bounds on volumes of hyperbolic Haken 3-manifolds*, J. Amer. Math. Soc. **20** (2007), no. 4, 1053–1077, With an appendix by Nathan Dunfield. MR 2328715
8. Mark Bell, Tracy Hall, and Saul Schleimer, *Twister (computer software)*, `https://bitbucket.org/Mark_Bell/twister/`, 2008–2014, Version 2.4.1.
9. John Berge, `heegaard3`, Available at `https://github.com/3-manifolds/heegaard3`, (2021-12-23).
10. Steven Boyer, *Dehn surgery on knots*, Handbook of geometric topology, North-Holland, Amsterdam, 2002, pp. 165 218. MR 1886670
11. Benjamin A. Burton, Ryan Budney, William Pettersson, et al., *Regina: Software for low-dimensional topology*, `http://regina-normal.github.io/`, 1999–2021.
12. Chun Cao and G. Robert Meyerhoff, *The orientable cusped hyperbolic* 3*-manifolds of minimum volume*, Invent. Math. **146** (2001), no. 3, 451 478. MR 1869847
13. David Coulson, Oliver A. Goodman, Craig D. Hodgson, and Walter D. Neumann, *Computing arithmetic invariants of 3-manifolds*, Experimental Mathematics **9** (2000), no. 1, 127 – 152.
14. Thomas Crawford, *A stronger gordon conjecture and an analysis of free bicuspid manifolds with small cusps*, Ph.D. thesis, Boston College, 2018, p. 62.
15. Marc Culler, Nathan M. Dunfield, Matthias Goerner, and Jeffrey R. Weeks, *SnapPy, a computer program for studying the geometry and topology of* 3*-manifolds*, Available at `http://snappy.computop.org` (2021-12-23).
16. David Futer, David Gabai, and Andrew Yarmola, *On the Margulis constant of hyperbolic 3-manifolds*, In progress.
17. David Futer, Jessica S. Purcell, and Saul Schleimer, *Effective bilipschitz bounds on drilling and filling*, 2021.
18. David Gabai, *The Smale conjecture for hyperbolic 3-manifolds:* $\mathrm{Isom}(M^3) \simeq \mathrm{Diff}(M^3)$, J. Differential Geom. **58** (2001), no. 1, 113–149. MR 1895350
19. David Gabai, Robert Haraway, Robert Meyerhoff, Nathaniel Thurston, and Andrew Yarmola, *Hyperbolic 3-manifolds of low cusp volume*, 2021.
20. David Gabai, Marc Lackenby, and Andrew Yarmola, *Bounds on exceptional Dehn surgery*, In preparation.
21. David Gabai, G. Robert Meyerhoff, and Nathaniel Thurston, *Homotopy hyperbolic 3-manifolds are hyperbolic*, Ann. of Math. (2) **157** (2003), no. 2, 335–431. MR 1973051
22. David Gabai, Robert Meyerhoff, and Peter Milley, *Minimum volume cusped hyperbolic three-manifolds*, J. Amer. Math. Soc. **22** (2009), no. 4, 1157–1215. MR 2525782
23. David Gabai and Maria Trnkova, *Exceptional hyperbolic 3-manifolds*, Comment. Math. Helv. **90** (2015), no. 3, 703 730. MR 3420467
24. Oliver Goodman, *Snap, a computer program for studying arithmetic invariants of hyperbolic 3-manifolds*, Available at `http://www.ms.unimelb.edu.au/~snap`.
25. C. McA. Gordon, *Dehn filling: a survey*, Knot theory (Warsaw, 1995), Banach Center Publ., vol. 42, Polish Acad. Sci. Inst. Math., Warsaw, 1998, pp. 129–144. MR 1634453
26. Robert Haraway, *Manifold enumeration software of "hyperbolic 3-manifolds of low cusp volume"*, Available at `https://github.com/bobbycyiii/low-cusp-volume` (2021-12-23).
27. Robert Haraway et al., *Coover: presentations to triangulations*, Available at `https://github.com/bobbycyiii/coover` (2021-12-23).
28. Craig D. Hodgson and Steven P. Kerckhoff, *Universal bounds for hyperbolic Dehn surgery*, Ann. of Math. (2) **162** (2005), no. 1, 367–421. MR 2178964
29. ———, *The shape of hyperbolic Dehn surgery space*, Geom. Topol. **12** (2008), no. 2, 1033 1090. MR 2403805
30. Marc Lackenby and Robert Meyerhoff, *The maximal number of exceptional Dehn surgeries*, Invent. Math. **191** (2013), no. 2, 341–382. MR 3010379

31. T. Marshall and Gavin Martin, *Minimal co-volume hyperbolic lattices, ii: Simple torsion in a Kleinian group*, Annals of Mathematics **176** (2012), no. 1, 261–301.
32. Robert Meyerhoff, *The cusped hyperbolic* 3*-orbifold of minimum volume*, Bull. Amer. Math. Soc. (N.S.) **13** (1985), no. 2, 154–156. MR 799800
33. ———, *Sphere-packing and volume in hyperbolic* 3*-space*, Comment. Math. Helv. **61** (1986), no. 2, 271–278. MR 856090
34. Walter D. Neumann and Don Zagier, *Volumes of hyperbolic three-manifolds*, Topology **24** (1985), no. 3, 307 332. MR 815482
35. Grisha Perelman, *The entropy formula for the ricci flow and its geometric applications*, 2002.
36. ———, *Ricci flow with surgery on three-manifolds*, 2003.
37. William P. Thurston, *Geometry and topology of three-manifolds*, Lecture notes available at `http://library.msri.org/books/gt3m/` (2021-12-23), 1978.
38. ———, *Three-dimensional geometry and topology. Vol. 1*, Princeton Mathematical Series, vol. 35, Princeton University Press, Princeton, NJ, 1997, Edited by Silvio Levy. MR 1435975
39. Maria Trnkova and Andrew Yarmola, *On maximal systoles of low-genus hyperbolic surfaces*, In progress.
40. Jeffrey R. Weeks, *Snap*, Available at `http://www.geometrygames.org/SnapPea/index.html`.
41. Andrew Yarmola, *Verification software for "Hyperbolic 3-manifolds of low cusp volume"*, Available at `https://github.com/andrew-yarmola/verify-cusp` (2021-12-23).
42. ———, *Visualization software for "Hyperbolic 3-manifolds of low cusp volume"*, Available at `https://github.com/andrew-yarmola/low-cusp-volume` (2021-12-23).

Department of Mathematics, Fine Hall, Princeton University, Princeton, NJ 08544
E-mail address: gabai@math.princeton.edu
URL: https://www.math.princeton.edu/people/david-gabai

Department of Mathematics, Maloney Hall, Boston College, Chestnut Hill, MA 02467
E-mail address: robert.meyerhoff@bc.edu

Department of Mathematics, Fine Hall, Princeton University, Princeton, NJ 08544
E-mail address: yarmola@math.princeton.edu
URL: https://web.math.princeton.edu/~yarmola/