

# NOTES ON AGOL'S VIRTUAL HAKEN THEOREM

DANNY CALEGARI

ABSTRACT. These are notes on Agol's proof of the Virtual Haken Conjecture and Virtual Fibration Conjecture, and the various mathematical tools that feed into the proof. There is no claim to originality. These notes follow a course given at the University of Chicago in Winter 2013.

## CONTENTS

1. Waldhausen's Conjecture and the Virtual Fibration Conjecture	1
2. Special Cube Complexes	5
3. Codimension 1 subgroups	13
4. Almost geodesic surfaces	20
5. Relative hyperbolicity	25
6. $\mathcal{MVH}$ and $\mathcal{QVH}$	31
7. Proof of the VHC	34
8. Other examples of cubulated groups	39
9. Acknowledgments	43
References	43

## 1. WALDHAUSEN'S CONJECTURE AND THE VIRTUAL FIBRATION CONJECTURE

1.1. **Basic definitions.** We assume here and throughout the sequel that 3-manifolds under consideration are always *oriented* and *connected*, unless we explicitly say otherwise. A property of a space (resp. group) is said to hold *virtually* if it holds for some finite covering space (resp. finite index subgroup).

**Definition 1.1** (Irreducible). A compact 3-manifold  $M$ , possibly with boundary, is said to be *irreducible* if every embedded 2-sphere in  $M$  bounds a 3-ball in  $M$ .

By the sphere theorem [23] an orientable 3-manifold  $M$  is irreducible if and only if  $\pi_2(M)$  is trivial. A reducible 3-manifold can be decomposed in an essentially unique way as a connect sum of  $S^2 \times S^1$  factors and irreducible factors.

**Definition 1.2** (Incompressible). A properly embedded surface  $S$  in a 3-manifold  $M$  is *incompressible* if for every embedded disk  $D$  in  $M$  intersecting  $S$  in a loop, the boundary  $\partial D$  bounds a disk in  $S$ . The surface  $S$  is said to be *boundary incompressible* if for every embedded disk  $D$  in  $M$  intersecting  $S$  only in a proper arc (with the rest of  $\partial D$  on  $\partial M$ ), the arc  $\partial D \cap S$  is (homotopically) inessential in  $S$ .

Surfaces which are incompressible and boundary incompressible are also said to be *essential*. By Dehn's Lemma and the Loop Theorem [23] an incompressible surface is  $\pi_1$ -injective.

A surface  $S$  properly embedded in  $M$  is boundary incompressible if and only if the double of  $S$  is incompressible in the double of  $M$ . One can further generalize (by doubling and induction) to surfaces with corners in 3-manifolds with corners; one says such surfaces (with corners) are essential.

**Definition 1.3** (Haken). A compact 3-manifold is said to be *Haken* if it is irreducible, and if it contains an essential properly embedded subsurface.

*Example 1.4* (Manifold with homology). Suppose  $M$  is irreducible, and that the rank of  $H_1(M)$  is positive. By duality there are nontrivial integral classes in  $H^1(M)$  and in  $H_2(M, \partial M)$ . A properly embedded surface  $S$  representing a homology class  $\alpha \in H_2(M, \partial M)$  is *Thurston norm minimizing* if it has no spherical or disk components, and minimizes  $-\chi(S)$  amongst all surfaces in the homology class  $\alpha$ . Norm minimizing surfaces exist, since  $-\chi(S)$  takes non-negative integer values. Norm minimizing surfaces are essential, since otherwise they can be compressed or boundary compressed, reducing  $-\chi(S)$ . Thus  $M$  is Haken.

*Example 1.5* (Action on a tree). Suppose  $M$  is irreducible, and that  $\pi_1(M)$  acts minimally and without inversions on a nontrivial (simplicial) tree  $T$ . By Bass–Serre theory (see [39]) there is such an action if and only if  $\pi_1(M)$  can be expressed as a nontrivial amalgam  $A *_B C$  or HNN extension  $A *_B$ . The quotient of the action is a graph  $\Gamma$ , and we can build an equivariant map from  $\tilde{M}$  to  $T$  covering a map from  $M$  to a graph  $\Gamma$ . The preimage of a regular value is an embedded surface in  $M$  which may be compressed until it is essential. Since  $M$  is irreducible, this essential surface has no sphere components, and therefore  $M$  is Haken.

As remarked earlier, the Loop Theorem implies that an incompressible embedded surface in a 3-manifold is  $\pi_1$ -injective. One may cut along such a surface and express  $\pi_1(M)$  as a nontrivial amalgam or HNN extension over a surface subgroup. Conversely, Example 1.5 shows that any expression of  $\pi_1(M)$  as a nontrivial amalgam or HNN extension gives rise to an incompressible surface. Thus for irreducible 3-manifolds, Hakenness is *equivalent* to the purely algebraic property of the existence of a nontrivial splitting of  $\pi_1(M)$  as  $A *_C B$  or  $A *_B$ .

*Example 1.6* (Manifold with boundary). Let  $M$  be irreducible, and let  $\partial M$  be nonempty. Then either  $M$  is a ball, or  $\partial M$  has no spherical components. The intersection pairing on homology defines a natural symplectic structure on  $H_1(\partial M)$ , and the kernel of the inclusion map  $H_1(\partial M) \rightarrow H_1(M)$  is a Lagrangian subspace. One way to see that the intersection pairing on  $H_1(\partial M)$  vanishes on the kernel of the inclusion is to represent two elements of the kernel as embedded 1-manifolds  $\alpha, \beta$  in  $\partial M$  bounding embedded surfaces in  $M$ . If we move these surfaces into general position, they intersect in proper 1-manifolds; each proper arc bounds a pair of points of  $\alpha \cap \beta$  with opposite sign, certifying that  $[\alpha] \cap [\beta] = 0$ .

It follows that the rank of  $H_1(M)$  is positive, and  $M$  is Haken as in Example 1.4. Note that the essential surface produced in this way represents a nontrivial class in  $H_2(M, \partial M)$  and is therefore not boundary parallel. Hence the result of cutting along this surface definitely reduces the complexity of  $M$ .

*Example 1.7* (Seifert fibered spaces). A 3-manifold is *Seifert fibered* if it is foliated by circles. Such a manifold admits the structure of an  $S^1$  bundle over a 2-dimensional *orbifold*  $O$ . A closed Seifert fibered space is *small* if the orbifold  $O$  is a sphere with at most three marked points. Any other closed orbifold contains an essential embedded loop, and the union of the circle fibers over such a loop is an essential torus. Hence every Seifert fibered space is either small, or Haken. An aspherical Seifert fibered space has a base orbifold  $O$  which is either Euclidean or hyperbolic; such an orbifold has a finite cover of positive genus. It follows that every aspherical Seifert fibered space is virtually Haken.

Haken manifolds are also sometimes said to be *sufficiently large*. A closed Haken manifold can be cut along an essential subsurface to produce a new Haken manifold (with boundary). Subsequent cuts produce irreducible manifolds always with boundary; such manifolds are necessarily Haken, and the process can be continued. Haken showed that after finitely many steps the result can be taken to be a collection of manifolds of the form  $S \times [0, 1]$  where  $S$  is a surface, possibly with boundary. This can be proved by arguing that the essential surfaces along which one cuts must all be of a particularly simple form in any triangulation of the original manifold (they can be taken to be *normal surfaces*), and therefore once there are sufficiently many of them (the number can be determined explicitly from the rank of homology and the number of simplices) some complementary piece must be a trivial  $I$ -bundle (i.e. a product).

The process of cutting a Haken manifold successively along essential surfaces is called a *hierarchy*. The resulting pieces at the end come together with a structure called a *boundary pattern* which indicates how the previous terms in the hierarchy may be recovered by gluing.

Wolfgang Haken used essential surfaces to solve the homeomorphism problem for the (Haken) 3-manifolds that contain them. This includes as a very important special case manifolds obtained as complements of (open tubular neighborhoods of) knots in the 3-sphere; hence Haken's methods solve the knot recognition problem — the problem of deciding when two different diagrams determine the same knot type.

The structure of a hierarchy opens the possibility of inductive proofs of theorems about Haken 3-manifolds; Thurston's famous geometrization theorem for Haken 3-manifolds is such an example.

The Virtual Haken Conjecture, formulated by Waldhausen in 1968 [42] is the following:

**Conjecture 1.8** (Waldhausen's Virtual Haken Conjecture). *Every aspherical closed 3-manifold has a finite cover which is Haken.*

By the sphere theorem, every closed 3-manifold admits a connect sum decomposition into factors with finite fundamental group,  $S^2 \times S^1$  factors, and aspherical factors. As part of his proof of Thurston's Geometrization Conjecture, Perelman [32, 33] proved the following:

**Theorem 1.9** (Perelman's Hyperbolization Theorem). *Let  $M$  be an aspherical closed 3-manifold which does not contain an essential torus. Then either  $M$  is a small Seifert fibered space, or  $M$  is hyperbolic.*

Consequently, Perelman's theorem reduces the Virtual Haken Conjecture to the case where the 3-manifold is closed and hyperbolic.

## 1.2. Surface bundles over the circle.

**Definition 1.10.** Let  $S$  be a surface, and  $\varphi : S \rightarrow S$  a homeomorphism. The *mapping torus* of  $\varphi$  is the manifold  $S *_\varphi := S \times [0, 1] / (s, 1) \sim (\varphi(s), 0)$ . A 3-manifold is said to *fiber over the circle*, to admit the structure of a *fibration*, or to be a *surface bundle over the circle* if it is homeomorphic to a manifold of the form  $S *_\varphi$ . The surfaces  $S \times t$  are called *fibers*.

A surface bundle is irreducible if the surface is not a sphere, and the fibers are all essential. So aspherical surface bundles are Haken.

*Example 1.11* (Positive braid). Let  $L$  be a link in  $S^3$  arranged as a braid in such a way that all crossings in the projection are *positive* (i.e. strands cross over from left to right as one moves down). Then  $S^3 - L$  is a surface bundle over the circle.

Recall from Example 1.4 the definition of a Thurston norm minimizing surface  $S$  in a homology class in  $H_2(M, \partial M; \mathbb{Z})$ . The quantity  $-\chi(S)$  is called the *Thurston norm* of the class. It extends to a pseudo-norm on  $H_2(M, \partial M; \mathbb{R})$  taking integer values on integer classes, and has a unit ball which is a finite sided rational polyhedron.

When  $M$  is hyperbolic, the Thurston norm is a genuine norm. If  $M = S *_\varphi$  is a surface bundle, the surfaces  $S$  are norm-minimizing. Thurston proved the following fact:

**Theorem 1.12** (Thurston norm ball). *Let  $M$  be a closed 3-manifold. The set of homology classes representing fibers of fibrations over the circle are precisely the integral classes which projectively intersect the interiors of certain top-dimensional faces of the unit ball in the Thurston norm.*

The faces described in Theorem 1.12 are called *fibered faces*. Thurston posed the following as a question in [41]:

**Conjecture 1.13** (Thurston's Virtual Fibration Conjecture). *Let  $M$  be a closed hyperbolic 3-manifold. Then  $M$  has a finite cover which fibers over the circle.*

The restriction to hyperbolic manifolds is important because of the following examples, due to Neumann [29]:

*Example 1.14* (Graph manifold). Let  $M$  be a closed oriented 3-manifold obtained by pasting two Seifert manifolds  $M_1, M_2$ , each having a torus as its boundary, along these tori. Suppose that neither half is the total space of the orientable circle bundle over the Möbius band. To each of  $M_i$  is associated a pair of numerical invariants  $e_i, \chi_i$  which are respectively the Euler number of the fibration and the orbifold Euler characteristic of the base. Let  $p$  be the intersection number within the gluing torus of the fibers of the two pieces of  $M$ . Then  $M$  is virtually fibered if and only if  $0 < p^2 e_1 e_2 \leq 1$  or  $e_1 = e_2 = 0$ .

This is proved by the classification of Seifert fibered spaces and their fibrations over the circle, and a careful analysis of when fibrations on covers of the two pieces can be matched up.

If  $M$  fibers over the circle with fiber  $S$ , the fundamental group  $\pi_1(M)$  fits into a short exact sequence

$$0 \rightarrow \pi_1(S) \rightarrow \pi_1(M) \rightarrow \mathbb{Z} \rightarrow 0$$

Stallings showed that  $M$  fibers over  $S^1$  if and only if  $\pi_1(M)$  fits into a short exact sequence  $A \rightarrow \pi_1(M) \rightarrow \mathbb{Z}$  where  $A$  is finitely generated.

In the sequel we aim to explain the main ingredients in Agol's proof of the Virtual Haken and Virtual Fibration Conjectures.

## 2. SPECIAL CUBE COMPLEXES

**2.1. Basic Definitions.** The material in this section is largely taken from [21], and the reader is referred to that paper for details.

**Definition 2.1** (CAT(0) metric space). A geodesic metric space  $X$  is CAT(0) if its triangles are at least as thin as triangles in Euclidean space. That is, if  $abc$  are three points in  $X$ , and  $\bar{a}\bar{b}\bar{c}$  are three points in the Euclidean plane with  $d_X(a, b) = d_{\mathbb{E}^2}(\bar{a}, \bar{b})$  and so on, then if we parameterize geodesics  $bc$  and  $\bar{b}\bar{c}$  by an arclength parameter  $t$ , we should have the inequality

$$d_X(a, bc(t)) \leq d_{\mathbb{E}^2}(\bar{a}, \bar{b}\bar{c}(t))$$

**Definition 2.2** (CAT(0) cube complex). A *cube complex* is a space obtained by gluing Euclidean cubes of edge length 1 along subcubes. A cube complex is CAT(0) if it is CAT(0) as a metric space. It is *nonpositively curved* if its universal cover is CAT(0).

We abbreviate nonpositively curved in the sequel by NPC.

The following observation is due to Gromov [17]:

**Lemma 2.3** (Flag condition). *A cube complex is NPC if and only if the link of every vertex is a flag complex (i.e. a complex in which every complete subgraph spans a simplex).*

*Proof.* The flag condition is necessary, since the boundary of an orthant in Euclidean space is not convex.

A complex is locally CAT(0) if the links of vertices are CAT(1). The links of vertices in a cube complex are spherical complexes made from right-angled spherical simplices. These vertex links themselves have vertex links which are also spherical complexes made from right-angled spherical simplices of one lower dimension, and so on. The flag condition is inherited by passing to links, so by induction on dimension we reduce to the case of a 2-dimensional complex. Hence the links are metric graphs made of segments of length  $\pi/2$ , and the flag condition is precisely the condition that there are no loops made up from fewer than 4 segments. It follows that every loop in these links have length at least  $2\pi$ , so there are no atoms of positive curvature in the 2-dimensional complexes, and they are CAT(1) if spherical or NPC if cubical.  $\square$

*Example 2.4* (Trees and graphs). Trees are the simplest (connected) examples of CAT(0) cube complexes. Graphs are NPC cube complexes.

*Example 2.5* (Surfaces). A pillowcase (the space obtained by doubling a square along its edges) gives the 2-sphere the structure of a cube complex, but it is not NPC. The 2-torus admits the structure of an NPC square complex by identifying opposite sides of a square by translation. Eight regular hyperbolic right-angled pentagons glue up to make a genus 2 surface; the dual cellulation is an NPC square complex with 10 squares. By taking covers, one sees that every surface of genus at least 1 admits the structure of an NPC square complex.

*Example 2.6* (Tree lattices). Products of CAT(0) cube complexes are CAT(0) cube complexes in an obvious way. There are many interesting examples of lattices in products  $\text{Aut}(T_1) \times \text{Aut}(T_2)$  for certain trees  $T_1, T_2$  constructed by Burger-Mozes [9]. The quotients are interesting examples of NPC square complexes.

*Example 2.7* (Heegaard diagrams). Aitchison-Rubinstein [5] describe the following example. Let  $M$  be a 3-manifold with a Heegaard splitting along a surface  $L$ . Choose families of loops  $C_i$  and  $C'_i$  on  $L$  bounding a complete set of compressing disks on either side. Suppose all complementary regions of  $L - \cup C_i - \cup C'_i$  are hexagons, and every curve  $C'_j$  meets  $\cup C_i$  in exactly four points and conversely. Then we may decompose  $M$  along  $L$  and the disks bounded by the  $C_i$  and  $C'_i$  into balls whose boundaries are tessellated by squares and hexagons; such balls can be further subdivided into cubes in such a way that the resulting cube complex is NPC.

*Example 2.8* (Dehn filling). Tao Li [27] showed that if  $M$  is a 3-manifold with boundary an incompressible torus which contains no non-peripheral embedded closed essential surface, then only finitely many Dehn fillings on  $M$  give rise to a closed 3-manifold which admits the structure of an NPC cube complex. This is despite the fact that if  $M$  as above is hyperbolic, all but finitely many Dehn fillings give rise to a hyperbolic 3-manifold.

We think of each  $n$ -cube in a cube complex as a copy of  $[-\frac{1}{2}, \frac{1}{2}]^n$ .

**Definition 2.9** (Hyperplane). A *midcube* in  $[-\frac{1}{2}, \frac{1}{2}]^n$  is the intersection with a coordinate plane  $x_i = 0$ . A midcube divides a cube into two equal pieces, and is parallel to a pair of opposite top-dimensional faces. An edge is *dual* to a midcube if it intersects it.

If  $X$  is a cube complex, form a new cube complex  $Y$  with one cube for each midcube of  $X$ , and with one vertex for each midpoints of a corresponding dual edge. The components of  $Y$  are the *hyperplanes* of  $X$ .

Each edge of  $X$  is dual to a unique hyperplane; we denote the hyperplane dual to an edge  $a$  as  $H(a)$ .

There is an equivalence relation on (oriented) edges of  $X$ , where two edges are equivalent if they are on opposite sides of some square. The equivalence classes of edges are called *walls*. There is an obvious correspondence between hyperplanes and walls. Note that if  $X$  is CAT(0), then hyperplanes are totally geodesic and therefore *embedded*.

**Definition 2.10** (Osculation). Two hyperplanes  $H_1, H_2$  (possibly equal) are said to *osculate* at a vertex  $v$  which is the endpoint of (oriented) edges  $e_1, e_2$ , not both adjacent in the same square, and dual to  $H_1$  and  $H_2$  respectively.

*Example 2.11*. A hyperplane can self-osculate in two different ways. *Direct* self-osculatation is when two oriented edges in the same equivalence class meet at the same vertex. *Indirect* self-osculatation is when an edge  $e$  meets an edge in the same equivalence class but with the opposite orientation.

Direct and indirect self-osculatation is illustrated in Figure 1. The hyperplane is in red, and the dual equivalence class of oriented edges is in blue.

*Interosculatation* between different hyperplanes occurs when the two hyperplanes both intersect and osculate; this can occur in only one way; see Figure 2. The two hyperplanes

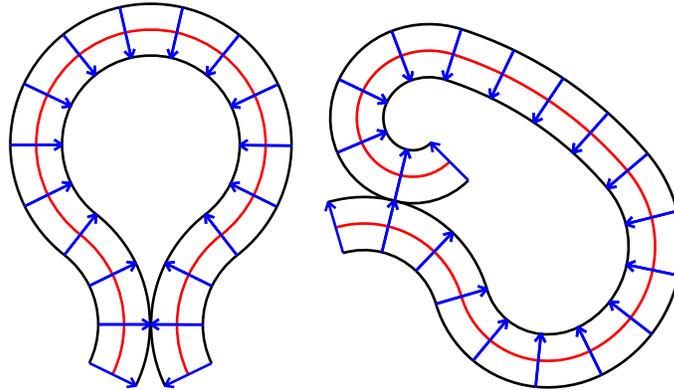


FIGURE 1. Direct and indirect self-osculation

are in red and yellow, and their dual equivalence classes of edges are in blue and green respectively.

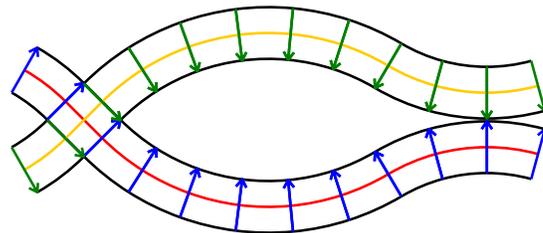


FIGURE 2. Interosculation between different hyperplanes

The following class of cube complexes were introduced by Haglund-Wise [21]:

**Definition 2.12** (Special cube complex). A cube complex is *special* if the following conditions hold:

- (1) every hyperplane is embedded and two-sided;
- (2) no hyperplane directly self-oscultates; and
- (3) no two hyperplanes interosculate (i.e. both intersect and osculate).

The property of being a special cube complex is preserved under taking products or covers. It is also preserved under taking locally convex subcomplexes.

If  $X$  is a cube complex, the simplicial length mod 2 of an edge in the 1-skeleton is a homotopy invariant rel. endpoints, since every homotopy is a composition of elementary homotopies across a square. Hence there is a homomorphism  $\ell : \pi_1(X) \rightarrow \mathbb{Z}/2\mathbb{Z}$  which gives the length of a simplicial representative mod 2. After passing to a 2-fold cover if necessary we can assume this map is trivial, and therefore vertices can be 2-colored.

Similarly, if  $H$  is a hyperplane, the parity of the number of intersections of a simplicial path with  $H$  is invariant under elementary homotopies, and there is a homomorphism

$\ell_H : \pi_1(X) \rightarrow \mathbb{Z}/2\mathbb{Z}$ . Suppose  $H$  is embedded. If  $H$  is one-sided, or if  $H$  indirectly self-oscillates, then we can construct a simplicial loop  $\gamma$  which intersects  $H$  exactly once. So if  $\hat{X}$  is the 2-fold cover of  $X$  associated to the kernel of  $\ell_H$ , the hyperplanes in the preimage of  $H$  are embedded, 2-sided, and do not indirectly self-oscillate. If  $X$  has finitely many hyperplanes  $H_i$  (for instance, if  $X$  is finite), we get a homomorphism  $\ell_i : \pi_1(X) \rightarrow \mathbb{Z}/2\mathbb{Z}$  for each  $i$ , and a finite cover  $\hat{X}$  corresponding to the intersection of the kernels of all the  $\ell_i$ . If every  $H_i$  is embedded, then this cover has the property that *every* hyperplane is embedded, 2-sided, and has no indirect self-oscillation.

## 2.2. Right-angled Artin groups.

**Definition 2.13.** Let  $\Gamma$  be a graph. The *right-angled Artin group*  $A_\Gamma$  associated to  $\Gamma$  is the group with generators given by the vertices of  $\Gamma$ , and with relations only that two generators commute if and only if the corresponding vertices are joined by an edge in  $\Gamma$ .

We abbreviate right-angled Artin group in the sequel by RAAG.

**Lemma 2.14** (Special cube complex from RAAG). *Let  $A_\Gamma$  be a RAAG associated to  $\Gamma$ . There is an NPC special cube complex  $Y_\Gamma$  whose fundamental group is  $A_\Gamma$ .*

*Proof.* A *cubical  $n$ -torus* is the torus obtained from a single  $n$ -cube by identifying opposite faces by translation. It has the structure of a cube complex in a natural way. We build a complex  $Y_\Gamma$  by taking one cubical  $n$ -torus for every complete graph on  $n$  vertices appearing in  $\Gamma$  and identifying them along subcubes. This is a cube complex with one vertex. Let  $F$  be the flag complex of  $\Gamma$ , i.e. the simplicial complex with one simplex for every complete graph on  $n$  vertices in  $\Gamma$ . The link of the vertex of  $Y_\Gamma$  is a complex  $\Sigma F$  obtained from  $F$  by the following functorial construction: a simplex of  $\Sigma F$  is a pair consisting of a simplex of  $F$  together with a function from its vertices to  $\{-1, 1\}$ . Evidently  $\Sigma F$  is a flag complex if  $F$  is (which it is, by definition). This shows that  $Y_\Gamma$  is NPC, by Lemma 2.3.

To see that  $Y_\Gamma$  is special, first observe that the complex  $Y_\Gamma$  is a subcomplex of a single cubical  $n$ -torus, and its hyperplanes are subcomplexes of the hyperplanes in the  $n$ -torus. It follows that hyperplanes are embedded and two-sided. There is a single dual edge to each hyperplane, which cannot end on the same vertex in two different ways; thus no hyperplane directly self-oscillates. Finally, if two hyperplanes intersect, their associated dual edges are adjacent in some square (any square containing a point of intersection of the hyperplanes in its center), so they do not oscillate.  $\square$

The space  $Y_\Gamma$  is sometimes called the *Salvetti complex* of  $\Gamma$ .

A cube complex  $X$  with embedded hyperplanes determines a graph  $\Gamma$  whose vertices are the hyperplanes, and whose edges are the pairs of hyperplanes that intersect. Associated to  $\Gamma$  is the RAAG  $A_\Gamma$  which in turn determines its own special cube complex  $Y_\Gamma$ .

The following is proved in [21], Lemma 3.15.

**Lemma 2.15** (Special immerses). *Let  $X$  be a special cube complex, and  $\Gamma$  the hyperplane graph. Let  $A_\Gamma$  be the RAAG associated to  $\Gamma$ , and  $Y_\Gamma$  its associated special cube complex. Then there is an immersion of  $X$  into  $Y_\Gamma$ .*

*Proof.* There is a map  $X \rightarrow Y_\Gamma$  defined as follows. Each vertex of  $X$  maps to the unique vertex of  $Y_\Gamma$ . Each (oriented) edge of  $X$  is dual to a hyperplane, which determines a

generator of  $A_\Gamma$  and thereby an edge of  $Y_\Gamma$ ; we map each edge of  $X$  to the corresponding edge of  $Y_\Gamma$ . Each  $n$ -cube of  $X$  determines a family of  $n$  distinct and intersecting hyperplanes (because hyperplanes are embedded and can therefore intersect a given cube at most once) and thereby determines a complete graph on  $n$  vertices in  $\Gamma$  and a cubical  $n$ -torus in  $Y_\Gamma$ ; we map each  $n$ -cube to the associated  $n$ -torus.

It remains to check that this map is an immersion. We examine the induced map on vertex links and check it is injective on vertices (which correspond to edges of the complexes). Since hyperplanes embed, adjacent vertices in a link are sent to distinct vertices in a link. Since hyperplanes do not directly self-osculate, nonadjacent vertices in a link are sent to distinct vertices in a link. This completes the proof.  $\square$

**Definition 2.16** (Word quasiconvex). Let  $G$  be a finitely generated group with a generating set  $S$ . A subgroup  $H$  is word quasiconvex (with respect to  $S$ ) if any geodesic path in the Cayley graph  $C_S(G)$  with endpoints in  $H$  stays within a bounded distance of  $H$ .

The following is the main theorem of [21], and paraphrases Theorem 4.2.

**Theorem 2.17** (NPC special group quasiconvex in RAAG). *Let  $X$  be an NPC special cube complex. Then the map from  $X$  to  $Y_\Gamma$  induces an injection of  $\pi_1(X)$  into the RAAG  $A_\Gamma$ , and the image is word quasiconvex with respect to the standard generating set.*

*Proof.* The map  $X \rightarrow Y_\Gamma$  is an immersion, by Lemma 2.15. To see that it is locally convex it suffices to show that nonadjacent vertices in a link of a vertex of  $X$  map to nonadjacent vertices in a link of a vertex of  $Y_\Gamma$ . Nonadjacent vertices in a link of a vertex of  $X$  correspond to a pair of osculating hyperplanes. By the definition of special, these hyperplanes must be disjoint, so the corresponding edges in  $Y_\Gamma$  do not span a square. It follows that the corresponding points in the link of the vertex of  $Y_\Gamma$  are not adjacent.

A locally convex connected subset of a CAT(0) space is globally convex, so the immersion of  $X$  to  $Y_\Gamma$  lifts to an embedding of universal covers  $\tilde{X} \rightarrow \tilde{Y}_\Gamma$  with convex image. In particular, the map on fundamental groups is injective, and the image is a word quasiconvex subgroup.  $\square$

### 2.3. Separation properties of RAAGs.

**Definition 2.18** (Residually finite). A group  $G$  is *residually finite* if the intersection of all finite index subgroups of  $G$  consists only of the identity. Equivalently, if for each nontrivial element  $g \in G$  there is a finite index subgroup  $H$  of  $G$  that does not contain  $g$ .

*Example 2.19* (Linear groups). Linear groups are residually finite. This is known as *Selberg's Lemma*. This is proved by arithmetic, and ultimately rests on the fact that every field which is finitely generated as a ring is finite.

**Definition 2.20** (Subgroup separable). A subgroup  $F$  of a group  $G$  is *separable* if for every nontrivial element  $g \in G - F$  there is a finite index subgroup  $H$  of  $G$  that contains  $F$  but does not contain  $g$ . A group  $G$  is LERF (locally extended residually finite) if every finitely generated subgroup is separable.

*Example 2.21* (Free groups). Stallings showed that free groups are LERF in the following way. Let  $F$  be a free group, and let  $X$  be a rose; i.e. a graph with a single vertex, and

one edge for each generator in a free generating set for  $F$ . Then  $F = \pi_1(X)$  (in fact, a free group is a very special kind of RAAG, one associated to a graph with no edges, and  $X$  is the associated Salvetti complex). Let  $G$  be a finitely generated subgroup of  $F$ , and let  $X_G$  be the covering space of  $X$  associated to  $G$ . Then  $X_G$  contains a compact core  $Y$  which is a maximal compact subgraph without 1-valent vertices. The graph  $Y$  immerses into  $X$ , and it is straightforward to add finitely many edges to form a new graph  $Y'$ , which still immerses into  $X$ , in such a way that the immersion is a covering. Then  $\pi_1(Y')$  is a finite index subgroup  $F'$  which retracts onto  $G$ ; in particular,  $G$  is separable.

*Example 2.22* (Surface groups). Scott [38] showed that fundamental groups of surfaces are LERF. He proved this in the following way. First he showed ([38] Lemma 1.4) that separability of a subgroup  $F$  of the fundamental group of a compact space  $X$  can be characterized by the following property: for any compact subset  $C$  of the covering space  $X_F$  associated to  $F$ , there is a finite covering  $X'$  of  $X$  such that  $C$  embeds in  $X'$ .

Now, any hyperbolic surface  $S$  can be decomposed into right-angled hyperbolic pentagons. Any compact  $C$  in a cover  $S_F$  can be engulfed in a convex union of right-angled pentagons. The group generated by reflections in the boundary of this union has a finite-index subgroup  $H$  which is finite index in  $G$  and has the property that  $C$  embeds into the cover  $S_H$ .

One can relate Stallings' method to Scott's method by thinking of a rose  $X$  as a double cover of a tree  $Z$  with mirrors on the 1-valent vertices. A subgraph  $Y$  of the cover  $X_G$  can be extended by adding mirrored half-edges at each vertex where the map is not locally onto. This produces a graph with mirrors  $Y'$  which finitely (orbifold) covers  $Z$ .

*Example 2.23* (Bianchi groups). Let  $d$  be a square-free positive integer, and let  $\mathcal{O}_d$  be the ring of integers in  $\mathbb{Q}(\sqrt{-d})$ . The *Bianchi groups* are groups of the form  $\mathrm{PSL}(2, \mathcal{O}_d)$  for various  $d$ . Agol-Long-Reid [1] showed by a number theoretic argument (essentially an application of the classification theorem for quadratic forms over  $\mathbb{Q}$ ) that the Bianchi groups are conjugate into a certain group  $G$  of isometries of  $\mathbb{H}^6$  in such a way that they stabilize a totally geodesic copy of  $\mathbb{H}^3$  in  $\mathbb{H}^6$ . The group  $G$  is the group generated by reflections in a right-angled semi-ideal polyhedron  $Q$  made from 51840 copies of a certain semi-ideal simplex. Geometrically finite subgroups of  $\mathcal{O}_d$  are also geometrically finite in  $G$ , and can be separated by essentially the same engulfing argument as Scott's argument in Example 2.22.

Any tessellation of (spherical, Euclidean, hyperbolic) space by (compact) all right-angled polyhedra is dual to a cubulation. One can think of the facets of the tessellating polyhedra as pieces of midcubes. Under suitable conditions this cubulation will be CAT(0). Compare with Example 2.5.

*Example 2.24* (Non-LERF cubed 3-manifolds). Matsumoto [28] gave examples of NPC cubed 3-manifolds whose fundamental groups are not LERF. This is shown by exhibiting a specific subgroup  $K$  which is not itself LERF. The subgroup is

$$K = \langle t, x, y \mid x^t = xy, y^t = y \rangle$$

where  $z^t := t^{-1}zt$  for any  $z$ ; this group was shown to be not LERF by Burns-Karrass-Solitar [10].

As an explicit example, first take the product  $M' := (T - 2 \text{ disks}) \times S^1$  where  $T$  is a torus. The surface  $(T - 2 \text{ disks})$  can be tiled by 13 squares as in Figure 3 and  $M'$  can be tiled with 52 cubes as a product, by subdividing the  $S^1$  factor into 4 intervals. The

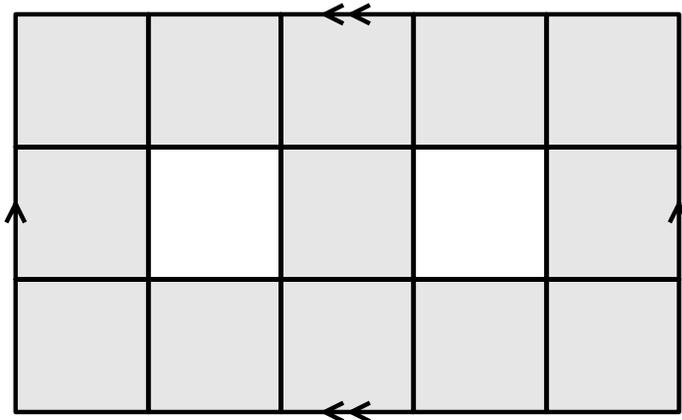


FIGURE 3. A torus with two holes tiled by 13 squares

boundary of  $M'$  consists of two square tori, each tiled by 16 squares. Gluing the boundary components by a  $\pi/2$  twist gives rise to an NPC cubulation of a closed 3-manifold  $M$ .

Let  $\theta$  be a graph in  $(T - 2 \text{ disks})$  consisting of the two boundary loops together with a proper embedded arc from one to the other. The product  $\theta \times S^1$  in  $M'$  is the union of two tori and an annulus. In  $M'$  the two tori are glued by a twist, and the result is a single torus with an annulus glued on one side to a meridian and on the other side to a longitude. The fundamental group of this object is evidently equal to  $K$ , and exhibits  $K$  as a subgroup of  $\pi_1(M)$ . Since  $K$  is not LERF, neither is  $\pi_1(M)$ .

The following theorem is proved by Haglund [20]:

**Theorem 2.25** (Quasiconvex subgroups are separable). *Let  $A_\Gamma$  be a RAAG. Then every word quasiconvex subgroup (with respect to the standard generating set) is separable.*

*Proof.* The proof is actually a consequence of the stronger statement that any word quasiconvex subgroup  $H$  of a right-angled Coxeter group  $C_\Gamma$  is a virtual retract. The proof follows Scott's proof for surface groups (see Example 2.22) closely. The Coxeter group  $C_\Gamma$  acts discretely and cocompactly on a CAT(0) cube complex  $\tilde{X}$  called the *Davis complex*, whose quotient  $X := \tilde{X}/C_\Gamma$  is analogous to the Salvetti complex for RAAGs. Let  $Hx$  be the orbit of some vertex in  $\tilde{X}$ ; this orbit is combinatorially quasiconvex, with respect to the natural combinatorial metric on the 1-skeleton, by the definition of word quasiconvexity. Then the orbit  $Hx$  is contained in a canonical convex subcomplex  $Y$  which is the intersection of all convex subcomplexes containing  $Hx$ . The main technical point is to show that  $Y$  is contained in a bounded neighborhood of  $Hx$ , and therefore the quotient  $Y/H$  is a compact convex subset of  $\tilde{X}/H$ . By adding mirrors to the faces of  $Y/H$  we obtain an orbifold  $Z$  whose orbifold fundamental group on the one hand is finite index in  $C_\Gamma$ , and on the other hand is the semidirect product of  $H$  and the group generated by reflections

in the faces of  $Y/H$ . It follows that this group retracts onto  $H$ ; i.e.  $H$  is a virtual retract, and consequently separable (but the separability can already be seen by the fact that  $Y/H$  “embeds” into the orbifold  $Z$  which is a finite cover of  $X$ ).

The standard embedding of  $\mathbb{Z}$  in the infinite dihedral group  $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$  extends in a natural way to an embedding of any RAAG into a right-angled Coxeter group. The RAAG acts convex cocompactly on the associated Davis complex; this observation is due to Davis-Januszkiewicz [14]. This proves the theorem.  $\square$

A more combinatorial argument, proving essentially the same thing, is the *canonical completion and retraction*, developed in § 6 of [21], especially Proposition 6.5 and Corollary 6.7.

**Theorem 2.26** (Haglund-Wise, Canonical completion and retraction). *Let  $X$  be a special cube complex, and let  $Y$  be a compact NPC cube complex that isometrically immerses in  $X$ . Then there is a finite cover  $\hat{X}$  of  $X$  such that  $Y$  embeds into  $\hat{X}$ , lifting the immersion to  $X$ . Moreover,  $\hat{X}$  retracts to (the image of)  $Y$ .*

*Proof.* We explain the idea of the construction. First, let’s describe the 1-skeleton of  $\hat{X}$ . The vertices consist of the product  $X^0 \times Y^0$ . There is an edge from  $(x, y)$  to  $(x', y')$  in  $\hat{X}$  in either of the following two cases:

- (1) if  $(x, x')$  and  $(y, y')$  both cross the same hyperplane in  $X$ ; or
- (2) if  $y = y'$ , and no edge of  $Y$  crosses the same hyperplane in  $X$  as  $(x, x')$ .

Note that since  $Y \rightarrow X$  is an immersion, and  $X$  is special, there can’t be two edges  $(y, y')$  and  $(y, y'')$  which both cross the same hyperplane.

One way to understand this construction is as follows. As in Theorem 2.17, there is a canonical map from  $X$  to a Salvetti complex  $Z$  whose edges are exactly the equivalence classes of edges in  $X$ . The map from  $Y$  to  $X$  can be composed  $X \rightarrow Z$ , so that we have maps from  $X$  and  $Y$  to  $Z$ . Restrict these maps to the 1-skeleta, and produce the fiber product; these are the edges of type (1) above. The result is a graph mapping both to the 1-skeleta of  $X$  and  $Y$ , but it is not yet a covering space; adding the edges of type (2) makes it a covering space (at least at the level of 1-skeleta).

Now one checks that for every square  $S$  in  $X$  and every vertex  $s$  of  $S$  and vertex  $s'$  of  $\hat{X}^1$  covering  $s$ , there is a unique lift  $S'$  of  $S$  through  $s'$ . This depends on the defining properties of a special cube complex in a straightforward way. Once one has built  $\hat{X}^2$ , the completion to the higher dimension skeleta is uniquely defined by requiring the map to be a covering.  $\square$

The following is a combination of Theorem 1.3 and Theorem 1.4 in [21]:

**Theorem 2.27** (Haglund-Wise, Quasiconvex separable is virtually special). *Let  $X$  be an NPC cube complex whose fundamental group is hyperbolic. Then  $X$  is virtually special if and only if every quasiconvex subgroup of  $\pi_1(X)$  is separable.*

*Proof.* Suppose  $X$  is virtually special. Pass to a finite special cover. By Theorem 2.17 the group  $\pi_1(X)$  is isomorphic to a word quasiconvex subgroup of a RAAG  $A_\Gamma$ . Quasiconvex subgroups of  $\pi_1(X)$  are also word quasiconvex in  $A_\Gamma$ , and are therefore separable.

Conversely, suppose every quasiconvex subgroup is separable. Then we can pass to a finite cover where all the hyperplanes are two-sided and embedded. Furthermore, since

hyperplane subgroups are quasiconvex, any self- or inter-osculation can be busted in a finite cover, again using separability.  $\square$

For complexes whose fundamental groups are also fundamental groups of hyperbolic 3-manifolds, there is a stronger conclusion:

**Corollary 2.28.** *Let  $X$  be an NPC special cube complex whose fundamental group is equal to the fundamental group of a closed hyperbolic 3-manifold. Then  $\pi_1(X)$  is LERF.*

*Proof.* By the tameness theorem [2, 11], every finitely generated subgroup of a Kleinian group is either geometrically finite or is virtually the fiber of a fibration over the circle. In the first case it is quasiconvex and therefore separable by the previous corollary; in the second case it is a virtual retract and therefore separable.  $\square$

### 3. CODIMENSION 1 SUBGROUPS

**3.1. Relative ends.** Hopf introduced the notion of the number of ends of a group.

**Definition 3.1.** If  $G$  is a finitely generated group, the *ends* of  $G$  are the ends of a Cayley graph of  $G$  with respect to a finite generating set. The number of ends is denoted  $e(G)$ .

Stallings [40] pioneered the study of the ends of groups, and developed the following picture:

- (1)  $e(G)$  is one of 0, 1, 2 or  $\infty$ ;
- (2)  $e(G) = 0$  if and only if  $G$  is finite;
- (3)  $e(G) = 2$  if and only if  $G$  contains  $\mathbb{Z}$  with finite index;
- (4) (**Stallings**, [40]).  $e(G) = \infty$  if and only if  $G$  splits nontrivially as a graph of groups with finite edge groups.

Houghton introduced the notion of the number of ends of a group relative to a subgroup, and this idea was further developed by Scott [37].

**Definition 3.2.** Let  $G$  be a finitely generated group and let  $H$  be a subgroup. The *ends of  $G$  relative to  $H$*  are the ends of the quotient of the Cayley graph of  $G$  by  $H$ . The number of relative ends is denoted  $e(G, H)$ .

A subgroup  $H$  is said to be *codimension 1* if  $e(G, H) > 1$ .

Note that  $e(G, H) = 0$  if and only if  $H$  is finite index in  $G$ . If  $G$  splits nontrivially over  $H$ , then  $e(G, H) > 1$ . However, the converse is not true. The problem is that  $H$  might not be separable in  $G$  (Scott [37] uses the terminology  *$H$ -residually finite* to mean that  $H$  is subgroup separable).

*Example 3.3* (Scott, [37]). Suppose  $G = A * H$ . Then either  $e(G, H) = \infty$  or both  $A$  and  $H$  have order two and  $e(G, H) = 1$ . This can be seen by considering the action of  $H$  (a vertex stabilizer) on the tree associated to the splitting; distinct  $H$ -orbits of ends of the tree correspond to distinct ends of  $G$  relative to  $H$ . On the other hand,  $G$  splits over  $H$  if and only if  $A$  is a nontrivial free product (in which case, if  $A = A_1 * A_2$ , the group  $G$  splits as  $G = (A_1 * H) *_H (H * A_2)$ ), or  $H$  is infinite cyclic. For, if  $G = X *_H Y$ , since no conjugate of  $A$  can meet  $H$ , we can freely decompose  $A$  into its intersection with conjugates of  $X$  and  $Y$ . On the other hand, let  $A$  and  $H$  be finitely generated infinite simple groups (a

famous example is Higman's group  $\langle a, b, c, d \mid b^a = b^2, c^b = c^2, d^c = d^2, a^d = a^2 \rangle$ , and let  $G = A * H$ . Then  $A$  is indecomposable and not infinite cyclic, so  $G$  can't split over  $H$  even though  $e(G, H) = \infty$ .

On the other hand, Scott showed that if  $G$  is finitely generated and  $H$  is a finitely generated subgroup which is separable, then  $e(G, H) \geq 2$  if and only if  $G$  *virtually* splits over  $H$ .

However, if we want to use codimension 1 subgroups to prove properties like subgroup separability, we must look at splittings of  $G$  itself (rather than finite index subgroups of it). Sageev [36] developed this idea further, especially when  $G$  is hyperbolic and  $H$  is quasiconvex.

**Definition 3.4.** Let  $H$  be a subgroup of  $G$ . We say  $H$  is *associated to a splitting* of  $G$  if  $G$  splits over a subgroup commensurable with a subgroup of  $H$ .

Sageev proved that when  $G$  is hyperbolic and  $H$  is quasiconvex of codimension 1 (i.e. so that  $e(G, H) \geq 2$ ) then either  $H$  is associated to a splitting of  $G$ , or else  $H$  itself contains a codimension 1 group  $H'$ , which can actually be taken to be the intersection of  $H$  with one of its conjugates. An important and motivating example is when  $G$  is the fundamental group of a closed hyperbolic surface, and  $H$  is the cyclic group associated to an immersed geodesic. In this case there is a splitting of  $G$  associated to  $H$  if and only if the geodesic is *virtually embedded*; otherwise it intersects itself nontrivially, and the intersections are (trivial!) subgroups of codimension 1 in  $H$ .

A quasiconvex subgroup of a hyperbolic group is itself hyperbolic, and the intersection of quasiconvex subgroups is quasiconvex, so  $H'$  is quasiconvex in  $H$ , and we can continue by induction; in particular, one can construct a descending sequence of groups  $G =: H_0, H_1, H_2, \dots, H_n$  where each  $H_{i+1}$  has codimension 1 in  $H_i$ . Sageev further proves that this procedure must *terminate*, so that  $H_n$  is associated to a splitting of  $H_{n-1}$ .

**3.2. Construction of cube complex.** Our main interest in codimension 1 subgroups is their relation to CAT(0) cube complexes. Sageev showed how to go from a collection of subgroups  $H_i$  of  $G$  with  $e(G, H_i) > 1$  to a natural action of  $G$  on a CAT(0) cube complex. We first describe the construction in the case of a single subgroup  $H$  with  $e(G, H) > 1$ .

Fix a finite generating set for  $G$ , and let  $\Gamma$  denote the Cayley graph of  $G$  with respect to this generating set, and let  $\Gamma_H$  denote  $\Gamma/H$ . By hypothesis, some compact subset  $C$  of  $\Gamma_H$  separates  $\Gamma_H$  into two unbounded sets. Let  $B$  be one of these unbounded components, and let  $A$  be the preimage of  $B$  in  $\Gamma$ . Note that  $A$  is  $H$ -invariant, and for any  $g \in G$  the symmetric difference  $A \Delta Ag$  intersects finitely many  $H$ -orbits.

Let  $\Sigma$  be a set of subsets of  $\Gamma$ , consisting of the (left) translates of  $A$  and of  $A^c$  (the complement of  $A$ ) by elements of  $G$ . This is a partially ordered set with respect to inclusion (as subsets of  $\Gamma$ ). By abuse of notation we use the letter  $A$  in what follows to mean a typical element of  $\Sigma$ , rather than the specific element  $A$  as above (when there is ambiguity the meaning should be clear from context). A *vertex*  $V$  is a subset of  $\Sigma$  satisfying that for all  $A \in \Sigma$  exactly one of  $A, A^c$  is in  $V$ ; and if  $A \in V$  and  $A \subset B$  for  $A, B \in \Sigma$  then  $B \in V$ .

*Example 3.5.* Pick  $g \in G$  and let  $V_g$  be the set of  $A \in \Sigma$  that contain  $g$ . Then  $V_g$  is a vertex.

If  $V$  and  $W$  are vertices, they span an *edge*  $(V, W)$  if and only if there is some  $A \in V$  such that  $W = (V - A) \cup A^c$ .

Lemma 3.2 in [35] is:

**Lemma 3.6** (Edge criterion). *The set of  $A \in V$  for which  $(V - A) \cup A^c$  is a vertex are precisely those  $A \in V$  which are minimal with respect to inclusion.*

*Proof.* If  $A' \subset A$  and  $A', A$  are both in  $V$ , then  $(V - A) \cup A^c$  is not a vertex, since it contains  $A'$  but not  $A$ .

Conversely, suppose  $A$  is minimal. If  $B' \subset B$  and  $B' \in (V - A) \cup A^c$  then either  $B' \in (V - A) \subset V$  in which case  $B \in V$  and therefore  $B \in V - A$  (because  $B \neq A$  since  $A$  is minimal and  $B$  isn't) or  $B' = A^c$  in which case  $B^c \subset A$  so  $B^c$  is not in  $V$  and therefore  $B$  is, and is also in  $(V - A)$ .  $\square$

This construction defines a graph; however, it is not yet the graph we want. We define the *subgraph*  $X^{(1)}$  to be the union of the components containing some  $V_g$  (as in Example 3.5) (it will shortly be proved that  $X^{(1)}$  is actually connected).

*Example 3.7.* Let  $G = \mathbb{Z}$  and  $H = \text{id}$ . Then the cosets of  $H$  are in bijection with the integers  $n$ , and the sets  $A, A^c$  are the subsets of  $\mathbb{Z}$  which consist either of all integers  $\geq$  some  $n$ , or all integers  $\leq$  some  $n$ . There is a vertex  $V_n$  for every  $n$  corresponding to the set of subsets of the form  $i \geq m$  for some  $m \leq n$  or  $i \leq m$  for some  $m \geq n$ . The  $V_i$  are the vertices of a connected graph isomorphic to  $\mathbb{R}$ . But there are also two “infinite” vertices  $V_+$  and  $V_-$ , defined by setting  $V_+$  equal to the set of subsets  $\geq$  some integer, and  $V_-$  equal to the set of subsets  $\leq$  some integer. There are no minimal elements in either  $V_{\pm}$ , so these vertices are disconnected from the rest of the graph.

Similar “infinite vertices” can be constructed whenever  $G$  is a hyperbolic group, corresponding to points in  $\partial_{\infty}G$ . These vertices will never be contained in  $X^{(1)}$ .

We can think of labeling the *oriented* edges of this graph by elements  $A \in \Sigma$  where the edge from  $V$  to  $(V - A) \cup A^c$  is labeled by  $A$ . Changing the orientation of an edge replaces the label  $A$  by  $A^c$ . Lemma 3.6 says that for each vertex  $V$ , the set of edges going out from  $V$  are labeled precisely by the minimal elements in  $V$ .

**Lemma 3.8** ( $X^{(1)}$  connected). *The graph  $X^{(1)}$  is connected.*

*Proof.* First observe that if  $V, V'$  are two vertices whose symmetric difference is a finite set  $S = \{A_1, \dots, A_n\} \subset \Sigma$ , then any  $A_i \in V$  minimal in  $S$  is actually minimal in  $V$ . For, suppose  $B \subset A_i$  and  $B \in V$ . If  $B \in V'$  then  $A_i \in V'$ , contrary to the fact that  $A_i \in V \Delta V'$ . So  $B \in V \Delta V'$ , as claimed. It follows that for any  $A_i \in V$  minimal in  $S$  the vertex  $(V - A_i) \cup A_i^c$  is adjacent to  $V$  and has a smaller symmetric difference with  $V'$ , so by induction we deduce that  $V$  and  $V'$  are in the same component.

Second, observe that for any  $g, k \in G$  the symmetric difference of  $V_g$  and  $V_k$  is finite. For, if  $fA \in V_g$  but  $fA^c \in V_k$  then  $f^{-1}g \in A$  but  $f^{-1}k \in A^c$  and therefore  $f^{-1}g \in A \Delta A g^{-1}k$  which is contained in one of finitely many  $H$ -orbits, so we can write  $f^{-1}g = he$  for some  $h \in H$  and one of finitely many  $e$ , and then  $fA = ge^{-1}h^{-1}A = ge^{-1}A$  is one of finitely many possibilities.  $\square$

Higher dimensional cubes are added to  $X^{(1)}$  inductively by a kind of “flag” construction: whenever the boundary of an  $n$ -cube appears in  $X^{(n-1)}$ , we glue in an  $n$ -cube. The union is a cube complex  $X$ . If we orient the edges of an  $n$ -cube so that parallel edges have the same orientations, then parallel edges will be labeled (as above) by the same  $A \in \Sigma$ . The orientations on edges as above determine a flow on the 1-skeleton with one source vertex and one sink vertex. Let  $A_1, \dots, A_n$  be the labels on the parallelism classes of edges in an  $n$ -cube, and let  $V$  be the source vertex. Then the  $A_i$  are all minimal in  $V$ , and no  $A_i$  is contained in or contains  $A_j^c$  for any  $i, j$ . Conversely, given a vertex  $V$  and given minimal elements  $A_1, \dots, A_n$  of  $V$  where no  $A_i$  is contained in or contains any  $A_j^c$ , there is an  $n$ -cube with source vertex  $V$  and sink vertex  $(V - \cup A_i) \cup (\cup A_i^c)$ . See Figure 4 for a 3-dimensional example.

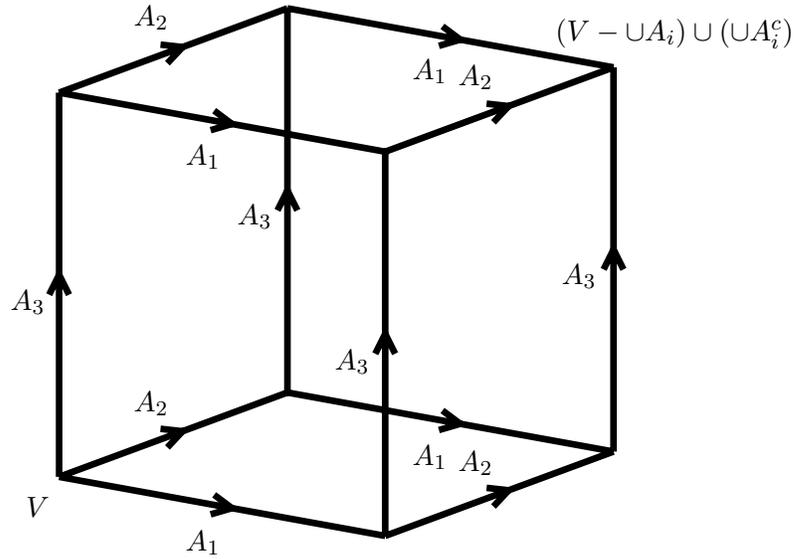


FIGURE 4. A 3-cube in  $X$

The following is Theorem 3.7 from [35].

**Theorem 3.9** (CAT(0)). *The complex  $X$  is CAT(0). In other words, it is simply connected, and NPC.*

*Proof.* An edge emanating from a vertex  $V$  is determined by a choice of minimal  $A \in V$ . We let  $(V; A)$  denote the vertex adjacent to  $V$  obtained by replacing  $A$  by  $A^c$ , and similarly use the notation  $V_i := (V; A_1, A_2, \dots, A_i)$  for the result of moving along a path of length  $i$ . Suppose  $V = (V; A_1, A_2, \dots, A_n)$  for some shortest non-contractible loop. There is necessarily some first index  $i > j$  for which  $A_i = A_j^c$  for some  $j$ . Now, by Lemma 3.6  $A_j^c$  is minimal in both  $V_j$  and in  $V_{i-1}$ , and since  $A_{j+1}^c \in V_{i-1}$  we have that  $A_j^c$  is actually minimal in  $V_{j+1}$ , and therefore  $A_{j+1}$  is minimal in  $V_{j-1}$ . So actually no two elements of  $A_j, A_j^c, A_{j+1}, A_{j+1}^c$  are comparable with respect to inclusion and there is a square in  $X^{(2)}$  with edges  $A_j, A_{j+1}$ . Pushing over this square we get another homotopic path from  $V$  to  $V_i = (V; A_1, \dots, A_{j-1}, A_{j+1}, A_j, \dots, A_j)$ . Repeating this argument inductively we see

that  $V_i = (V; A_1, \dots, A_{j-1}, A_{j+1}, \dots, A_j, A_j^c)$  so the loop can be shortened by a homotopy, contrary to hypothesis. This proves  $X$  is simply connected.

The argument that  $X$  is NPC (equivalently, that vertex links are flag) is straightforward from the construction. Let  $V$  be a vertex and suppose there is a  $K_n$  in the link of  $V$ . Each vertex in the  $K_n$  corresponds to an oriented edge (which we may suppose to point out from  $V$ ) and therefore a minimal  $A_i$  in  $V$ . Each edge of the  $K_n$  corresponds to a square spanned by  $A_i, A_j$  and therefore  $A_i$  is not contained in, nor does it contain  $A_j^c$  and conversely. But this is exactly the condition that all the  $A_i$  together span an oriented  $n$ -cube with source vertex  $V$ , giving rise to an  $(n - 1)$ -simplex in the link filling in the  $K_n$ .  $\square$

This construction can be generalized to a finite collection of subgroups  $H_i$  with  $e(G, H_i) \geq 2$ . For each  $H_i$  we choose some  $H_i$ -invariant set  $A_i$  as above, and define  $\Sigma$  to be the union of the set of  $G$  translates of some  $A_i$  or  $A_i^c$ , then define vertices, edges, etc. as above. The proofs go through essentially without change.

There is a natural simplicial action of  $G$  on  $X$ , but there is no reason at this point to think that the action is proper. This is addressed in the next section.

*Example 3.10* (Embedded loops on a surface). Let  $S$  be a closed surface, and  $\Gamma$  a collection of disjoint embedded loops. The universal cover  $\tilde{S}$  is a plane, and  $\tilde{\Gamma}$  is a collection of proper disjoint lines. Vertices correspond to complementary regions of  $\tilde{S} - \tilde{\Gamma}$  and edges to components of  $\tilde{\Gamma}$ . Thus  $X$  is a tree, and  $X/\Gamma$  is a finite graph giving the decomposition of  $\pi_1(S)$  over cyclic subgroups corresponding to the elements of  $\Gamma$ .

*Example 3.11* (Immersed loops on a surface). Let  $S$  be a closed hyperbolic surface, and  $\Gamma$  an immersed collection of geodesics. Now  $\tilde{S}$  is decomposed by  $\tilde{\Gamma}$  into many regions. There is one cube of dimension  $n$  in  $X$  for each family of  $n$  components of  $\tilde{\Gamma}$  that intersect pairwise. For example, if  $\Gamma$  can be partitioned into  $\Gamma_1, \Gamma_2$  where each  $\Gamma_i$  is embedded, then  $X$  is a CAT(0) square complex, and  $X/\pi_1(S)$  is the square complex dual to the cellulation of  $S$  by components of  $S - \Gamma$ . If  $\tilde{\Gamma}$  contains a triple of elements  $\gamma_1, \gamma_2, \gamma_3$  which intersect pairwise, there are two (nondegenerate) combinatorial possibilities for the intersection, related by a Reidemeister 3 move. The 8 vertices of the associated 3-cube correspond to the 8 complementary regions — the 6 unbounded regions (which are present in each combinatorial realization) and the 2 bounded regions, one of which appears in each of the two combinatorial realizations. See Figure 5.

At least for totally geodesic codimension 1 submanifolds in a hyperbolic manifold there is a standard construction that lets one “see” the cubes directly. A hyperplane arrangement in a projective space  $\mathbb{P}V$  determines a set of (dual) unit vectors in  $V^*$  which span a zonohedron  $Z$  (by taking Minkowski sum). The cellulation of  $\partial Z$  by facets is dual to a cellulation of the unit sphere  $UV$ , which can be thought of as the double cover of  $\mathbb{P}V$ , so zones in the zonohedron  $Z$  (i.e. equivalence classes of parallel edges) are dual to the hyperplanes in the arrangement. Now think of the zonohedron as the projection to  $V^*$  of a high dimensional cube. Each hyperplane in the arrangement corresponds to a midcube, dual to an equivalence class of edges.

In the case of immersed geodesics  $\Gamma$  in a hyperbolic surface, the cubes in  $X$  correspond (in this way) to finite sets of geodesics in  $\tilde{\Gamma}$  which pairwise intersect.

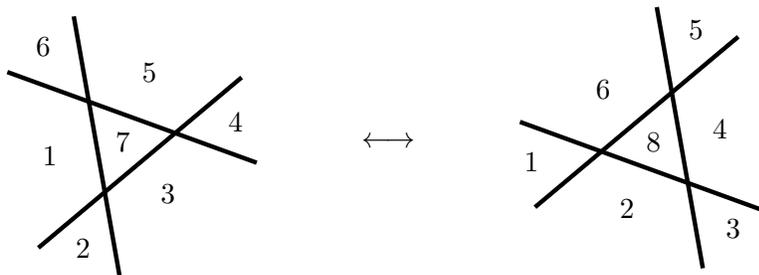


FIGURE 5. The 8 complementary regions corresponding to the 8 vertices of a 3-cube

The converse of Sageev’s construction is straightforward (under suitable hypotheses); namely, if  $G$  acts without a global fixed point (equivalently, without a bounded orbit) on a finite-dimensional CAT(0) cube complex, then  $G$  has a codimension 1 subgroup, namely the stabilizer of an oriented hyperplane. This is Theorem 5.1 in [35], but is superfluous for applications to the VHC.

**3.3. Boundary criterion.** Now let’s restrict attention to *quasiconvex* codimension 1 subgroups of *hyperbolic* groups. The following is Theorem 3.1 in Sageev [36] (but the reader is referred to [16] for part of the proof). The argument we give is a bit different from that in the literature.

First we begin with a lemma.

**Lemma 3.12** (Quasiconvex Helly’s Theorem). *Let  $G$  be a hyperbolic group, and fix some  $K > 1$ . For any integer  $k$  and for any  $R$ , there is an  $R'$  so that if  $L_1, \dots, L_k$  are  $K$ -quasiconvex subsets of  $G$  with  $N_R(L_i) \cap N_R(L_j)$  is nonempty for each pair  $i, j$  then  $\cap_i N_{R'}(L_i)$  is nonempty.*

*Proof.* Helly’s theorem says that if  $X$  is an  $n$ -dimensional CAT(0) space, any collection of convex subsets whose  $(n + 1)$ -fold intersections are nonempty has a nonempty common intersection. The asymptotic cone of a hyperbolic group is an  $\mathbb{R}$ -tree; in particular it is a 1-dimensional CAT(0) space. In the asymptotic cone, quasiconvex subsets become convex. Translating back to  $G$  proves the desired claim.  $\square$

**Proposition 3.13** (Cocompact action). *Let  $G$  be hyperbolic, and let  $H_i$  be a finite family of codimension 1 quasiconvex subgroups of  $G$ . Then  $X$  is finite dimensional, and  $X/G$  is compact.*

*Proof.* The translates of the  $H_i$  are a collection of quasiconvex subsets of  $G$  which coarsely separate; we refer to these (somewhat informally) as *hyperplanes*. A collection of  $n$  hyperplanes is dual to the edges of an  $n$ -cube if (roughly speaking) each pair  $L, L'$  of hyperplanes “intersects essentially”. This does not imply that  $L$  and  $L'$  literally intersect, but it does imply that they *coarsely intersect*, so that there is a universal constant  $R$  and a point within distance  $R$  of both  $L$  and  $L'$ .

If  $H$  is a quasiconvex subgroup of a hyperbolic group  $G$ , the translates of  $H$  do not *accumulate*; i.e. for any open cover of  $\partial_\infty G$  there are only finitely many translates whose

limit set is not contained in some element of the open cover. The same is true if one considers translates of finitely many quasiconvex  $H_i$ . Given constants  $R, T$  and we can cover  $\partial_\infty G$  by open sets so that if  $U, U'$  are two elements of the cover which can be joined by a geodesic that comes within distance  $T$  of the origin, and if  $L, L'$  are hyperplanes whose limit sets are contained in  $U$  and  $U'$  respectively, then no point is within distance  $R$  of both  $L$  and  $L'$ .

Now, given any collection of hyperplanes  $L_1, \dots, L_k$  we translate the collection by left multiplication so that origin is the point that minimizes  $\sum_i d(p, L_i)$ . There is a fixed  $k'$  so that at most  $k'$  of the  $L_i$  have limit set not contained in some  $U$  from the open cover as above. If we choose  $k \gg k'$  then I claim that there are at least two  $L_i, L_j$  with limit sets contained in  $U, U'$  so that  $U$  and  $U'$  may be joined by a geodesic  $\gamma$  coming within distance  $T$  of the origin. For, otherwise, all but  $k'$  of the  $L_i$  have limit sets contained in a collection of open sets which are all “in the same approximate direction” as seen from the origin; in particular, by moving the basepoint in the direction where these open sets are clustered, we can find a new point  $p$  with  $\sum_i d(p, L_i) < \sum_i d(\text{id}, L_i)$ , contrary to hypothesis. It follows that there is no point within distance  $R$  of both  $L_i, L_j$  for some indices  $i, j$  and therefore the dimension of  $X$  is  $< k$ ; i.e. it is *finite*.

To deduce cocompactness, let  $l$  be the dimension of  $X$ , and let  $L_1, \dots, L_l$  be a collection of hyperplanes dual to a top dimensional cube of  $X$ . By Lemma 3.12 there is some point  $p$  within distance  $R'$  of every  $L_i$ . Translating this point back to the origin, we see that there are only finitely many choices for the collection of hyperplanes up to the left  $G$ -action. Hence there are only finitely many  $l$ -cubes in  $X/G$ , as claimed.  $\square$

However, without further hypotheses, we cannot conclude that the action will be proper; for example, if  $G$  splits over  $H$  (so that  $X^{(1)}$  is a tree) the action is not cocompact unless  $H$  and the vertex groups are both finite (so that  $G$  is virtually free).

Since cube stabilizers are (virtually) subgroups of the stabilizers of their faces, the action of  $G$  on  $X^{(1)}$  is proper whenever vertex stabilizers are finite. We now give a simple criterion, observed by Bergeron-Wise, which ensures that vertex stabilizers are finite, and therefore that the action of  $G$  on  $X^{(1)}$  is proper and cocompact. Thus, in this case,  $G$  is (virtually) isomorphic to the fundamental group of a compact NPC cube complex.

**Proposition 3.14** (Bergeron-Wise boundary criterion). *Let  $G$  be word hyperbolic. Suppose that for each pair of distinct points  $p, q$  in  $\partial_\infty G$  there is a quasiconvex codimension 1 subgroup  $H$  such that  $p$  and  $q$  are in distinct components of  $\partial_\infty G - \partial_\infty H$ . Then there is a finite collection  $H_i$  of quasiconvex codimension 1 subgroups such that the action of  $G$  on the corresponding CAT(0) cube complex  $X$  is proper and cocompact.*

*Proof.* It suffices to show that we can choose a finite collection of subgroups so that the stabilizer of a cube of maximal dimension (in the associated complex  $X$ ) is finite. For every bi-infinite geodesic  $\gamma$  through the origin we can find a translate of a subgroup  $H_\gamma$  which separates the endpoints  $\gamma^\pm$ . By compactness of the space of geodesics through the origin we can find finitely many  $H_i$  such that the  $H_\gamma$  may all be taken to be translates of some  $H_i$ .

A maximal cube corresponds to a collection of translates  $L_1, \dots, L_k$ . Their coarse intersection is quasiconvex, as in the proof of Proposition 3.13. If the stabilizer of this cube is

infinite, this coarse intersection is noncompact, and coarsely contains an infinite geodesic  $\gamma$ . But now some  $H_\gamma$  as above will intersect all the  $L_i$  essentially, showing that the cube was not maximal after all.  $\square$

#### 4. ALMOST GEODESIC SURFACES

In this section we will prove the following theorem, after Kahn-Markovic [25]:

**Theorem 4.1** (Kahn-Markovic surface theorem). *Let  $M$  be a closed hyperbolic 3-manifold, let  $\epsilon > 0$ , let  $p \in M$  and  $v \in UT_p M$ . Then there is a  $(1 + \epsilon)$ -quasigeodesic closed immersed surface  $S$  in  $M$  passing through  $p$  and perpendicular to  $v$ .*

If  $\epsilon$  is sufficiently small, the surface  $S$  is necessarily  $\pi_1$ -injective. Since it is closed, to show that  $M$  is virtually Haken it would suffice to show that the subgroup  $\pi_1(S)$  is separable.

In any case,  $\pi_1(S)$  is evidently a codimension 1 quasiconvex subgroup. Since  $p$  and  $v$  can be chosen arbitrarily, it follows that pairs of points in  $\partial_\infty \pi_1(M)$  can be separated by codimension 1 quasiconvex subgroups. By Proposition 3.14 we conclude:

**Corollary 4.2.** *Let  $M$  be a closed hyperbolic 3-manifold. Then  $\pi_1(M)$  is isomorphic to the fundamental group of a compact NPC cube complex.*

*Proof.* There is a CAT(0) cube complex associated to a sufficiently big collection of subgroups of the form  $\pi_1(S)$ , and the action of  $\pi_1(M)$  on this complex is proper and cocompact. Since  $\pi_1(M)$  is torsion-free, the action is free.  $\square$

**4.1. The frame flow.** The surface  $S$  will be obtained by gluing up almost geodesic pairs of pants along their geodesic boundary. These pairs of pants are constructed using ergodic theory, in particular (sufficiently fast) mixing of the frame flow.

Let  $\xi := (\xi_1, \xi_2, \xi_3)$  be a (positive) orthonormal frame at a point  $p$  on a hyperbolic 3-manifold. There is a unique geodesic  $\gamma$  with  $\gamma(0) = p$  and  $\gamma'(0) = \xi_1$ . Parallel transport of  $\xi$  along  $\gamma$  determines a 1-parameter family  $\mathcal{F}_t(\xi)$  of frames on  $M$ ; this is the *frame flow* on the *orthonormal frame bundle*  $FM$ .

Another way to see it is to observe that  $\mathrm{PSL}(2, \mathbb{C})$  acts transitively and with trivial stabilizers on the orthonormal frame bundle of  $\mathbb{H}^3$ , and one may therefore identify this bundle with  $\mathrm{PSL}(2, \mathbb{C})$  (after choosing a baseframe), and identify  $FM$  with the left quotient  $\pi_1(M) \backslash \mathrm{PSL}(2, \mathbb{C})$ . Then frame flow  $\mathcal{F}_{2t}$  on  $FM$  is given by the *right* action of the matrix  $\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$ .

Let  $\iota$  and  $\tau$  respectively denote the effect on a frame  $\xi$  of a twist through angle  $\pi$  and  $2\pi/3$  respectively in the (oriented) plane perpendicular to  $\xi_3$ . A left-invariant metric on  $\mathrm{PSL}(2, \mathbb{C})$  descends to a metric on  $FM$ .

Let  $\xi$  and  $\xi'$  denote two frames in  $M$ , and let  $\theta := (\theta_0, \theta_1, \theta_2)$  be a triple of geodesic segments joining the basepoints of  $\xi$  and  $\xi'$ .

**Definition 4.3.** The triple  $(\xi, \xi', \theta)$  is  $(T, \epsilon)$ -*matched* if the following conditions hold:

- (1) the angle between  $\tau^i(\xi_1)$  and  $\theta_i$  should be  $< \epsilon$  for each  $i$ , and similarly for  $\tau^{3-i}(\xi'_1)$ ;
- (2) the length of each  $\theta$  should be  $\epsilon$ -close to  $T$ ; and
- (3) parallel transport of  $\xi_3$  along each  $\theta_i$  should give a vector  $\epsilon$ -close to  $\xi'_3$ .

The three geodesics  $\theta$  together make up a  $\theta$ -graph, which is the spine of a(n almost geodesic) pair of pants  $P$ . The (geodesic) cuff lengths of  $P$  are  $\epsilon$ -close to  $2R$  where (when  $\epsilon$  is sufficiently small) the law of hyperbolic cosines gives

$$\cosh(R) \sim \cosh^2(T/2) + \sinh^2(T/2)/2$$

so that  $R \sim T + \log(3/4)$ . See Figure 6.

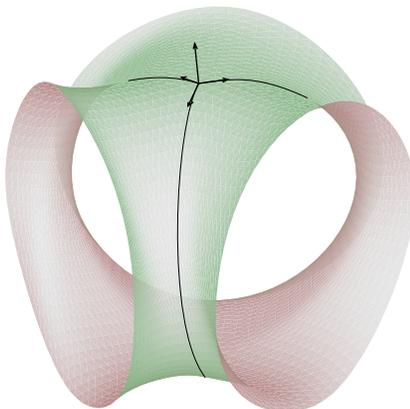


FIGURE 6. The pants  $P$  associated to  $\theta$ . Also  $\xi_3$  and  $\tau^i(\xi_1)$  are indicated.

It is important to stress that not only are  $P$ 's cuff lengths  $\epsilon$ -close to  $2R$ , but parallel transport around these cuffs is  $\epsilon$ -close to the identity transformation. One summarizes this by saying that the *complex lengths* of the cuffs of  $P$  are  $\epsilon$ -close to  $2R$ .

The graph  $\theta$  divides  $P$  into three (almost totally geodesic) annuli, which we call *lunes*. For each  $R, T$  related as above there is a “model” (totally geodesic) lune whose inner boundary is a geodesic of length  $2R$  and whose outer boundary is composed of two geodesic segments of length  $2T$  meeting at angle  $2\pi/3$ . There are two geodesic arcs from the inner to outer boundary meeting both at right angles; by symmetry, these have the same length, which is of order  $e^{-T}$ . We call the tangents at the end of these two arcs the *feet*.

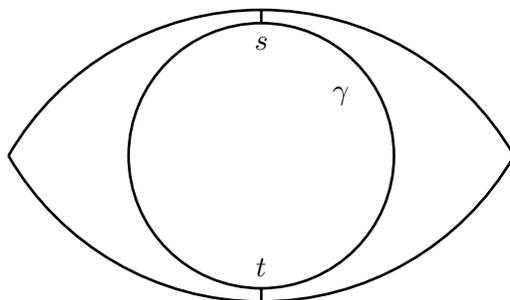


FIGURE 7. A model lune with inner boundary  $\gamma$  and feet at  $s, t$ .

Thus, a  $(T, \epsilon)$ -matched triple  $(\xi, \xi', \theta)$  determines a pair of feet  $s, t$  in the unit normal bundle  $N(\gamma)$  of each cuff  $\gamma$  of  $P$ . Let  $\Gamma$  denote the set of geodesics of complex length  $\epsilon$ -close to  $2R$ , and let  $N(\Gamma)$  denote the union of the normal bundles of the various  $\gamma \in \Gamma$ . We

think of  $N(\Gamma)$  as a disjoint union of tori. Let  $\Pi$  denote the space of  $(T, \epsilon)$ -matched triples. There is a measure  $\lambda$  on the space  $\Pi$ , which is just the Liouville measure on  $FM$  for each of the coordinates  $\xi, \xi'$ , and is the counting measure on the set of homotopy classes of graphs  $\theta$ . Then assigning to each triple  $(\xi, \xi', \theta)$  the feet of the associated pants  $P$  defines a (multi-valued) map from  $\Pi$  to  $N(\Gamma)$ , and we can let  $\nu$  denote the pushforward of  $\lambda$ .

Now, although the total  $\nu$ -measure on the unit normal bundles of different  $\gamma \in \Gamma$  can vary greatly, for a fixed  $\gamma$ , the restriction of  $\nu$  to  $N(\gamma)$  is very well equidistributed, as follows.

**Lemma 4.4.** *There is a positive constant  $C$  so that for any  $\epsilon > 0$ , for any  $T > C \log(\epsilon^{-1})$ , and for any  $\gamma \in \Gamma$  there is there is some positive constant  $D := D(\gamma)$  so that for every  $\epsilon$ -ball  $B$  in  $N(\gamma)$  there is an estimate*

$$(1 - \epsilon)D \leq \nu(B) \leq (1 + \epsilon)D$$

*Proof.* This follows from exponential mixing of the frame flow, due to Pollicott [34]. Intuitively, it can be seen from the fact that if  $\alpha$  is a geodesic segment of length  $T$ , then if we fix one endpoint of  $\alpha$  and move the other through distance 1 the tangent vector of  $\alpha$  at the fixed endpoint moves through distance  $O(e^{-T})$ . Given a geodesic  $\gamma$  and a pair of antipodal feet  $s, t$  we can construct a lune  $L$  with feet at  $s, t$ , with edges of length  $\epsilon$ -close to  $T$ , and making an angle  $\epsilon$  close to  $2\pi/3$  at the two endpoints. The outer edges of the lune are two geodesics  $\theta_0, \theta_1$  as above. Mixing implies that the set of configurations of a third geodesic  $\theta_2$  for which there is a compatible  $(T, \epsilon)$ -matched triple  $(\xi, \xi', \theta)$  depends (up to error of order  $\epsilon$ ) only on the geometry of  $\theta_0$  and  $\theta_1$ , and not on how they sit in  $M$ , at least for  $T$  big enough depending on  $\epsilon$ .

The unit normal bundle  $N(\gamma)$  is topologically a torus, which can be identified with (the connected component of the identity of) its group of isometries. These isometries extend to isometries of the infinite cover of  $M$  with cyclic fundamental group  $\langle \gamma \rangle$ , and we can move the lune  $L$  around by this group. The mass of the set of  $(T, \epsilon)$ -matched triples compatible with each translate of this lune is almost independent of the choice of translation; the result follows.  $\square$

*Remark 4.5.* Actually, there is a subtle point in the proof of Lemma 4.4 which is that the constant  $D$  in question will really depend (very dramatically) on the choice of  $\gamma$  in  $\Gamma$ , so that equidistribution of  $\nu$  holds (with small error) only in each individual  $N(\gamma)$ , and not in  $N(\Gamma)$  as a whole.

To see why this is true, consider a geodesic  $\gamma$  with complex length  $2R + s$  for some real number  $s$ . For any antipodal pair of feet  $\nu, \nu'$  the lune  $L$  with these feet becomes more and more constrained, so that at some critical  $S$  the edges are both forced to have length *exactly*  $T + \epsilon$  and to meet at an angle of exactly  $2\pi/3 - 2\epsilon$ . For such a lune  $L$  there are a *unique* pair of frames  $\xi, \xi'$  at the vertices of the lune which might be part of a  $(T, \epsilon)$ -matched triple  $(\xi, \xi', \theta)$ , where two of the legs of  $\theta$  are the two sides of  $L$ . However, even in this critical configuration, there is *still* a big open family of third legs  $\theta_2$  whose length is  $\epsilon$ -close to  $T$  and whose tangent vector is  $\epsilon$ -close to  $\tau^2(\xi_1)$  and  $\tau^1(\xi'_1)$  at the endpoints. Thus, the number of homotopy classes of configurations  $\theta$  compatible with  $L$  is almost independent of the choice of  $\nu, \nu'$  no matter *how small* the volume of the set of compatible  $\xi, \xi'$ .

The space  $\Pi$  is an open manifold locally modeled on  $FM \times FM$  with a finite total (Liouville) measure  $\lambda$ . We can approximate the continuous measure  $\lambda$  by an atomic measure  $\lambda'$  in such a way that for every  $\delta$ -ball  $B$  in  $\Pi$  there is an estimate  $(1 - \delta) < \lambda(B)/\lambda'(B) < (1 + \delta)$ , where we pick some  $0 < \delta \ll \epsilon$ . It is also convenient for the masses of the atoms of  $\lambda'$  to be rational. Now multiply through by a big integer to clear denominators; this gives a new measure which by abuse of notation we also denote  $\lambda'$ . Each atom determines a spine  $\theta$  of some pants  $P$  with feet in  $N(\Gamma)$ , and we can think of the (integral) mass of each atom as determining a positive integral weight on  $P$ . Let  $\mathcal{P}$  denote the formal positive integral linear combination of pants  $P$  arising in this way. We will show how to glue up  $2\mathcal{P}$  along pairs of cuffs to build a closed  $(1 + \epsilon)$ -quasigeodesic surface.

**4.2. Gluing with a twist.** The atomic measure  $\lambda'$  pushes forward to an atomic measure  $\nu'$  on  $N(\Gamma)$ . For each  $\gamma$  the atoms of  $\nu'$  on  $N(\gamma)$  mark the feet of the pants in  $\mathcal{P}$  that land on  $\gamma$  (with multiplicity). We think of  $\nu'|_{N(\gamma)}$  as a finite subset with multiplicity. A fixed-point free involution  $j$  on  $\nu'|_{N(\gamma)}$  determines a pairing of the cuffs that land on  $\gamma$ ; gluing pairs by this pairing for each  $\gamma$  will determine a closed surface  $S$  from  $\mathcal{P}$ .

Actually, it is convenient to work with  $2\nu'$  instead of  $\nu'$  for the simple reason that the total mass of  $2\nu'$  on each  $N(\gamma)$  is even, which is a prerequisite for the existence of a fixed-point free involution. There is an additional reason to work with  $2\nu'$  instead of  $\nu'$ , which is that we want to build an *oriented* surface, and therefore for each pants we take two copies of it, one with either orientation, and glue pairs of pants compatibly with the orientations.

We parameterize  $N(\gamma)$  as  $\mathbb{C}/(2\pi i\mathbb{Z} \oplus \ell\mathbb{Z})$  where  $\ell$  is the complex length. In these coordinates, the group of isometries acts on  $N(\gamma)$  by (complex) addition.

**Definition 4.6.** Let  $\Sigma$  be a finite subset of  $N(\gamma)$  (possibly with multiplicity). Let  $2\Sigma$  denote the set  $\Sigma \times \pm 1$ . An involution  $j : 2\Sigma \rightarrow 2\Sigma$  is  $\epsilon$ -well-matched if it interchanges  $\Sigma \times +$  and  $\Sigma \times -$ , and for each  $s \in \Sigma \times +$  we have  $|j(s) - s - (1 + i\pi)| < \epsilon$ .

If  $\Sigma$  is the set of feet of a collection of pants, we can think of  $\pm$  as a choice of coorientation on each pant. If  $j(s, +) = (t, -)$  the pants  $P_s$  and  $P_t$  are on almost opposite sides of  $\gamma$ , and for each pant, the foot on the opposite pant is approximately distance 1 to the right, as measure along  $\gamma$  with the induced orientation. This makes sense, since the orientations on  $\gamma$  coming from the two pants *disagree*; note that this means that the surface  $P_s \cup P_t$  is oriented and  $(1 + \epsilon)$ -quasigeodesic.

**Lemma 4.7.** *Let  $\Sigma$  be a finite subset of  $N(\gamma)$  for which there is some positive constant  $D$  so that for every  $\epsilon$ -ball  $B$  in  $N(\gamma)$  there is an estimate*

$$(1 - \epsilon)D \leq \#(\Sigma \cap B) \leq (1 + \epsilon)D$$

*Then there is a  $5\epsilon$ -well-matched involution  $j$  from  $2\Sigma$  to  $2\Sigma$ .*

*Proof.* This can be proved by Hall's marriage lemma; this is the method favored by Kahn-Markovic. But actually it is easy to construct an explicit involution, perhaps at the cost of multiplying  $\Sigma$  by some big fixed integer. We explain this construction. Note that since  $j$  is an involution, we just need to define a map  $j : \Sigma \rightarrow \Sigma$ , and let this stand for  $j : \Sigma \times + \rightarrow \Sigma \times -$ .

First, let's decompose the torus  $N(\gamma)$  into annuli  $A_i$  of the form  $N(\sigma_i)$  for  $\sigma_i$  the elements of a partition of  $\gamma$  into disjoint intervals, so that each  $\sigma_i$  has width  $\sim \epsilon$ , and each  $A_i$  contains

the same number of points of  $\Sigma$  (this is the point where we need the number of points in  $\Sigma$  to be divisible by a big integer). Pick one interval and call it  $A_0$ , label the others with cyclic indices, and let  $k$  be the index such that some point in  $\sigma_k$  is distance 1 from  $\sigma_0$  as measured along  $\gamma$  (in some fixed direction). The widths of the intervals from  $\sigma_i$  to  $\sigma_{i+k}$  are contained in  $(1 \pm \epsilon)$ , so it suffices to choose  $j : \sigma_i \rightarrow \sigma_{i+k}$  for each  $i$ . Now order the elements of each  $\sigma_i$  cyclically by their imaginary part, and define  $j$  by first taking  $p \in \sigma_i$ , to  $q \in \sigma_{i+k}$  whose imaginary parts differ by  $\sim i\pi$ , and then extending in cyclic order. This map has the desired properties.  $\square$

**4.3. Thin parts don't accumulate.** By Lemma 4.7 we can glue up  $2\mathcal{P}$  to a closed oriented surface  $S$  immersed in  $M$ . We now show that  $S$  as above is  $(1 + \epsilon)$ -quasigeodesic.

A pair of pants with cuffs of length  $2R$  has legs of length  $\sim e^{-R}$ . So for big  $R$ , when several such pants are glued together, there is *a priori* a danger that the thin parts will accumulate. Even if the gluing angle is of order  $\epsilon$  along each cuff, and  $R \sim C \log(1/\epsilon)$  the resulting surface might not be quasigeodesic.

However, if we glue cuffs by a well-matched involution, a geodesic transverse to the cuffs will enter the thick part of a pants after crossing only  $O(T)$  cuffs. Figure 8 shows a collection of hyperbolic geodesics where each is obtained from the previous one by an orthogonal displacement through  $\epsilon$  followed by a shear of length 1.

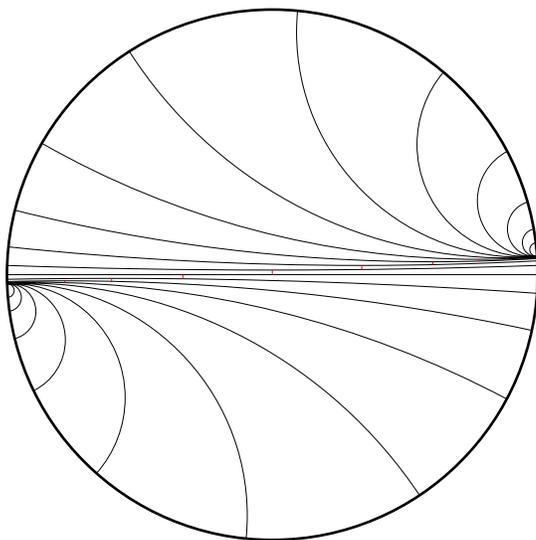


FIGURE 8. Shearing by 1 at each cuff spreads out the thin parts

If  $\epsilon \sim e^{-T}$ , a geodesic of length  $O(1)$  will only cross  $O(T)$  cuffs. So if the bending angle at each cuff is of order  $\epsilon$ , the resulting surface will be  $1 + \epsilon^{1-\delta}$ -quasigeodesic for any  $\delta$ . After adjusting  $\epsilon$ , this completes the proof of Theorem 4.1.

5. RELATIVE HYPERBOLICITY

**5.1. Hyperbolic Dehn filling.** Relative hyperbolicity for groups generalizes hyperbolicity in the same way that finite volume complete hyperbolic manifolds generalize compact hyperbolic manifolds. The ends of such manifolds are cusps, which are the quotient of horoballs by parabolic groups of isometries.

If  $G$  is the fundamental group of a complete finite hyperbolic manifold  $M$ , and  $P$  is the parabolic subgroup, the Cayley graph of  $G$  is not typically quasi-isometric to the universal cover of  $M$ , but it can be completed by gluing in “combinatorial” horoballs along the conjugates of  $P$ .

**Definition 5.1** (Combinatorial horoball). Let  $\Gamma$  be a simplicial graph with vertices  $V$ . The *combinatorial horoball* associated to  $\Gamma$ , denoted  $\mathcal{H}(\Gamma)$ , is the graph with vertex set  $V \times \mathbb{N}$ , and with two kinds of edges:

- (1) a vertical edge from  $(v, i)$  to  $(v, i + 1)$  for each  $v \in V$  and  $i \in \mathbb{N}$ ; and
- (2) a horizontal edge from  $(v, i)$  to  $(w, i)$  whenever  $d_\Gamma(v, w) \leq 2^i$ .

Combinatorial horoballs are defined, and their properties are established in [19]. The graph  $\mathcal{H}(\Gamma)$  can be canonically completed to a simply-connected space, by gluing in a disk along every loop of length  $\leq 5$ . This new space satisfies a linear isoperimetric inequality. An arc  $\delta$  in  $\mathcal{H}(\Gamma)$  consisting of two vertical edges together with a horizontal path joining their upper vertices can be pushed down to a horizontal path  $\delta'$  with  $|\delta'| \leq |\delta|/2 - 1$ . A completely horizontal loop  $\gamma$  can be pushed vertically downwards one level to a loop  $\gamma'$  with  $|\gamma'| \leq |\gamma|/2 + 1$  until  $|\gamma| \leq 3$  when it can be filled in with a single disk. By induction therefore every simplicial loop  $\gamma$  can be filled in with a disk of area  $\sim 2|\gamma|$ , and it follows that  $\mathcal{H}(\Gamma)$  is  $\delta$ -hyperbolic (for some universal  $\delta$ ).

Any two vertices  $v, w$  of  $\mathcal{H}(\Gamma)$  can be joined by a *model geodesic*, consisting of the union of at most two vertical paths, and at most one horizontal path of length at most 3. Notice for  $u, v \in V$  that  $d_{\mathcal{H}(\Gamma)}((u, 0), (v, 0)) \sim \log(d_\Gamma(u, v))$ , so the inclusion of  $\Gamma$  into  $\mathcal{H}(\Gamma)$  is highly distorted.

Let  $G$  be a finitely generated group, and let  $P$  be a finitely generated subgroup. It is possible to choose a finite generating set  $S$  for  $G$  such that  $S \cap P$  is a generating set for  $P$  (just choose any finite generating set for  $P$  and add it to any finite generating set for  $G$ ). For such a  $P$  the Cayley graph  $C_{S \cap P}(P)$  is a subgraph of the Cayley graph  $C_S(G)$  and it makes sense to attach a combinatorial horoball to each (left) translate of  $C_{S \cap P}(P)$  in  $C_S(G)$  to get a new graph  $X(G, \mathcal{P}, S)$  (or just  $X$  for short if  $G$  and  $\mathcal{P}$  are understood) on which  $G$  acts isometrically by left translation (although not cocompactly; the quotient is quasi-isometric to a ray).

**Definition 5.2** (Relatively hyperbolic). Let  $G$  be a finitely generated group, and let  $\mathcal{P} := \{P_i\}$  be a finite collection of finitely generated subgroups of  $G$ . We say  $(G, \mathcal{P})$  is *relatively hyperbolic* (or:  $G$  is *hyperbolic relative to*  $\mathcal{P}$ ) if the space  $X$  obtained from  $C_S(G)$  by attaching a combinatorial horoball to each translate of each  $C_{S \cap P_i}(P_i)$  is hyperbolic.

*Example 5.3.* If each of the  $P_i$  is finite, then  $(G, \mathcal{P})$  is relatively hyperbolic if and only if  $G$  is hyperbolic.

*Example 5.4* (Isoperimetric inequalities propagate). If  $G$  is hyperbolic relative to  $\mathcal{P}$ , then Farb [15], Thm. 3.8 showed that any isoperimetric function for the  $\mathcal{P}$  is actually an isoperimetric function for  $G$ . It follows that if each  $P_i$  in  $\mathcal{P}$  is itself hyperbolic, then  $G$  is hyperbolic.

*Example 5.5* (Almost malnormal). A collection of subgroups  $\mathcal{P} := \{P_i\}$  of a group  $G$  is said to be *almost malnormal* if, for all  $g \in G$ , the intersection  $P_i^g \cap P_j$  is infinite only when  $i = j$  and  $g \in P_i$ . Bowditch [8] proved that if  $G$  is hyperbolic, and  $\mathcal{P}$  is a collection of almost malnormal quasiconvex subgroups of  $G$ , then  $G$  is hyperbolic relative to  $\mathcal{P}$ .

*Example 5.6* (Cusped 3-manifold). Let  $M$  be a complete noncompact finite-volume hyperbolic 3-manifold ( $M$  is also called a *cusped* hyperbolic 3-manifold). Then  $M$  is homeomorphic to the interior of a compact 3-manifold  $\bar{M}$  whose boundary is a union of tori  $T_i$ . The fundamental group  $\pi_1(M)$  is hyperbolic relative to the collection of subgroups  $\pi_1(T_i)$ .

Thurston famously showed that almost every Dehn filling on a cusped hyperbolic 3-manifold  $M$  gives rise to a closed hyperbolic 3-manifold  $N$ . Topologically, a Dehn filling is obtained by gluing a solid torus along its boundary to each of the cusps of  $M$ . At the level of fundamental groups,  $\pi_1(N)$  is obtained from  $\pi_1(M)$  by killing the meridian loop in  $\pi_1(T_i)$  for each  $T_i$ .

The meridian loops are primitive elements in each  $\pi_1(T_i)$ . The group obtained by killing the power of a primitive element is (again, with finitely many exceptions for each cusp) the fundamental group of a hyperbolic orbifold; in particular, it is still word-hyperbolic. So a much weaker (but still very useful) statement than Thurston's theorem is just the observation that for all but finitely many Dehn fillings on a cusped hyperbolic 3-manifold the fundamental group of the result is a hyperbolic group. This latter statement is easier to make effective.

*Example 5.7* ( $2\pi$ -Theorem [7]). Let  $M$  be a cusped hyperbolic 3-manifold, with one cusp for convenience. Let  $T$  be an embedded (Euclidean) horotorus section of the cusp. Let  $\alpha$  be a simple geodesic in  $T$  (with its Euclidean metric) of length  $> 2\pi$ . Then Gromov and Thurston showed that the result of Dehn filling on  $M$  with slope  $\alpha$  admits a metric of negative curvature.

This can be done with an explicit metric. Let  $S$  be a hyperbolic surface with a cusp, and let  $S'$  be obtained by removing a neighborhood of the cusp so that  $\partial S'$  is a horocycle of length  $T > 2\pi$  and with constant geodesic curvature 1 (pointing into the cusp). Let  $D(K, r)$  be a disk of radius  $r$  in a plane of constant curvature  $K$  (negative). Then

$$\text{length}(\partial D) = 2\pi \sinh(r|K|)/|K| = 2\pi(r + r^3K^2/6 + \dots)$$

If we vary  $K$  between  $-1$  and  $0$  and choose  $r$  so that  $\text{length}(\partial D) = 2\pi$ , then  $r$  varies monotonically between  $0.88137$  and  $1$ , and the geodesic curvature along the boundary varies monotonically between  $\sqrt{2}$  and  $1$ . Increasing  $r$  slightly so that  $\text{length}(\partial D) = T > 2\pi$  we see that there is some strictly negative  $K$  with  $k = 1$ . Gluing in this disk along  $\partial S'$  gives rise to a  $\text{CAT}(K)$  metric on a closed surface. To do a Dehn filling on  $M$ , just glue a copy of  $D$  along each straight horocircle in the homotopy class of  $\alpha$ , and put a suitable metric on the transverse direction to make these disks totally geodesic, and the result  $\text{CAT}(K)$ .

Example 5.7 illustrates the crucial geometric idea behind hyperbolic Dehn surgery. The geodesic curvature along the (concave) boundary of a horoball complement is *bounded away*

from zero, whereas the geodesic curvature along the (convex) boundary of a Euclidean ball of sufficiently large radius is as close to zero as desired. Thus a sufficiently long curve on the boundary of a cusp can be killed in  $\pi_1$  while keeping the curvature negative, and without distorting the geometry much away from the cusp.

At the level of fundamental groups,  $\pi_1(M)$  is hyperbolic relative to the cusp group  $\pi_1(T)$ . A primitive element  $\alpha \in \pi_1(T)$  generates a (normal) subgroup  $\langle\langle \alpha \rangle\rangle_{\pi_1(T)} \trianglelefteq \pi_1(T)$ . When we Dehn fill  $\alpha$  we get a closed 3-manifold  $M(\alpha)$ , and  $\pi_1(M(\alpha)) = \pi_1(M)/\langle\langle \alpha \rangle\rangle_{\pi_1(M)}$ ; i.e. we kill the normal closure of  $\alpha$  in  $\pi_1(M)$ . *A priori*, without any geometric hypothesis, this normal closure in  $\pi_1(M)$  might intersect  $\pi_1(T)$  in a bigger subgroup than  $\langle\langle \alpha \rangle\rangle_{\pi_1(T)}$ ; for example,  $M$  might be a knot complement in  $S^3$ , and  $\alpha$  might be the meridian. But under the geometric hypothesis that  $\alpha > 2\pi$  we see that the natural map

$$\pi_1(T)/\langle\langle \alpha \rangle\rangle_{\pi_1(T)} \rightarrow \pi_1(M)/\langle\langle \alpha \rangle\rangle_{\pi_1(M)}$$

is injective. Furthermore, if  $\alpha$  is long enough, loops in the thick part of  $M$  stay in the thick part of  $M(\alpha)$ , and their image in  $\pi_1(M(\alpha))$  is nontrivial, and not conjugate into  $\pi_1(T)/\langle\langle \alpha \rangle\rangle_{\pi_1(T)}$ .

This can be translated into purely combinatorial language, with combinatorial horoballs playing the role of cusps, and curvature being controlled on a mesoscopic scale by isoperimetric inequalities (where Gauss-Bonnet translates isoperimetric inequalities into curvature bounds in the Riemannian world).

Let  $G$  be hyperbolic relative to  $\mathcal{P} := \{P_i\}$  as above, and let  $N_i \trianglelefteq P_i$  be a normal subgroup for each  $i$ . Let  $\mathcal{N} = \langle\langle N_i \rangle\rangle$  be the normal subgroup of  $G$  generated by all the  $N_i$ . We say that  $G/\mathcal{N}$  is the result of *Dehn filling*  $G$  along the  $N_i$ .

The next theorem follows directly from the main theorems of [31] and [19] except for the last statement, which is explained in the appendix of [4].

**Theorem 5.8** (Groves-Manning, Osin; Hyperbolic Dehn Surgery Theorem). *Suppose that  $G$  is hyperbolic relative to  $\mathcal{P}$ . Then there is a finite subset  $A$  of  $G - \text{id}$  so that if no  $N_i$  meets  $A$ , then*

- (1) *the natural map  $P_i/N_i \rightarrow G/\mathcal{N}$  is injective for all  $i$ ; and*
- (2) *the group  $G/\mathcal{N}$  is hyperbolic relative to the  $\{P_i/N_i\}$ .*

*Furthermore, if  $F$  is any finite subset of  $G$  we can choose  $A$  as above so that  $\phi : G \rightarrow G/\mathcal{N}$  is injective on  $F$ , and  $\phi(F) \cap \phi(P_i) = \phi(F \cap P_i)$  for all  $i$ .*

It is impossible to give a really detailed proof of Theorem 5.8 in a few words, but we can at least explain how relative hyperbolicity of  $(G, \mathcal{P})$  can be propagated to relative hyperbolicity of  $(G/\mathcal{N}, \{P_i/N_i\})$  under the hypothesis that the elements in the  $N_i$  are “sufficiently long”.

*Proof.* Osin’s argument is a kind of generalization of small cancellation theory. In the classical small cancellation theory, one considers a group of the form

$$G = \langle x_1, \dots, x_k \mid r_1, \dots, r_s \rangle$$

where the  $r_i$  (thought of as cyclicly reduced words in the  $x_i$ ) satisfy a *small cancellation condition*  $C(\lambda)$  for some  $\lambda < 1/6$ . This means that no two  $r_i$  shares a common subword with any  $r_j^\pm$  of more than  $\lambda$  of its length (except for the trivial case of  $i = j$  and the subwords

occurring at the same place). This implies that if  $w$  is a trivial (cyclically reduced) word in  $G$  in the generators, the interior vertices in the dual graph to a van Kampen diagram for  $w$  have valence at least 7, and therefore  $G$  satisfies a linear isoperimetric inequality and is (infinite) hyperbolic.

Now think about the result of a filling  $G/\mathcal{N}$ , and imagine a van Kampen diagram for some element  $w$  in  $G$  which is killed in  $G/\mathcal{N}$ . We can think of this diagram as a disk  $D$  containing subdisks  $E_i$  whose boundaries are decorated by words which are generators in the  $N_i$ . We would like to show that  $G/\mathcal{N}$  is hyperbolic by showing that such  $D$  satisfy a linear isoperimetric inequality. This is immediate for  $D - \cup_i E_i$ , since that part of the disk lives (locally) in  $X$ , which is hyperbolic by hypothesis. If we think of a hyperbolic space on a large scale as much like a tree, this part of the diagram can be more or less discounted when the length of the elements of  $N_i$  are  $\gg \delta$ .

The problem is, exactly as in the small cancellation picture, when too many of the  $\partial E_i$  have big segments in their boundaries which are “close” in  $D$ . Now think of the  $\partial E_i$  as undistorted segments lying in the boundary of a pair of horoballs in  $X$ . If two such boundaries come close along long segments in  $D$  this means that there is part of  $X$  which looks metrically like two long horocycles glued together along their exteriors. But each horocycle is exponentially distorted in  $X$  into the interior of the horoball it lies on, as we have seen. So the space  $X$  could not have been hyperbolic after all, contrary to hypothesis. This contradiction shows, contrapositively, that  $G/\mathcal{N}$  satisfies a (relative) linear isoperimetric inequality and is (relatively) hyperbolic.

The other claims can be proved by similarly translating geometric arguments into the language of van Kampen diagrams and isoperimetric inequalities. The details are nontrivial, but can be found in the references above.  $\square$

If each  $P_i$  is residually finite, then each  $P_i$  contains a finite index normal subgroup  $P'_i$  so that the conditions of the theorem are satisfied whenever  $N_i \trianglelefteq P'_i$ .

A Dehn filling is *peripherally finite* if each  $N_i$  is finite index in  $P_i$ . Note that if each  $P_i$  is residually finite, Theorem 5.8 implies that there is a peripherally finite Dehn filling which is hyperbolic.

## 5.2. The weak separation theorem.

**Theorem 5.9** (Agol-Groves-Manning; Weak Separation Theorem). *Let  $G$  be a hyperbolic group, let  $H$  be a quasiconvex subgroup of  $G$  which is isomorphic to the fundamental group of a virtually special NPC cube complex, and let  $g \in G - H$ . Then there is a group  $G'$  and a surjection  $\phi : G \rightarrow G'$  so that*

- (1)  $G'$  is hyperbolic;
- (2)  $\phi(H)$  is finite; and
- (3)  $\phi(g)$  is not contained in  $\phi(H)$ .

Again, giving a complete proof is beyond the scope of this survey, but we can explain some of the key ideas.

Let's first consider the case where  $H$  is almost malnormal. Then by Bowditch (Example 5.5) the group  $G$  is hyperbolic relative to  $H$ . We let  $F$  as in Theorem 5.8 consist only of the element  $g$ . Since  $H$  is isomorphic to the fundamental group of a virtually special

NPC cube complex, it is linear and therefore residually finite, so it contains a finite index normal subgroup  $N$  satisfying the conclusion of the theorem.

If  $H$  is not malnormal, one must induct on an invariant called the *height*. The following definition is taken from [16].

**Definition 5.10** (Height). Let  $G$  be hyperbolic and let  $H$  be quasiconvex in  $G$ . The *height* of  $H$  is the least integer  $n$  so that if there are elements  $g_1, \dots, g_n$  so that  $H, H^{g_1}, \dots, H^{g_n}$  are distinct, then the intersection of the conjugates  $H \cap H^{g_1} \cap \dots \cap H^{g_n}$  is finite.

With this definition,  $H$  has height 0 if and only if it is finite, and has height 1 if and only if it is almost malnormal and infinite.

**Proposition 5.11** (Finite height). *Every quasiconvex subgroup of a hyperbolic group has finite height, and for any quasiconvex  $H$  there are only finitely many  $H$ -conjugacy classes of infinite groups of the form  $H \cap H^{g_1} \cap \dots \cap H^{g_k}$ .*

*Proof.* The proof uses the same kinds of geometric ideas as went into Sageev's proof of Proposition 3.13. First observe that if  $H \cap H^g = H \cap gHg^{-1}$  is infinite, then there are infinitely many points in  $H$  and  $gH$  within distance  $|g|$  of each other. Let  $p, q \in H$  and  $p', q' \in gH$  satisfy  $d_G(p, p') \leq |g|$  and  $d_G(q, q') \leq |g|$  but  $d_G(p, q) \gg 1$ . Since  $H$  and  $gH$  are both  $K$ -quasiconvex for some *fixed*  $K$  (independent of  $g$ ) it follows that there are points in  $H$  and  $gH$  at distance  $2K + 2\delta$  from each other; in particular, this distance does *not* depend on  $g$ . Hence, the translates  $H, g_1H, g_2H, \dots, g_nH$  (which are all  $K$ -quasiconvex for fixed  $K$ ) have the property that they pairwise coarsely intersect, and therefore just as in the proof of Proposition 3.13 there is a uniform bound  $n \leq k$ , and the set of points within a bounded distance of all  $g_iH$  is finite up to the left action of  $\cap_i H^{g_i}$ .  $\square$

*Remark 5.12.* In fact, the relationship between Proposition 5.11 and Proposition 3.13 is even closer. One can define a complex whose  $k$ -simplices are the  $(k + 1)$ -fold infinite intersections of distinct conjugates of  $H$ , and then height is the dimension of this complex plus one. The argument above shows this complex is finite dimensional and the conjugation action of  $G$  is cocompact.

Using this proposition, we let  $\mathcal{P}$  denote the collection of  $H$ -conjugacy classes of minimal infinite intersections of the form  $H \cap H^{g_1} \cap \dots \cap H^{g_k}$  (these are conjugacy classes of subgroups of  $H$ ). These subgroups are quasiconvex. Replace each  $P$  in  $\mathcal{P}$  by its commensurator  $P'$  in  $H$ . Note that since each  $P$  in  $\mathcal{P}$  is quasiconvex in  $H$ , each  $P$  is finite index in  $P'$ . Now choose one element  $P'$  per  $H$ -conjugacy class to produce a new collection  $\mathcal{P}'$ . By construction each  $P'$  in  $\mathcal{P}'$  is almost malnormal in  $H$ , so the pair  $(H, \mathcal{P}')$  is relatively hyperbolic, by Bowditch.

We further replace each  $P'$  in  $\mathcal{P}'$  by its commensurator  $P''$  in  $G$ . Again, each  $P'$  is quasiconvex in  $G$ , so each  $P'$  is finite index in  $P''$ , and we get (by choosing one subgroup per  $G$ -conjugacy class) a collection of subgroups  $\mathcal{P}''$  which are almost malnormal in  $G$ , so that  $(G, \mathcal{P}'')$  is relatively hyperbolic.

We would like to do Dehn filling on subgroups  $N_i$  of the  $P_i''$ . However we would like to do this in such a way that we get at the same time a Dehn filling of  $(G, \mathcal{P}'')$  and a Dehn filling of  $(H, \mathcal{P}')$ .

**Definition 5.13** ( $H$ -filling). A collection of normal subgroups  $N_i \trianglelefteq P_i''$  gives rise to an  $H$ -filling if whenever  $P_j' \cap (P_i'')^g$  is infinite for some  $P_j' \in \mathcal{P}'$ , then  $N_i^g$  is contained in  $P_j'$ .

The collection  $\{N_i\}$  gives rise to a Dehn filling  $\phi : G \rightarrow G/\mathcal{N}$  of  $(G, \mathcal{P}'')$ . A  $G$ -conjugacy class  $P_i''$  contains several  $H$ -conjugacy classes of finite index  $P'_{i,j}$ . If  $\{N_i\}$  gives rise to an  $H$ -filling, the subgroup  $N_i$  is contained in, and normal in each  $P'_{i,j}$ , and induces a Dehn filling of  $(H, \mathcal{P}')$  which we say is *induced* by the filling of  $G$ .

The following is proved in the appendix of [4] by Agol-Groves-Manning:

**Proposition 5.14** (*H*-filling). *Let  $G$  be hyperbolic, let  $H$  be quasiconvex in  $G$  of height at least 1, let  $(G, \mathcal{P}'')$  and  $(H, \mathcal{P}')$  be as above, and let  $g \in G - H$ . Then for all sufficiently long peripherally finite  $H$ -fillings  $\phi : G \rightarrow G/\mathcal{N}$ ,*

- (1)  $\phi(H)$  is isomorphic to the result of the induced filling of  $H$ ;
- (2)  $\phi(H)$  is quasiconvex in  $G/\mathcal{N}$ ;
- (3)  $\phi(g)$  is not contained in  $\phi(H)$ ; and
- (4) the height of  $\phi(H)$  in  $G/\mathcal{N}$  is strictly less than the height of  $H$  in  $G$ .

The expression ‘‘sufficiently long’’ just means that every nontrivial element of every  $P_i''$  should be sufficiently long (in the word metric in  $G$ ), which is just to say that finitely many (‘‘short’’) elements should not be allowed in  $P_i''$ .

One would like to apply this proposition repeatedly to the result until the height of  $H$  goes to 0, which is exactly the conclusion of Theorem 5.9. The problem is that we do not know that  $H$  admits peripherally finite fillings, since we do not know that the  $\mathcal{P}_i''$  are residually finite. Since the  $\mathcal{P}_i''$  are virtually subgroups of  $H$  it suffices to know that  $H$  itself is residually finite. The hypothesis of Theorem 5.9 is that  $H$  is isomorphic to the fundamental group of a virtually special NPC cube complex; this implies that *at the first step*  $H$  is linear and therefore residually finite. But *a priori* we do *not* know that  $\phi(H)$  (as in the conclusion of Proposition 5.14) is linear, so the induction cannot be continued without more work.

**5.3. The malnormal special quotient theorem.** At this point one must appeal to a substantial black box, the *Malnormal Special Quotient Theorem* of Wise (in the sequel we refer to this as the MSQT). The statement of the theorem is as follows:

**Theorem 5.15** (Wise; Malnormal Special Quotient Theorem). *Let  $G$  be hyperbolic, and let  $\mathcal{P} := \{P_i\}$  be a family of almost malnormal and quasiconvex subgroups so that  $(G, \mathcal{P})$  is relatively hyperbolic. Suppose  $G$  is the fundamental group of a virtually special NPC cube complex. Then there are finite index subgroups  $P'_i$  in the  $P_i$  so that if  $\phi : G \rightarrow G/\mathcal{N}$  is any peripherally finite filling with each  $N_i$  contained in  $P'_i$ , then  $\phi(G)$  is the fundamental group of a virtually special NPC cube complex.*

Now it is clear how to combine the MSQT with Proposition 5.14 to complete the induction and prove Theorem 5.9. Under the assumption that  $H$  is (virtually) the fundamental group of a special NPC cube complex, we may find finite index subgroups of the  $\mathcal{P}'$  so that any  $H$ -filling with  $N_i$  contained in these finite index subgroups will have the property that  $\phi(H)$  is again isomorphic to the fundamental group of a virtually special NPC cube complex. Thus one may inductively reduce the height of  $H$  to zero (by repeated  $H$ -filling) and obtain Theorem 5.9 as desired.

*Remark 5.16.* The MSQT is a truly remarkable theorem, in that it combines the combinatorial ‘‘specialness’’ of the special NPC property for cube complexes, with the flexibility

and robustness of hyperbolic Dehn surgery. The proof is not easy, and takes up much of the technical part of Wise's substantial preprint [44]. On the other hand, the virtual specialness of the groups  $\phi(H)$  at each stage is only used for the relatively weak conclusion that these groups are residually finite. Is it possible to prove Theorem 5.9 under the weaker hypothesis that  $H$  is residually finite? Can Proposition 5.14 be strengthened to add the conclusion that  $\phi(H)$  can be taken to be residually finite if  $H$  is?

Wise's proof of the MSQT is complicated, and expressed largely in combinatorial language. Agol-Groves-Manning have developed a new proof of the MSQT (borrowing very heavily from Wise's argument) which is expressed in more geometric language (their preprint is not yet available, but they have given several lectures explaining the main idea). Both proofs depend on an inductive characterization of hyperbolic virtually special groups as those that admit malnormal quasiconvex virtual hierarchies, a subject we shall discuss in more detail in the next section. Roughly speaking, a malnormal quasiconvex virtual hierarchy is an inductive description of a (hyperbolic) group as an iterated amalgam or HNN extension over simpler groups, where at each stage the edge groups are quasiconvex and malnormal in the amalgam. Haglund-Wise (building on Hsu-Wise) show that hyperbolic groups with such hierarchies are virtually special (and conversely); we outline their argument in § 6.2.

The very crudest sketch one can give of the proof of the MSQT is to say that one first needs to adjust the given malnormal quasiconvex virtual hierarchy of  $G$  so that the pieces interact in a geometrically controlled way with the subgroups  $P_i$ . This can be accomplished by a kind of "engulfing" argument, so that one pushes the pieces of the hierarchy over  $P_i$ . Then when one does relatively hyperbolic Dehn filling, one can ensure that the pieces of the (virtual) hierarchy of  $G$  stay quasiconvex and malnormal (in their respective summands) after filling, so that they give rise to a malnormal quasiconvex virtual hierarchy of  $\phi(G)$ . I learned this sketch from Daniel Groves.

## 6. $\mathcal{MVH}$ AND $\mathcal{QVH}$

**6.1. Hierarchies.** The statement and the proof of the MSQT (i.e. Theorem 5.15) uses a different characterization of hyperbolic groups that are the fundamental groups of virtually special cube complexes. The characterization is in terms of an *inductive* definition, which is therefore well suited to arguments to prove that certain groups are in this class. There are two, *a priori* different, versions of this definition; and it is the MSQT itself that is used to prove the equivalence.

The following definition is due to Wise [44], Def. 11.5:

**Definition 6.1** ( $\mathcal{MVH}$  and  $\mathcal{QVH}$ ). A hyperbolic group  $G$  has a *quasiconvex virtual hierarchy* (we say  $G$  is in  $\mathcal{QVH}$ ) if it is obtained inductively by the following operations:

- (1) the trivial group is in  $\mathcal{QVH}$ ;
- (2) if  $G = A *_B C$  where  $A, C \in \mathcal{QVH}$  and  $B$  is finitely generated and quasiconvex in  $G$ , then  $G$  is in  $\mathcal{QVH}$ ;
- (3) if  $G = A *_B$  where  $A \in \mathcal{QVH}$  and  $B$  is finitely generated and quasiconvex in  $G$ , then  $G$  is in  $\mathcal{QVH}$ ; and
- (4) if  $H$  is finite index in  $G$ , and  $H \in \mathcal{QVH}$  then  $G \in \mathcal{QVH}$ .

A hyperbolic group  $G$  has a *malnormal quasiconvex virtual hierarchy* (we say  $G$  is in  $\mathcal{MVH}$ ) if it has a quasiconvex virtual hierarchy, and if at each stage of the hierarchy the edge group  $B$  is malnormal in  $A$  and  $C$  (or in  $A$  in the case of an HNN extension).

Evidently from the definition  $\mathcal{MVH}$  implies  $\mathcal{QVH}$ . One of the main applications of the MSQT is the proof of the converse.

*Remark 6.2.* If  $G = A *_B C$  or  $G = A *_B$  where  $A, C$  are hyperbolic, and  $B$  is quasiconvex and almost malnormal in  $G$ , then  $G$  is hyperbolic. This follows from the Bestvina-Feighn combination theorem [6].

*Example 6.3.* Finite groups are in  $\mathcal{QVH}$ . Free groups and closed surface groups with  $\chi < 0$  are in  $\mathcal{QVH}$ . A Haken 3-manifold has a fundamental group which is  $\mathcal{QVH}$  if its fundamental group is hyperbolic, and if the decomposing surface subgroups at each step of the hierarchy are quasiconvex. So, for example, a closed hyperbolic 3-manifold with an embedded essential surface that is not the fiber of a fibration has a fundamental group which is  $\mathcal{QVH}$ .

*Example 6.4.* Wise [44], Thm. 18.1 shows that a 1-relator group with torsion is  $\mathcal{QVH}$ . 1-relator groups were already known to have a certain kind of hierarchy, called the *Magnus-Moldavanskii hierarchy*. Wise shows this hierarchy is quasiconvex whenever the 1-relator group has torsion.

Wise's main motivation to study these groups is that among hyperbolic groups, they precisely capture the class of groups which are virtually special. The following is proved for torsion-free groups in [44], Thm. 13.3, and without the torsion-free assumption in the appendix of [4]:

**Theorem 6.5** ( $\mathcal{QVH}$  is virtually special). *A hyperbolic group is in  $\mathcal{QVH}$  if and only if it has a finite index subgroup which is isomorphic to the fundamental group of a special cube complex.*

Informally, we say that a hyperbolic group is  $\mathcal{QVH}$  if and only if it is *virtually special*. We will show how to deduce Theorem 6.5 from the corresponding theorem for groups with a *malnormal* quasiconvex virtual hierarchy, given the MSQT.

**6.2.  $\mathcal{MVH}$  is virtually special.** An intermediate step to the proof of Theorem 6.5 is to prove it for groups in  $\mathcal{MVH}$ . This is accomplished by combining work of Haglund-Wise with (unpublished) work of Hsu-Wise.

First we quote the following result of Haglund-Wise [22], Thm. 1.2 which is used as a key ingredient:

**Theorem 6.6** (Combination theorem). *Let  $Y$  and  $Y'$  be compact virtually special cube complexes with hyperbolic fundamental groups. Let  $M \rightarrow Y$  and  $M \rightarrow Y'$  be locally isometric inclusions of a cube complex such that  $\pi_1(M)$  is quasiconvex and malnormal in  $\pi_1(Y)$  and  $\pi_1(Y')$ . Then the cube complex  $X := Y \cup_{M \times [-1,1]} Y'$  is virtually special.*

*Proof.* The first (and in a way, the main) step is to show that for any finite covers of  $Y$  and  $Y'$ , there are further *regular* covers which restrict to *the same* covers on  $M$ . This is

accomplished by means of the canonical completion and retraction operation (i.e. Theorem 2.26), which is morally a generalization of Scott's "engulfing" method to show that surface groups are LERF. This step accomplishes part of the claim, since one can find a finite cover of  $Y$  or  $Y'$  extending any given finite cover of  $M$ . However, *a priori* such a cover need not be *regular*; in particular, different preimages of  $M$  in the cover might cover  $M$  in different ways. The key is to be able to find a cover in which the distinct preimages of  $M$  do not "interfere" with each other, so that the canonical completion and retraction associated to each lift has a further compatible regular cover which restricts to the given cover on each preimage of  $M$ . It is here that malnormality plays a role: under the hypothesis of malnormality, we can find covers in which the *wall projection* of each preimage of  $M$  to each other is trivial. If  $A$  and  $B$  are subcomplexes of  $X$ , the *wall projection* of  $A$  to  $B$  is equal to  $B^0$  together with the union of all cubes of  $B$  whose 1-cubes are parallel to cubes of  $A$ ; it is *trivial* if every closed loop in the projection is homotopically trivial in  $X$ . If wall projections are trivial, a suitable cover can be guaranteed to have a controlled effect on each preimage of  $M$ . Thus we can assemble the covers of  $Y$  and  $Y'$  along the covers of  $M$  to produce a cover of  $X$ .

The second step is to observe that every quasiconvex subgroup of  $X$  is separable. For, such subgroups can be separated by first restricting to each side  $Y$  and  $Y'$ , separating the restrictions using virtual specialness of  $Y$  and  $Y'$ , and then gluing up the result using the first claim. This lets us pass to a finite cover of  $X$  in which every hyperplane is embedded. Since  $X$  is NPC, self-oscultations or inter-oscultations are associated to loops which are essential in  $\pi_1$ , and which can be busted by passing to a further finite cover (using separability of the hyperplanes, as above). This exhibits a finite cover of  $X$  which is special, as claimed.  $\square$

*Remark 6.7.* A similar result holds for HNN extensions, with a similar proof.

From Theorem 6.6 one expects to be able to conclude (by induction) that any hyperbolic group with a malnormal quasiconvex virtual hierarchy is virtually special. This is almost true, but one must be careful, since even if  $G = A *_B C$  where  $A, C \in \mathcal{MVH}$  and  $B$  is finitely generated, quasiconvex and malnormal in  $G$  (so that  $B$  is quasiconvex in  $A$  and  $C$  and therefore also in  $\mathcal{MVH}$ ) it is not *a priori* obvious that  $B$  can be cubulated consistently with cubulations of  $A$  and  $C$ , in order to be able to apply Theorem 6.6. This is exactly what Hsu-Wise prove. Explicitly, they show ([24], Theorem A):

**Theorem 6.8** (Hsu-Wise). *Let  $G$  be hyperbolic, and split as  $G = A *_B C$ , or as  $A *_B$  where  $A$  and  $C$  are virtually special hyperbolic, and  $B$  is quasiconvex and malnormal in  $G$ . Then  $G$  is cubulated.*

From these two theorems together one deduces that any hyperbolic group  $G$  in  $\mathcal{MVH}$  is virtually special.

**6.3.  $\mathcal{QVH}$  is virtually special.** We now sketch the proof of Theorem 6.5 (I learned this argument from Daniel Groves):

*Proof.* One direction of the argument is easy. Suppose  $Y$  is a special cube complex with hyperbolic fundamental group. Then the hyperplanes are embedded and quasiconvex. Let  $Z$  be a hyperplane. Then  $Y$  can be cut along  $Z$ , and  $\pi_1(Y)$  can be expressed either as an

amalgamated product or an HNN extension over  $\pi_1(Z)$ , with vertex groups isomorphic to the fundamental groups of special cube complexes with fewer cubes of maximal dimension.

We now indicate how to show that a hyperbolic  $\mathcal{QVH}$  group is virtually special. So let's suppose (for simplicity) that  $G = A *_C B$  where (by induction),  $A, B, C$  are virtually special, and  $C$  is quasiconvex. We claim that it suffices to show that  $C$  is separable in  $G$ . To see this, first note that if  $C$  were malnormal, we would be done by Theorem 6.6 and Theorem 6.8. If  $C$  is not malnormal, it is nevertheless true that it has finite height. So there are only finitely many cosets  $g_i C$  that are close to  $C$  on infinite subsets that include  $\text{id}$  (say), and if we can pass to a finite index subgroup  $G'$  containing  $C$  but not containing any  $g_i$  we will get a splitting of  $G'$  over copies of  $C$  where now the copies of  $C$  are malnormal.

So it suffices to show that  $C$  is separable in  $G$ . Let's show how to separate  $C$  from some element  $g$  using the MSQT. By Proposition 5.11 we can find finitely many  $g_i$  with  $D := C \cap C^{g_1} \cap \dots \cap C^{g_k}$  infinite and malnormal. By the MSQT we can find a further  $D'$  of finite index in  $D$  and separate from  $g$  so that the hypotheses of the MSQT apply to each pair  $(A, D')$ ,  $(B, D')$  separately, and the hypotheses of Proposition 5.14 apply to the pair  $(G, D')$ . The result is a group  $G' = A' *_C B'$  where  $C' = C/D'$  is separate from the image  $g'$  of  $g$ , and is quasiconvex of smaller height than  $C$ . Moreover,  $A'$  and  $B'$  are still virtually special (by the MSQT) so by induction on the height of the amalgamating subgroup, we can conclude that  $G'$  is virtually special too. But this means that  $g'$  can be separated from  $C'$  in  $G'$ , and therefore  $g$  can be separated from  $C$  in  $G$ . Since  $g$  was arbitrary,  $C$  is separable in  $G$ , so  $G$  is virtually special.  $\square$

## 7. PROOF OF THE VHC

**7.1. Quotient with compact hyperplanes.** We now have essentially all the necessary background to prove Agol's theorem. Agol proves his theorem in the following form:

**Theorem 7.1** (Virtually special; Agol [4], Thm. 1.1). *Let  $G$  be a hyperbolic group which acts properly and cocompactly on a CAT(0) cube complex  $X$ . Then  $G$  has a finite index subgroup  $G'$  so that  $X/G'$  is a special cube complex.*

We point out that Theorem 7.1 implies the Virtual Haken Theorem, in combination with Corollary 4.2, Theorem 4.1 and Theorem 2.27.

The next few sections are devoted to an exposition of Agol's proof of Theorem 7.1, following [4].

If  $G$  is torsion free, the quotient  $Y := X/G$  is an NPC cube complex; otherwise we can think of the quotient as an "orbifold" NPC cube complex, which adds technical complications. The hyperplanes of  $Y$  are *a priori* immersed; the goal is to show that there is a finite cover in which they are embedded. Note that the hyperplanes of  $Y$  are NPC cube complexes of strictly smaller dimension than that of  $X$ , so by induction we may assume that the fundamental group of each hyperplane is virtually special. Let  $W_1, \dots, W_m$  be  $G$ -orbit classes of hyperplanes in  $X$ , and let  $H_1, \dots, H_m$  be  $G$ -conjugacy classes of hyperplane stabilizers, so  $W_i/H_i$  is a virtually special NPC (orbifold) cube complex for each  $i$ .

By the Weak Separation Theorem (i.e. Theorem 5.9) there is a homomorphism  $\phi : G \rightarrow \mathcal{G}$  (where  $\mathcal{G}$  is *not* assumed to be finite) so that  $\phi(H_i)$  is finite for each  $i$  and so that

$\phi(H_i)$  is separated from any given finite collection of elements not contained in the  $H_i$ . Geometrically,  $\mathcal{X} := X/\mathcal{G}$  is an NPC cube complex (which is *not* assumed to be compact) which is a regular cover of  $Y$  such that every hyperplane is *compact*. We can further assume (by separating the  $H_i$  from finitely more elements if necessary) that the hyperplanes in  $\mathcal{X}$  are two-sided, and do not self-osculte or inter-osculte. In fact, we may assume if we like that for any fixed  $R$ , the  $R$ -neighborhood of each hyperplane  $W_i/H_i$  in  $X/H_i$  embeds in  $\mathcal{X}$ . We choose  $R$  sufficiently large so that if any two hyperplanes  $W, W'$  in  $X$  have  $|H \cap H'| = \infty$  (where  $H, H'$  are the stabilizers in  $G$  of  $W, W'$  respectively) then  $W$  and  $W'$  contain points which are distance at most  $R$  apart. Since  $X$  is  $\delta$ -hyperbolic for some  $\delta$ , such an  $R$  exists.

Form a graph  $\Gamma$  with vertices  $V(\Gamma)$  corresponding to the hyperplanes of  $\mathcal{X}$ , and edges  $E(\Gamma)$  corresponding to pairs of hyperplanes in  $\mathcal{X}$  which contain points at distance  $\leq R$ . Notice that  $\mathcal{G}$  acts simplicially on  $\Gamma$ .

**7.2. Invariant coloring measures.** We follow Agol [4], § 5 very closely.

Let  $\Gamma$  be a graph with vertex set  $V$  and with bounded valence  $\leq k$ , and let  $G$  be a group acting simplicially and cocompactly on  $\Gamma$ . We assume  $\Gamma$  has no loops or multiple edges between pairs of vertices, although  $\Gamma/G$  might have. Denote the vertices of  $\Gamma$  by  $V(\Gamma)$  and the edges by  $E(\Gamma)$ , so  $E(\Gamma)$  is a symmetric subset of  $V(\Gamma) \times V(\Gamma) - \Delta$ .

For each integer  $n$ , let  $\hat{C}_n(\Gamma)$  be the set of functions from  $V(\Gamma)$  to a set with  $n$  elements, and let  $C_n(\Gamma)$  be the subset of functions which take distinct values on the endpoints of every edge. Informally, we can say that  $C_n(\Gamma)$  is the set of “colorings” of  $\Gamma$  with at most  $n$  colors, in the usual sense of graph theory. For any  $n$  we topologize  $\hat{C}_n(\Gamma)$  with the product topology, as a product of  $V(\Gamma)$  copies of a discrete  $n$ -element set. Thus  $\hat{C}_n(\Gamma)$  is compact, and homeomorphic to a Cantor set for any  $n > 1$ . The subset  $C_n(\Gamma)$  is closed and therefore compact; the easiest way to see this is to observe that a sequence of functions  $c_i$  from  $V(\Gamma)$  to  $\{1, \dots, n\}$  converges if it is eventually constant on any finite set. If  $c_i$  is an infinite sequence of elements in  $C_n(\Gamma)$ , some subsequence must converge on finite subsets of  $V(\Gamma)$ , and the limit is therefore also in  $C_n(\Gamma)$ .

The group  $G$  acts on  $\hat{C}_n(\Gamma)$  by homeomorphisms, and preserves the subspace  $C_n(\Gamma)$ . If  $X$  is a compact Hausdorff  $G$ -space, we denote by  $\mathcal{M}(X)$  the space of probability measures on  $X$ , and by  $\mathcal{M}_G(X)$  the subspace of  $G$ -invariant probability measures, with the weak\* topology. Note that  $\mathcal{M}(X)$  is convex, compact and metrizable, and  $\mathcal{M}_G(X)$  is a convex, compact subset. Further, if  $Y \subset X$  is compact and  $G$ -invariant,  $\mathcal{M}(Y)$  is a convex, compact subset of  $\mathcal{M}(X)$ . Probability measures in  $\mathcal{M}(X)$  are Radon measures. The cases we have in mind are  $X = \hat{C}_n(\Gamma)$  and  $Y = C_n(\Gamma)$ .

**Proposition 7.2** (Invariant coloring measure). *With notation as above,  $\mathcal{M}_G(C_{k+1}(\Gamma))$  is nonempty, where  $k$  is a bound on the valence of  $\Gamma$ .*

*Proof.* Let  $\nu$  be a  $G$ -invariant probability measure on  $\hat{C}_n(\Gamma)$ . Since  $G$  acts cocompactly on  $\Gamma$ , we can choose a finite set of orbits  $e_1, \dots, e_m$  amongst the set of edges. We define the *weight* of  $\nu$ , denoted  $\text{weight}(\nu)$ , to be the expected number of  $e_i$  such that a  $\nu$ -random function  $c \in \hat{C}_n(\Gamma)$  has the same color at both endpoints of  $e_i$ . In other words, if  $B_e$  denotes the subset of  $\hat{C}_n(\Gamma)$  consisting of functions  $c$  which take the same values on the endpoints

of  $e$ , we define

$$\text{weight}(\nu) = \sum \nu(B_{e_i})$$

We remark that each  $B_e$  is *open and closed*, since membership in  $B_e$  is determined by only finitely many values. It follows that the indicator functions of the  $B_{e_i}$  are continuous, and therefore that weight is continuous as a function of  $\nu$  in  $\mathcal{M}(\hat{C}_n(\Gamma))$ . Note that since  $\nu$  is  $G$ -invariant,  $\text{weight}(\nu)$  does not depend on the choice of representatives  $e_i$ . Evidently if  $\nu \in \mathcal{M}_G(C_n(\Gamma))$  then the support of  $\nu$  is disjoint from every  $B_e$ , and therefore  $\text{weight}(\nu) = 0$ . Conversely, suppose  $\text{weight}(\nu) = 0$ . The *support* of  $\nu$  is the smallest compact subset of  $\hat{C}_n(\Gamma)$  of full  $\nu$ -measure. Any open set with  $\nu$  measure 0 is in the complement of the support. Hence  $B_{e_i}$  is in the complement of the support for every  $i$ . But since  $\nu$  is  $G$ -invariant,  $B_e$  is in the complement of the support for every  $e$ , and therefore the support of  $\nu$  is contained in  $\hat{C}_n(\Gamma) - \cup_e B_e = C_n(\Gamma)$  by definition. Since  $\nu$  is  $G$ -invariant,  $\nu \in \mathcal{M}_G(C_n(\Gamma))$ . So to prove the proposition it suffices to find a  $G$ -invariant probability measure of weight 0.

The first step is to find measures in  $\mathcal{M}_G(\hat{C}_n(\Gamma))$ . There is one obvious measure, namely the *uniform measure*  $\mu_n$ , which is the product of the uniform probability measures on each copy of  $\{1, \dots, n\}$  in the product  $\hat{C}_n(\Gamma)$ . In the uniform measure,  $\mu_n(B_e) = 1/n$  for every edge  $e$ , so  $\text{weight}(\mu_n) = m/n$ .

Now, for every  $n > k+1$ , there is a  $G$ -equivariant measurable map  $p_n : \hat{C}_n(\Gamma) \rightarrow \hat{C}_{n-1}(\Gamma)$  defined as follows. Let  $c \in \hat{C}_n(\Gamma)$  be a function from  $V(\Gamma)$  to  $\{1, \dots, n\}$ . Define  $p_n(c)(v) = c(v)$  if  $c(v) < n$ , and if  $c(v) = n$  define  $p_n(c)(v)$  to be the smallest number in  $\{1, \dots, n-1\}$  which is not among the values of  $c(v')$  on the neighbors  $v'$  of  $v$ . Because by hypothesis the valence of  $v$  is at most  $k$  which is strictly less than  $n-1$ , such a value exists. This map has the following key property: if  $v, v'$  are adjacent and  $p_n(c)(v) = p_n(c)(v')$  then  $c(v) = c(v')$ . This means that if  $p_n(c) \in B_e(n-1)$  then  $c \in B_e(n)$  (by abuse of notation). If  $\mu$  is a measure in  $\mathcal{M}_G(\hat{C}_n(\Gamma))$ , then  $(p_n)_*\mu$  is a measure in  $\mathcal{M}_G(\hat{C}_{n-1}(\Gamma))$ , and by the key property,  $\text{weight}((p_n)_*\mu) \leq \text{weight}(\mu)$ . Define  $\nu_n$  to be the measure in  $\mathcal{M}_G(\hat{C}_{k+1}(\Gamma))$  obtained from  $\mu_n$  by pushing forward by  $p_{k+1} \circ \dots \circ p_{n-1} \circ p_n$ . Then  $\text{weight}(\nu_n) \leq m/n$ . By compactness, some subsequence of the  $\nu_n$  converges in  $\mathcal{M}_G(\hat{C}_{k+1}(\Gamma))$  to some limit  $\nu_\infty$ . By continuity of weight we have  $\text{weight}(\nu_\infty) = 0$  so  $\nu_\infty \in \mathcal{M}_G(C_{k+1}(\Gamma))$ , as desired.  $\square$

We apply this theorem to the graph  $\Gamma$  with its  $\mathcal{G}$  action, whose vertices correspond to the hyperplanes in  $\mathcal{X}$  and whose edges correspond to pairs of hyperplanes coming within distance  $R$  of each other. We deduce there is a  $\mathcal{G}$ -invariant probability measure  $\nu$  on colorings of  $\Gamma$  (using only finitely many colors).

*Remark 7.3.* Suppose the support of  $\nu$  consisted of a finite number of atoms. Then some finite index subgroup  $\mathcal{G}'$  of  $\mathcal{G}$  would fix an atom. This atom would be a  $\mathcal{G}'$ -invariant *coloring*. It would follow that the  $R$ -neighborhood of every hyperplane in  $\mathcal{X}$  would map to  $\mathcal{X}/\mathcal{G}'$  by an embedding; in particular, hyperplanes in  $\mathcal{X}/\mathcal{G}'$  would be two-sided, embedded and without self- or inter-osculation; i.e. this quotient would be special, and we would be done.

**7.3. Boundary patterns.** Consider the result of splitting  $\mathcal{X}$  along its hyperplanes. One obtains a union of compact pieces, where each piece is the star on a vertex of  $\mathcal{X}$ . The pieces

themselves are cubical polyhedra, where the cubes in question have side lengths half as long as the cubes in  $\mathcal{X}$ .

The cubical polyhedra we obtain have a *boundary pattern*, which is a decomposition of its “boundary” into locally convex subcomplexes  $\{\partial_1, \dots, \partial_n\}$  which inductively have the structure of a cubical polyhedron with boundary pattern of its own, where the boundary pattern in the “boundary” of  $\partial_i$  consists of the collection of nonempty  $\partial_i \cap \partial_j$ . The meaning of “boundary” here is simply that each  $\partial_i$  should have an open product collar neighborhood in the cubical polyhedron. The motivating example is the pieces into which a Haken 3-manifold is decomposed by cutting along a hierarchy. We call the  $\partial_i$  the *facets* of the cubical polyhedron.

Suppose  $X$  is a cubical polyhedron with boundary pattern (possibly disconnected), and  $\partial_n$  is one of the facets. Given an isometric involution  $\tau : \partial_n \rightarrow \partial_n$  without fixed points, and such that  $\tau(\partial_i \cap \partial_n) = \partial_i \cap \partial_n$  for each  $i$ , we can glue  $\partial_n$  to itself by  $\tau$  to obtain a new cubical polyhedron, with an induced boundary pattern where the new facets  $\partial'_i$  are the images of the old facets  $\partial_i$  under the quotient map.

Let  $c$  be a coloring on  $V(\Gamma)$  with  $k+1$  colors. Remember that the vertices of  $\Gamma$  correspond to the hyperplanes of  $\mathcal{X}$ .

**Definition 7.4** (Supercolor). Two colors  $c, c'$  induce the same *supercolor* on a vertex  $v$  if they satisfy the following inductive definition:

- (1) if  $c(v) = 1$  and  $c'(v) = 1$ ; or
- (2) if  $c(v) = c'(v) > 1$  and, for all neighbors  $v'$  of  $v$  with  $c(v') < c(v)$ , the supercolors of  $v'$  induced by  $c$  and  $c'$  agree.

If we think of a color as a (discrete) Morse function on  $\Gamma$ , two colors induce the same supercolor at a vertex  $v$  if they have the same descending submanifold from  $v$ , with the same values on it. Notice that the set of possible supercolors is *finite*, since a supercolor at  $v$  is determined by the values of a color on a ball around  $v$  of radius at most  $k$ .

If we cut up  $\mathcal{X}$  along all its hyperplanes, we obtain a collection of cubical polyhedra with facets. Each facet inherits a supercolor from the hyperplane that contained it; a *supercolored polyhedron* is a  $\mathcal{G}$ -equivalence class of polyhedron together with a collection of supercolors on its facets, and a *supercolored facet* is a  $\mathcal{G}$ -equivalence class of facets together with a choice of supercolor.

**7.4. Gluing equations.** Let  $\mathcal{P}$  denote the finite set of supercolored polyhedra, and  $\mathcal{F}$  the finite set of supercolored facets. Let  $\mathbb{R}[\mathcal{P}]$  be the vector space spanned by the set  $\mathcal{P}$ , and  $\mathbb{R}[\mathcal{F}]$  the vector space spanned by the set  $\mathcal{F}$ . If we choose a co-orientation for each facet in  $\mathcal{F}$ , there is an integral linear map  $\partial : \mathbb{R}[\mathcal{P}] \rightarrow \mathbb{R}[\mathcal{F}]$  taking each supercolored polyhedron to a signed sum of its supercolored facets, where the sign is  $\pm 1$  according to whether the induced co-orientation agrees or disagrees with the global choice.

We say that a vector  $p \in \mathbb{R}[\mathcal{P}]$  *satisfies the gluing equations* if  $\partial p = 0$ . This is a finite rational linear system of equalities, so if there is a nonzero, non-negative solution  $p$ , there is a nonzero, non-negative rational solution, and (by clearing denominators by scaling) a nonzero non-negative integral solution.

Now, let  $\nu$  denote the  $\mathcal{G}$ -invariant probability measure on  $C_{k+1}(\Gamma)$ . For each vertex  $v$  and each supercolor  $s$ , the probability that a  $\nu$ -random color induces the supercolor  $s$  on  $v$

depends only on the  $\mathcal{G}$ -orbit of  $v$ . If we cut up  $\mathcal{X}$  into cubical polyhedra, then for each facet and each supercolor there is a well-defined probability that a  $\nu$ -random color induces the given supercolor on the facet. Similarly, for each cubical polyhedra, there is a well-defined probability that a  $\nu$ -random color induces a given collection of supercolors on the facets of the polyhedron. Since  $\nu$  is  $\mathcal{G}$ -invariant, this vector of probabilities only depends on the  $\mathcal{G}$  orbit of the polyhedron, so we obtain a well-defined vector  $p_\nu$  in  $\mathbb{R}[\mathcal{P}]$ . This vector is nonzero and non-negative, and by construction it satisfies the gluing equations. So we conclude that there is a nonzero non-negative integral solution to the gluing equations  $p_{\mathbb{Z}}$  in  $\mathbb{R}[\mathcal{P}]$ . This vector defines a (disconnected) cube complex with (supercolored) boundary pattern, by simply taking a disjoint union of (supercolored) cubical polyhedra, where the number of each kind of polyhedron is given by the corresponding coefficient of  $p_{\mathbb{Z}}$ . We denote this polyhedron  $\mathcal{V}_{k+1}$ . In the next section we explain how to *virtually* glue up  $\mathcal{V}_{k+1}$  by pairing boundary facets inductively to obtain a cube complex  $\mathcal{V}_0$  which will be special, and which will finitely cover  $X$ . The gluing is only virtual, since at each stage we might need to take a finite cover of the complex already obtained before being able to glue up to obtain the next stage.

**7.5. The complexes  $\mathcal{Y}_j$ .** We now have a finite collection  $\mathcal{V}_{k+1}$  of supercolored cubical polyhedra satisfying the gluing equations, and we would like to glue them up along supercolored facets to build a complex finitely covering  $X/G$  and with embedded hyperplanes. Note that  $\mathcal{V}_{k+1}$  is special (actually, each component is contractible).

We think of  $\mathcal{V}_{k+1}$  as a cubical polyhedron with a boundary pattern

$$\{\partial_1(\mathcal{V}_{k+1}), \dots, \partial_{k+1}(\mathcal{V}_{k+1})\}$$

where  $\partial_i$  is the union of the facets colored (not supercolored)  $i$ . Note that each component of  $\partial_i$  consists of a single facet, since adjacent facets of  $\mathcal{V}_{k+1}$  correspond to intersecting hyperplanes of  $\mathcal{X}$ , which must be colored differently.

The first step of the gluing is completely straightforward: for each class of facet with a given supercolor inducing the color  $k+1$ , there are an equal number in the boundary of our polyhedra with each co-orientation. We may therefore glue them (arbitrarily) in pairs, with opposite co-orientations to produce  $\mathcal{V}_k$ . Note that  $\mathcal{V}_k$  is virtually special, since it is built from  $\mathcal{V}_{k+1}$  and we can invoke Theorem 6.5. Since gluing respects supercolors, and the supercolor of a facet determines the colors of lower values on adjacent facets, this gluing is compatible with supercolors, and  $\mathcal{V}_k$  inherits a boundary pattern  $\{\partial_1(\mathcal{V}_k), \dots, \partial_k(\mathcal{V}_k)\}$  where  $\partial_i(\mathcal{V}_k)$  is the union of facets colored  $i$ .

At the next step we would like to glue up  $\partial_k(\mathcal{V}_k)$ . Each  $\partial_k(\mathcal{V}_k)$  decomposes into facets colored  $k$ , and as before there are an equal number of each supercolor and each co-orientation. But the gluing at this second step must be done compatibly with the induced gluings on boundaries of facets coming from the identification of the facets supercolored  $k+1$ .

Since the gluing at the first step respected supercolors and  $\mathcal{G}$ -orbits, adjacent facets in each  $\partial_i(\mathcal{V}_k)$  correspond to adjacent  $\mathcal{G}$ -orbits in  $\mathcal{X}$ . We would like to construct a compact cube complex  $\mathcal{Y}_k$  and an immersion  $\partial_k \rightarrow \mathcal{Y}_k$  of “degree 0”; i.e. each cube in  $\mathcal{Y}_k$  is in the image of the same number of cubes in  $\partial_k$  with either co-orientation. The complex  $\mathcal{Y}_k$  is obtained as follows. Take a supercolored hyperplane  $H$  in  $\mathcal{X}$  with color  $k$ , and let  $H'$  be the result when  $\mathcal{X}$  is cut open along all hyperplanes of color  $\leq k$  (this makes sense, since the

supercolor of  $H$  determines precisely which hyperplanes intersecting  $H$  have color  $\leq k$ . For each component of  $H'$  we can take the quotient by its stabilizer in  $\mathcal{G}$ . The union, over all  $\mathcal{G}$ -orbits, is  $\mathcal{Y}_k$ . The key observation is that the  $\mathcal{G}$ -orbit of each facet in  $\partial_k(\mathcal{V}_k)$  corresponds to a unique facet of  $\mathcal{Y}_k$ . This is immediate from the definition, but it implies that there is a well-defined map  $\partial_k \rightarrow \mathcal{Y}_k$  taking each facet to the corresponding facet of  $\mathcal{Y}_k$ . In fact, this map is a *covering map*; this can be seen by first developing a map from the universal cover of a component of  $\partial_k$  to a component of  $H'$  (as above) and then projecting to  $\mathcal{Y}_k$ . The point is that the supercolor and the  $\mathcal{G}$ -orbit type of each facet determines the germ of the coloring (on cut open facets) near  $H'$  so non-trivial elements in  $\pi_1(\partial_k)$  determine elements in the stabilizer of  $H'$ .

We would now like to pass to a finite cover  $\hat{\mathcal{V}}_k$  of  $\mathcal{V}_k$  so that the induced covering map  $\hat{\partial}_k \rightarrow \mathcal{Y}_k$  is regular. Since  $\mathcal{V}_k$  is virtually special, this cover can be found by essentially the argument of Theorem 6.5. After passing to this cover, we can glue up pairs of components of  $\hat{\partial}_k$  mapping to the same component of  $\mathcal{Y}_k$  with opposite co-orientation, and obtain  $\mathcal{V}_{k-1}$ . We would like to conclude by Theorem 6.5 that  $\mathcal{V}_{k-1}$  is virtually special. For this we need to show that the inclusion of  $\partial_k$  in  $\mathcal{V}_k$  is *acylindrical*, and therefore that after gluing, the result will still be hyperbolic, and the amalgamating subgroup is quasiconvex (by Bestvina-Feighn). But recall that pairs of hyperplanes of  $X$  which contain points at distance  $\leq R$  must have distinct colors in any coloring, and  $R$  was chosen so that any two hyperplanes in  $X$  with an infinite common stabilizer should have points at distance  $\leq R$  apart. The core of an essential cylinder in  $\mathcal{V}_k$  would stabilize two distinct hyperplanes in  $X$ , so we conclude to the contrary that  $\partial_k$  is acylindrical in  $\mathcal{V}_k$ . Thus  $\mathcal{V}_{k-1}$  is virtually special.

Each subsequent step is essentially the same as the second step. Each  $\mathcal{Y}_j$  is obtained from a set of hyperplanes with supercolors inducing the color  $j$  by cutting them open along their intersection with hyperplanes of color  $< j$ , and taking the quotient by the stabilizer in  $\mathcal{G}$ . Each  $\partial_j$  in  $\mathcal{V}_j$  immerses to  $\mathcal{Y}_j$  with degree 0, and each  $\partial_j$  is acylindrical in  $\mathcal{V}_j$ . After passing to a finite cover we can ensure that  $\hat{\partial}_j \rightarrow \mathcal{Y}_j$  is regular, and can therefore be paired and glued up. At every stage  $\mathcal{V}_j$  is virtually special by Theorem 6.5 and induction. Finally we obtain  $\mathcal{V}_0$  which by construction is virtually special, and finitely covers  $X/G$ . This proves the Virtual Haken Theorem.

## 8. OTHER EXAMPLES OF CUBULATED GROUPS

Since virtually special groups have so many remarkable properties, and since cubulated (hyperbolic) groups are relatively straightforward to construct, Agol's theorem has spectacular implications throughout geometric group theory, many of which presumably remain to be worked out. In this section we briefly describe a few examples of classes of groups that are known to be cubulated, mainly by the work of Wise.

8.1.  $C'(1/6)$  **groups.** Let  $G$  be a group with finite presentation

$$G := \langle x_1, x_2, \dots, x_k \mid r_1, r_2, \dots, r_s \rangle$$

Where we can take the  $r_i$  to be cyclically reduced. A *piece* is a subword  $\sigma$  of some  $r_i$  that is also a subword of some other  $r_j$  or its inverse, or appears as a subword of  $r_i$  in a different place.

Recall that a group presentation is said to satisfy the condition  $C''(\lambda)$  (and one says  $G$  is  $C''(\lambda)$  if it admits such a presentation) if for every piece  $\sigma$  in some  $r_i$  one has  $\text{length}(\sigma)/\text{length}(r_i) < \lambda$ .

Let  $G$  be  $C''(1/6)$  with respect to some presentation, and let  $K$  be a presentation 2-complex; i.e. a complex with one vertex, one edge for each generator, and one 2-cell for each relator, in such a way that  $\pi_1(K) = G$ . The condition  $C''(1/6)$  ensures (by Gauss-Bonnet) that  $K$  is aspherical and hyperbolic.

Wise [43] shows how to cubulate such groups  $G$  as follows. Each cell  $D$  in  $K$  corresponds to a relator  $r$ , and can be thought of as a polygon with  $\text{length}(r)$  sides. For simplicity let's assume in the sequel that every  $\text{length}(r)$  is even (otherwise one can just declare that every edge of  $K$  has length 2 and proceed verbatim in what follows). Then every edge on  $D$  has a corresponding antipodal edge, and we can join pairs of antipodal edges by arcs in  $D$ . Take the disjoint union of all these edges over all disks  $D$  in  $K$ , and glue them by their endpoints, thus building a graph  $X$  together with an immersion  $X \rightarrow K$ . The graph  $X$  may have many components; each such component sits in  $K$  in much the same way that an immersed hyperplane sits in an NPC cube complex. Under the hypothesis that the presentation is  $C''(1/6)$ , the components of  $X$  give rise to codimension 1 subgroups of  $\pi_1(K)$ , and there are enough of them to cubulate.

Note in this case that the codimension 1 subgroups are all free (since they are  $\pi_1$  of graphs).

**8.2. Random groups at density  $< 1/6$ .** If we fix a free group  $F_k$  of rank  $k$ , and fix a free generating set, there are roughly  $(2k - 1)^n$  reduced words of length  $n$ . If we pick some  $0 < D < 1$ , a *random group at density  $D$  and length  $n$*  is a group with a presentation of the form

$$G := \langle F_k \mid r_1, r_2, \dots, r_s \rangle$$

where  $s = (2k - 1)^{nD}$ , and the relators  $r_i$  are all reduced words of length  $n$ , chosen independently and randomly with the uniform distribution. For a fixed density  $D$ , one considers random groups of length  $n$  and asks which properties hold with probability going to 1 as  $n \rightarrow \infty$ . Informally, one says that such a property holds for a random group at density  $D$  *with overwhelming probability*.

Gromov [18], § 9 introduced this model of a random group and proved several fundamental facts about them, including:

- (1) at density  $D > 1/2$ , a random group is either trivial or isomorphic to  $\mathbb{Z}/2\mathbb{Z}$  (wop);
- (2) at density  $0 < D < 1/2$ , a random group is non-elementary hyperbolic (wop); and
- (3) at density  $D$ , a random group (with its random presentation) satisfies  $C''(2D)$  but not  $C''(\lambda)$  for any  $\lambda < 2D$  (wop).

It follows that at density  $D < 1/12$ , a random group is  $C''(1/6)$  with overwhelming probability, and therefore is cubulated. In fact, Ollivier-Wise [30] showed that at density  $D < 1/6$  a random group is cubulated. The construction of the codimension 1 subgroups is essentially the same as that in § 8.1, but showing that the resulting subgroups are codimension 1, and that there are enough to cubulate, is more subtle.

*Remark 8.1.* At density  $1/3 < D < 1/2$  work of Zuk [45] (further elaborated by Kotowski-Kotowski [26]) shows that random groups have property (T) with overwhelming probability;

a group with property (T) cannot act without a global fixed point on a CAT(0) cube complex at all, so these groups should be thought of as being as “far from cubulated as possible”.

*Remark 8.2.* Calegari-Walker [12] show that at any density  $D < 1/2$  a random group contains (many) quasiconvex surface subgroups (wop). The construction has some interesting points of similarity to the method of Kahn-Markovic described in § 4, and it raises the question of whether similar methods might apply to all one-ended hyperbolic groups. It follows by Ollivier-Wise and Agol that at any density  $D < 1/6$  a random group virtually retracts to a surface group (wop).

**8.3. One-relator groups with torsion.** We recall the construction of the (so-called) Magnus-Moldovanskii hierarchy for a 1-relator group. When the group has torsion, it is hyperbolic, and the amalgamating subgroups are quasiconvex. Thus such groups are  $\mathcal{QVH}$  and therefore virtually special.

Fix a finite generating set  $a, b, c, \dots$  and a cyclically reduced primitive word  $w$ . We consider a group  $G$  with presentation

$$G := \langle a, b, c, \dots \mid w^n \rangle$$

for some  $n \geq 2$ .

**Definition 8.3.** An *appearance* of a generator  $x$  in  $w$  is a copy of  $x$  or  $x^{-1}$ . The *complexity* of  $G$  (with the given presentation) is equal to  $|w|$  minus the number of distinct generators that appear in  $w$ .

Thus the complexity is non-negative, and is equal to zero if and only if each generator that appears in  $w$  appears exactly once. Geometrically, if  $X$  is a rose for the given generators, and  $w$  is represented by an immersed loop  $\gamma \rightarrow X$ , complexity zero means that  $\gamma$  goes over each edge of  $X$  at most once.

Suppose there is some generator  $a$  that appears in  $w$  exactly once, so  $w = aw'$  where  $w'$  does not contain  $a$  or  $a^{-1}$ . Recall that the Tietze moves allow us to simultaneously add (or remove) a new generator together with a relation with exactly one appearance of that new generator.

We can rewrite the presentation by Tietze moves:

$$\langle a, b, c, \dots \mid w^n \rangle = \langle a, b, c, \dots, w \mid aw' = w, w^n \rangle = \langle b, c, \dots, w \mid w^n \rangle = F * \mathbb{Z}/n\mathbb{Z}$$

where  $F$  is free on  $b, c, \dots$ . In particular, if  $w$  has complexity 0, the group  $G$  can be written as a free product of a free group and a cyclic group.

**Proposition 8.4.** *Suppose  $G$  is a one-relator group, with a presentation as above. Then either the complexity is zero, or  $G$  splits nontrivially as an HNN extension over a one-relator group of strictly smaller complexity.*

We will see that if  $G$  has torsion the splitting subgroup in Proposition 8.4 is quasiconvex.

*Proof.* It turns out to be easier to construct a splitting of a free product  $G' := G * \mathbb{Z}$ . We let  $t$  denote the generator of the new  $\mathbb{Z}$  factor, so  $G' := \langle a, b, c, \dots, t \mid w^n \rangle$ . Let  $F$  be the free group generated by  $a, b, c, \dots$ , and let  $\phi : F \rightarrow \mathbb{Z}$  be a (nontrivial) homomorphism

such that  $\phi(w) = 0$ . For each generator  $x$  we define a new generator  $\bar{x}$  by  $x = \bar{x}t^{\phi(x)}$ . Rewriting  $w$  in terms of these new generators we get a new 1-relator presentation

$$G' := \langle \bar{a}, \bar{b}, \bar{c}, \dots, t \mid \bar{w}^n \rangle$$

The condition  $\phi(w) = 0$  means that  $\phi$  extends to a new homomorphism  $G' \rightarrow \mathbb{Z}$  sending  $t \rightarrow 1$  and  $\bar{x} \rightarrow 0$  for all  $\bar{x}$ .

Let  $X'$  be a rose for the generators and  $X$  a subrose for the generators excluding  $t$ . Let  $\hat{X}$  be the cyclic cover of  $X'$  associated to  $\phi$ , so that  $\hat{X}$  is made from  $\mathbb{Z}$  copies of  $X$  joined by arcs labeled  $t$ . We label the  $\mathbb{Z}$  copies of  $X$  as  $\dots, X_{-1}, X_0, X_1, \dots$ . The loop  $w$  lifts to a loop in  $\hat{X}$ . Once we pick a lift, each copy of  $\bar{x}$  or  $\bar{x}^{-1}$  in  $w$  lifts to some  $\bar{x}_i$  or  $\bar{x}_i^{-1}$  in  $X_i$ . For each  $\bar{x}$ , let  $L_x$  be the least index  $i$  and  $R_x$  the greatest index  $i$  occurring as above. After identifying  $X_0$  with  $X$ , we can identify  $\bar{x}_i = t^i \bar{x} t^{-i}$ . We therefore obtain a new presentation for  $G'$  in the form

$$G' := \langle t, \bar{a}_i \text{ for } L_a \leq i \leq R_a, \bar{b}_i \text{ for } L_b \leq i \leq R_b, \dots \mid \bar{v}^n, \bar{x}_{i+1}^t = \bar{x}_i \text{ for each } \bar{x}, i \rangle$$

where  $\bar{v}$  is obtained from  $\bar{w}$  by rewriting it in the generators  $\bar{x}_i$  (omitting  $t$ ). Thus  $G'$  is an HNN extension of the one-relator group

$$K := \langle \bar{a}_i, \bar{b}_i, \dots \mid \bar{v}^n \rangle$$

conjugating the free subgroup generated by the  $\bar{x}_i$  with  $i > L_x$  to the free subgroup generated by the  $\bar{x}_i$  with  $i < R_x$ .

Evidently  $\bar{v}$  has the same length as  $w$ , and its complexity is no bigger than that of  $w$ . In fact, the complexity will strictly *decrease* if there is some  $\bar{x}$  with  $L_x < R_x$ ; i.e. so that the lift of  $w$  to  $\hat{X}$  runs over at least two different  $\bar{x}_i$ . We show that this can always be arranged by suitable choice of  $\phi$ .

For each generator  $x$  let  $\#_x$  be the homomorphism from the free group on the generators to  $\mathbb{Z}$  sending  $x$  to 1 and all other generators to 0. Suppose there is a generator  $a$  with  $\#_a(w) = 0$ . If two successive appearances of  $a$  have the same sign, we can just take  $\phi = \#_a$ , since this will give rise to distinct  $\bar{a}_i, \bar{a}_{i+1}$ . So it must be that successive appearances of  $a$  alternate between  $a$  and  $a^{-1}$ , i.e.  $w = aA_1a^{-1}A_2aA_3a^{-1}A_4\dots$ .

Call a letter  $b$  a *zero letter* if  $\#_b(w) = 0$ . If some  $A_i$  consists entirely of zero letters, then for every innermost expression of the form  $b^{\pm 1}Cb^{\pm 1}$  in  $A_i$  with  $b^{\pm 1}$  not in  $C$ , either  $C$  is empty (in which case  $\#_b$  separates the adjacent  $bb$  or  $b^{-1}b^{-1}$ ) or there is  $c \in C$ , and  $\#_c$  separates the surrounding  $bs$ . So we can suppose that every  $A_i$  contains some nonzero letter.

Let  $b$  be a nonzero letter in  $A_2$  and  $c$  a nonzero letter in  $A_1$ . Note that each appearance of  $b$  is in an odd  $A_i$  and each appearance of  $c$  is in an even  $A_i$ , or else distinct appearances would be separated by  $\#_a$ . Let  $A_i$  be such that  $\#_b(A_i) \neq 0$ . Then  $\phi := \#_c(w)\#_b - \#_b(w)\#_c$  separates the adjacent  $as$  surrounding  $A_i$ . So we are reduced to the case that  $\#_a(w)$  is nonzero for *every*  $a$ .

Suppose some letter occurs with both signs. Then there is some innermost  $aBa^{-1}$ , and  $\#_b(B) \neq 0$  for any  $b$  in  $B$ , so  $\#_a(w)\#_b - \#_b(w)\#_a$  separates the two outer copies of  $a$ . We reduce finally to the case that each generator always appears positively. Write  $w = aA_1aA_2\dots$ . If  $\#_b(A_i) = 0$  for some  $b$  with  $\#_b(w) \neq 0$  then  $\#_a(w)\#_b - \#_b(w)\#_a$  separates the bounding  $a$ 's in  $aA_ia$ . If  $\#_b(A_i) \geq 2$  then  $\#_a(w)\#_b - \#_b(w)\#_a$  separates the

occurrences of  $b$  in  $A_i$ . So we must have  $\#_b(A_i) = 1$  for each  $b$ . This quickly implies that  $w = u^m$  for some  $u$  where each letter occurs exactly once in  $u$ . Since  $w$  is primitive by hypothesis,  $m = 1$  so the complexity of  $w$  is zero.  $\square$

## 9. ACKNOWLEDGMENTS

Danny Calegari was supported by NSF grant DMS 1005246. I would like to thank Ian Agol, Daniel Groves, Frédéric Haglund, Jason Manning, Vlad Markovic and Piotr Przytycki for helpful comments.

## REFERENCES

- [1] I. Agol, D. Long and A. Reid, *The Bianchi groups are separable on geometrically finite subgroups*, Ann. Math. (2) **153** (2001), no. 3, 599–621
- [2] I. Agol, *Tameness of hyperbolic 3-manifolds*, preprint; arXiv:math/0405568
- [3] I. Agol, *Criteria for virtual fibering*, J. Topol. **1** (2008), no. 2, 269–284
- [4] I. Agol (with an appendix by D. Groves and J. Manning), *The virtual Haken conjecture*, preprint; arXiv:1204.2810
- [5] I. Aitchison and H. Rubinstein, *An introduction to polyhedral metrics of non-positive curvature on 3-manifolds*, Geometry of low-dimensional manifolds, 2 (Durham, 1989), 127–161, LMS Lecture Note Ser. 151, Cambridge Univ. Press, Cambridge, 1990
- [6] M. Bestvina and M. Feighn, *A combination theorem for negatively curved groups*, J. Diff. Geom. **35** (1992), no. 1, 85–101
- [7] S. Bleiler and C. Hodgson, *Spherical space forms and Dehn filling*, Topology **35** (1996), 809–833
- [8] B. Bowditch, *Relatively hyperbolic groups*, preprint, 1999
- [9] M. Burger and S. Mozes, *Lattices in product of trees*, Publ. Math. IHES **92** (2000), 151–194
- [10] R. Burns, A. Karrass and D. Solitar, *A note on groups with separable finitely generated subgroups*, Bull. Aust. Math. Soc. **36** (1987), no. 1, 153–160
- [11] D. Calegari and D. Gabai, *Shrinkwrapping and the taming of hyperbolic 3-manifolds*, J. AMS **19** (2006), no. 2, 385–446
- [12] D. Calegari and A. Walker, *Random groups contain surface subgroups*, preprint, arXiv:1304.2188
- [13] R. Charney, *An introduction to right-angled Artin groups*, Geom. Ded. **125** (2007), 141–158
- [14] M. Davis and T. Januszkiewicz, *Right-angled Artin groups are commensurable with right-angled Coxeter groups*, J. Pure Appl. Algebra **153** (2000), 229–235
- [15] B. Farb, *Relatively hyperbolic groups*, Geom. Func. Anal. **8** (1998), no. 5, 810–840
- [16] R. Gitik, M. Mitra, E. Rips and M. Sageev, *Widths of subgroups*, Trans. AMS **350** (1998), no. 1, 321–329
- [17] M. Gromov, *Hyperbolic groups*, In *Essays in group theory*, volume 8 of Math. Sci. Res. Inst. Publ., 75–263. Springer, New York, 1987
- [18] M. Gromov, *Asymptotic invariants of infinite groups*, Geometric group theory, vol. 2, LMS Lecture Notes **182** (Niblo and Roller eds.) Cambridge University Press, Cambridge, 1993
- [19] D. Groves and J. Manning, *Dehn filling in relatively hyperbolic groups*, Israel J. Math. **168** (2008), 317–429
- [20] F. Haglund, *Finite index subgroups of graph products*, Geom. Ded. **135** (2008), 167–209
- [21] F. Haglund and D. Wise, *Special cube complexes*, Geom. Funct. Anal. **17** (2008), no. 5, 1551–1620
- [22] F. Haglund and D. Wise, *A combination theorem for special cube complexes*, Ann. of Math. **176** (2012), no. 3, 1427–1482
- [23] A. Hatcher, *Notes on basic 3-manifold topology*, preprint; available from the author’s webpage
- [24] T. Hsu and D. Wise, *Cubulating malnormal amalgams*, preprint; available from the authors’ webpages
- [25] J. Kahn and V. Markovic, *Immersing almost geodesic surfaces in a closed hyperbolic three manifold*, Ann. Math. (2) **175** (2012), no. 3, 1127–1190

- [26] M. Kotowski and M. Kotowski, *Random groups and property (T): Zuk's theorem revisited*, Jour. LMS (2013)
- [27] T. Li, *Boundary curves of surfaces with the 4-plane property*, Geom. Topol. **6** (2002), 609–647
- [28] S. Matsumoto, *Non-separable surfaces in cubed manifolds*, Proc. AMS **125** (1997), no. 11, 3439–3446
- [29] W. Neumann, *Commensurability and virtual fibration for graph manifolds*, Topology **36** (1997), no. 2, 355–378
- [30] Y. Ollivier and D. Wise, *Cubulating random groups at density less than 1/6*, Trans. AMS **363** (2011), no. 9, 4701–4733
- [31] D. Osin, *Peripheral fillings of relatively hyperbolic groups*, Invent. Math. **167** (2007), 295–326
- [32] G. Perelman, *The entropy formula for the Ricci flow and its geometric applications*, preprint; arXiv:math/0303109
- [33] G. Perelman, *Ricci flow with surgery on three-manifolds*, preprint; arXiv:math/0307245
- [34] M. Pollicott, *Exponential mixing for the geodesic flow on hyperbolic three-manifolds*, J. Stat. Phys. **67** (1992), nos. 3–4, 667–673
- [35] M. Sageev, *Ends of group pairs and non-positively curved cube complexes*, Proc. London Math. Soc. (3) **71** (1995), no. 3, 585–617
- [36] M. Sageev, *Codimension-1 subgroups and splittings of groups*, J. Algebra **189** (1997), 377–389
- [37] P. Scott, *Ends of pairs of groups*, J. Pure Appl. Algebra **11** (1977), 179–198
- [38] P. Scott, *Subgroups of surface groups are almost geometric*, J. London Math. Soc. (2) **17** (1978), no. 3, 555–565
- [39] J.-P. Serre, *Trees*, Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003
- [40] J. Stallings, *On torsion-free groups with infinitely many ends*, Ann. of Math. (2) **88** (1968), 312–334
- [41] W. Thurston, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. **6** (1982), no. 3, 357–381
- [42] F. Waldhausen, *On irreducible 3-manifolds which are sufficiently large*, Ann. of Math. (2) **87** (1968) 56–88
- [43] D. Wise, *Cubulating small cancellation groups*, GAFA **14** (2004), 150–214
- [44] D. Wise, *The structure of groups with a quasiconvex hierarchy*, preprint; available from the author's webpage
- [45] A. Zuk, *Property (T) and Kazhdan constants for discrete groups*, GAFA **13** (2003), no. 3, 643–670

UNIVERSITY OF CHICAGO, CHICAGO, ILL 60637 USA

*E-mail address:* dannyc@math.uchicago.edu