GEOMETRIC GROUP THEORY AND 3-MANIFOLDS HAND IN HAND: THE FULFILLMENT OF THURSTON’S VISION

MLADEN BESTVINA

Dedicated to Bill Thurston (1946–2012), who taught us how to think about mathematics

Abstract. In the late 1970s, Thurston revolutionized our understanding of 3-manifolds. He stated a far-reaching geometrization conjecture and proved it for a large class of manifolds, called Haken manifolds. He also posed 24 open problems, describing his vision of the structure of 3-manifolds.

Pieces of Thurston’s vision have been confirmed in the subsequent years. In the meantime, Dani Wise developed a sophisticated program to study cube complexes and, in particular, to promote immersions to embeddings in a finite cover. Ian Agol completed Wise’s program and, as a result, essentially all problems on Thurston’s list are now solved. In these notes I will outline a proof that closed hyperbolic 3-manifolds are virtually Haken.

1. Introduction

One way to understand surfaces is to successively cut them along incompressible circles and arcs until a collection of disks is obtained. Figure 1 shows this process for the torus.

![Figure 1. Cutting a torus.](image)

By *incompressible* we mean that circles are $\pi_1$-injective (or equivalently do not bound disks), and that arcs, whose boundaries are always in the boundary of the surface, do not cobound disks with arcs in the boundary of the surface. The collection of surfaces obtained by successive cuts is the *hierarchy* of the original surface. This process does not quite work for the 2-sphere (there are no incompressible circles!) but can be used, for example, to prove that a homotopy equivalence between two *aspherical* closed surfaces is homotopic to a homeomorphism by inducting on the hierarchy.

--

Received by the editors May 22, 2013.

2010 Mathematics Subject Classification. 57M50, 57N10.

A manifold is aspherical if its universal cover is contractible.
In the 1960s Wolfgang Haken [Hak62] introduced the analogous notion of hierarchies for aspherical 3-manifolds. The cuts are required to be along aspherical incompressible surfaces. Haken established that if a closed aspherical 3-manifold admits the first cut, then it has a (Haken) hierarchy terminating in a collection of 3-balls. Manifolds with a hierarchy are called Haken manifolds. Waldhausen [Wal68] proved that homotopy equivalent Haken manifolds are homeomorphic, and in the same paper he points out the possibility that non-Haken aspherical 3-manifolds might have a finite-sheeted covering space which is Haken. This weaker condition would suffice in many applications and has become known as Waldhausen’s virtual Haken conjecture, although I should point out that at the time all known aspherical non-Haken manifolds were of very simple type (small Seifert fibered spaces, i.e. admitting the structure of a circle fibration with three singular fibers and 2-sphere quotient).

In the late 1970s William P. Thurston completely reshaped 3-manifold theory. He stated his Geometrization Conjecture and proved it for Haken manifolds. He also showed that in some sense generic 3-manifolds are non-Haken, and in particular constructed the first examples of hyperbolic manifolds that were provably non-Haken.

In [Thu82] Thurston outlined his vision for 3-manifolds and posed 24 problems. The most famous of these, the Geometrization Conjecture, was settled by Perelman. Other major conjectures that followed were the Tameness Conjecture (Agol [Agoa], and Calegari and Gabai [CG06]), Ending Lamination Conjecture (Brock, Canary, and Minsky [BCM]), and the Surface Subgroup Conjecture (Kahn and Marković [KM]).

The virtual Haken conjecture, using Perelman’s work, reduces quickly to closed hyperbolic manifolds $M$, and in that case the Kahn–Marković theorem says that $\pi_1(M)$ contains many (quasi-convex) surface subgroups. Peter Scott observed in 1978 [Sco78] that in this situation one could prove that $M$ is virtually Haken, provided the surface subgroup $H$ is separable: for every $g \in \pi_1(M) - H$ there is a finite index subgroup $G < \pi_1(M)$ such that $H < G$ but $g \not\in G$. In one dimension lower, this is illustrated in Figure 2.

In the meantime, geometric group theorists have been considering groups and spaces in part motivated by 3-manifolds and hyperbolic geometry. Gromov [Gro87] introduced the concept of hyperbolic groups, a rich class of groups that contains fundamental groups of closed hyperbolic manifolds. Dani Wise, with coauthors, had been developing the theory of special cube complexes. The important thing is that the hyperbolic fundamental group of a compact special cube complex has the property that all of its quasi-convex subgroups are separable. On the other hand, using the Kahn–Marković theorem and an old construction of Sageev, Bergeron and Wise [BW] proved that for every closed hyperbolic 3-manifold there is a non-positively curved (NPC) compact cube complex with the same fundamental group (notice the absence of the adjective “special”).

The last step was accomplished by Ian Agol in April 2012 [Agob]: Every compact NPC cube complex with hyperbolic fundamental group is virtually special. Putting

---

2One should really say homotopy 3-balls, as the Poincaré conjecture was not known at the time.
Figure 2. The immersed curve \( C \) with \( H = \pi_1(C) \) can be lifted to an embedded curve in the double cover of the surface obtained by cutting along the two embedded curves, taking two copies, and regluing the boundaries of different copies with each other. The group element \( g \notin H \) is represented by a half of \( C \), forming a loop based at the intersection point.

it all together, the following four theorems essentially finish off Thurston’s list of problems. Let \( M \) be a hyperbolic 3-manifold, complete with finite volume.

1. \( M \) is virtually Haken.
2. \( M \) is large, i.e. it has a finite cover whose fundamental groups map onto the free group \( F_2 \).
3. \( \pi_1(M) \) is locally extended residually finite (LERF), i.e. it has the property that every finitely generated subgroup is separable.
4. \( M \) has a finite cover that fibers over the circle.

When \( M \) is not closed, but has at least one cusp, it is automatically Haken and Cooper, Long, and Reid [CLR97] proved that such \( M \) is large. It follows from the work of Wise [Wis] (which in turn uses the Tameness Conjecture, Agol’s fibering criterion and [CLR97]) that \( M \) virtually fibers and that \( \pi_1(M) \) is LERF. We will focus on the case when \( M \) is closed.

I will discuss (special) cube complexes and Sageev’s construction in detail, with the general mathematician in mind. I will then give an outline of Agol’s argument, really more of a roadmap for those who wish to take a closer look at the proofs.

Casson had early ideas about the usefulness of cube complexes for 3-manifolds. Gromov’s Link Condition (see below) put the theory on the firm footing. Aitchison and Rubinstein considered 3-manifolds that can be given the structure of an NPC cube complex [AR90]. Sageev’s thesis [Sag95], with his seminal construction of an NPC cube complex from a codimension 1 subgroup, marks the beginning of a systematic study of NPC cube complexes.

I will not define the concepts of hyperbolic groups or CAT(0) spaces. They have become ubiquitous in modern mathematics since Gromov’s groundbreaking paper [Gro87], and those readers unfamiliar with them are referred to [BH99] for a thorough introduction.

2. NPC CUBE COMPLEXES

An \( n \)-cube is an isometric copy of \([-1, 1]^n\). It has faces obtained by fixing some of the coordinates to be \( \pm 1 \). Each face is naturally a cube of appropriate dimension, e.g. a 2-cube (i.e. a square) has four faces of dimension 1 (edges) and four faces of dimension 0 (vertices). A cube complex is a space obtained from a collection of

\footnote{There is another problem left open, about volumes of hyperbolic manifolds.}
cubes by isometrically identifying some of the faces. For example, the 2-torus is a cube complex obtained from a square by identifying opposite faces. The reader is referred to Sageev’s lecture notes [Sag] for a gentle introduction to NPC cube complexes; see also Figure 3.

We equip each cube with the natural Euclidean metric, and a cube complex with the path metric induced by this metric on the cubes.

The link of a vertex in an n-cube is the \((n - 1)\)-simplex of tangent vectors that point into the cube. The link of a vertex in a cube complex is naturally the union of simplices. The cube complex \(X\) is NPC (nonpositively curved) if:

- the link of every vertex in \(X\) is a simplicial complex, which is also a flag complex.

So for example, the double of a square along its boundary is not NPC, since the link of a vertex is the double of an edge along its boundary, and this is not a simplicial complex. A simplicial complex \(L\) is a flag complex if it is determined by its 1-skeleton, i.e. if \(v_1, \ldots, v_k\) are distinct vertices such that every pair bounds an edge, then \(L\) contains the simplex with vertices \(v_1, \ldots, v_k\). For example, the boundary of the 3-cube is not NPC since the links of vertices are hollow triangles. This definition is based on Gromov’s Link Condition [Gro87,BH99] which says that a cube complex is NPC if it is locally CAT(0). In particular, by (a version of) the Cartan–Hadamard theorem [Gro87,BH99,Lea], the universal cover of an NPC cube complex is \(CAT(0)\).

2.1. Hyperplanes. Two parallel edges of a square in a cube complex are square equivalent, and we extend this to an equivalence relation on the set of all edges. In any cube \([-1, 1]^n\), a midcube is the subset obtained by fixing one of the coordinates to be 0. Thus an \(n\)-cube has \(n\) midcubes, and each midcube intersects \(2^{n-1}\) edges.
all of which are square equivalent. The hyperplane dual to an equivalence class of oriented edges is the union of all midcubes that intersect only the edges contained in the equivalence class. A hyperplane is embedded if it intersects each cube in at most one midcube. Each midcube has a natural normal bundle, which is a trivial line bundle. These bundles glue together to form a normal bundle of a hyperplane. The hyperplane is 2-sided (1-sided) if the normal bundle is trivial (nontrivial, respectively). See Figures 4 and 5.

If $X$ is CAT(0), then hyperplanes are embedded and they have the structure of CAT(0) cube complexes.

3. Sageev’s construction

The great thing about NPC cube complexes is that they keep track of intersections between hypersurfaces. First recall that if $S \subset M$ is an incompressible surface in a 3-manifold (or a circle in a surface), then $\pi_1(M)$ acts on a tree $T$ encoding the pattern of the components of the preimage $\tilde{S}$ of $S$ in the universal cover $\tilde{M}$ of $M$ (these components are copies of the universal cover of $S$). The vertices of $T$ are represented by the components of $\tilde{M} - \tilde{S}$ and the edges by the components of $\tilde{S}$. The stabilizer of an edge of $T$ is $\pi_1(S)$, and the tree $T$ encodes the pattern seen in $\tilde{M}$. This construction is part of the Bass–Serre theory (see e.g. [SW79]). See Figure 6.

Now suppose that $S$ is only immersed in $M$, but the preimage $\tilde{S} \subset \tilde{M}$ consists of (embedded, but possibly intersecting) copies of the universal cover of $S$. 

![Figure 5. Self-intersecting hyperplane.](image1)

![Figure 6. Construction of the Bass–Serre tree.](image2)
Michah Sageev [Sag95] discovered that this time there is a CAT(0) cube complex $\tilde{X}$ encoding the pattern in $\tilde{M}$.

Here is a more abstract version of Sageev’s construction, due to Haglund and Paulin [HP98].

Let $Y$ be a set. A wall is a partition of $Y$ into two subsets. A wall set is a pair $(Y, W)$ where $W$ is a collection of walls. We require that any two points of $Y$ are separated by finitely many, and at least one, wall.

A half-space determined by a wall is one of the two subsets in the partition. An ultrafilter $\omega$ is a collection of half-spaces so that:

- if $W = \{A, A^c\}$, then exactly one of $A, A^c$ is in $\omega$;
- if $W = \{A, A^c\}$, $W' = \{B, B^c\}$, $A \subset B$, and $A \in \omega$, then $B \in \omega$.

We think of $\omega$ as making a choice of a half-space determined by each wall. For every $y \in Y$ we have the ultrafilter $\omega_y$ that chooses the half-space containing $y$. If $\omega, \omega'$ are two ultrafilters, define the distance $d(\omega, \omega')$ as the number of walls where $\omega, \omega'$ made different choices (this is possibly infinite).

For any $y, y' \in Y$ the ultrafilters $\omega_y, \omega_{y'}$ are a finite distance apart. Let $V$ be the set of all ultrafilters a finite distance apart from each $\omega_y$. Then $V$ is the 0-skeleton of the Sageev cube complex $\tilde{X}$ associated with the wall space. Two vertices $v, v'$ are connected by an edge if they differ in exactly one wall. An $n$-cube is determined by $n$ walls $W_1, \ldots, W_n$ and its vertices all agree on every other wall, but all $2^n$ choices on $W_i$’s represent vertices of the cube.

**Example 3.1.** Let $Y$ have seven points with three walls; see Figure [7]. There are eight ultrafilters: seven of the form $\omega_y$ plus an “ideal” ultrafilter that always picks the half-space that does not contain the middle point. The Sageev cube complex for this example is the 3-cube.

![Figure 7. The seven points example.](image)

Two walls $W = \{A, A^c\}$, $W' = \{B, B^c\}$ intersect if all four intersections of half-spaces are nonempty: $A \cap B \neq \emptyset, A \cap B^c \neq \emptyset, A^c \cap B \neq \emptyset, A^c \cap B^c \neq \emptyset$.

**Proposition 3.2.**

- $\tilde{X}$ is a CAT(0) cube complex.
- There is a natural injection $Y \to V$ and bijection $W \to \text{hyperplanes and the latter is stabilizer-preserving}$.
- If $Y$ is the set of the vertices of a CAT(0) cube complex $\tilde{X}$ and walls are hyperplanes, the Sageev cube complex is $\tilde{X}$.
- If there is an $n$-cube corresponding to the walls $W_1, \ldots, W_n$, then $W_1, \ldots, W_n$ pairwise intersect.
We will apply Sageev’s construction in the context of a closed hyperbolic 3-manifold (or a surface) \( M \) and a finite collection of incompressible surfaces (or circles) \( S_i \) immersed in \( M \). We assume that in the universal cover \( \tilde{M} = \mathbb{H}^3 \) the preimage of the \( S_i \)'s consists of embedded copies of the universal covers of the \( S_i \)'s and that they intersect transversally. Let \( Y' \) be the set of complementary components of the preimage of \( \bigcup S_i \) in \( \tilde{M} \), and for the walls take the collection of copies of universal covers of \( S_i \). There may be distinct points in \( Y' \) not separated by any walls, so for \( Y \) take the set of equivalence classes of points in \( Y' \), with two points equivalent if they are not separated by any walls.

**Proposition 3.3** ([Sag97]). In this situation, if each \( \pi_1(S_i) \) is quasi-convex in \( \pi_1(M) \), the action of \( \pi_1(M) \) on \( \tilde{X} \) is cocompact.

The quasi-convexity condition means the following: if \( W \) is a copy of the universal cover of \( S_i \) in \( \tilde{M} \), then there is \( R > 0 \) so that any geodesic in \( \tilde{M} = \mathbb{H}^3 \) with endpoints in \( W \) stays in the \( R \)-neighborhood of \( W \). We recall a theorem from [GMRS98]: for any \( R \) there is \( k \) so that given \( k \) walls at least two are distance > \( R \) apart (they prove this more generally in the context of hyperbolic groups and quasi-convex subgroups). In particular, the collection of walls satisfies Scott’s \( k \)-plane property: for any \( k \) distinct walls there are two that are disjoint. Proposition 3.3 now follows from the last bullet of Proposition 3.2.

### 4. Special cube complexes

Here is the key definition of Haglund and Wise [HW08].

**Definition 4.1.** An NPC cube complex is **special** if

- hyperplanes are embedded and 2-sided, and
- there are no direct self-osculations nor inter-osculations.

Two oriented edges with the same initial vertex are **perpendicular** if they are connected in the link, i.e. span a square, otherwise they **osculate**. Orienting the normal bundle of a 2-sided hyperplane \( H \) assigns an orientation to every edge dual to \( H \). A 2-sided embedded hyperplane **directly self-osculates** if it has dual osculating (oriented) edges. It **indirectly self-osculates** if there are osculating dual edges intersecting a hyperplane with opposite orientations. Two hyperplanes **inter-osculate** if they have perpendicular dual edges that are square equivalent to osculating dual edges. Hyperplanes can be made 2-sided by passing to double covers. Likewise, indirect self-osculations can be removed by subdivision or passing to a double cover. See Figure 8.

Achieving embedded hyperplanes and removing the other two kinds of osculations in a finite cover requires separability properties of fundamental groups of hyperplanes and their double cosets. For hyperbolic groups we have the following criterion: \textit{a compact NPC cube complex with hyperbolic fundamental group is virtually special if all quasi-convex subgroups are separable.} See [HW08]; this uses [Min05] (for one direction, see Proposition 4.3 below). In particular, virtual specialness of a compact NPC cube complex with hyperbolic \( \pi_1 \) depends only on \( \pi_1 \).

Notice the analogy between special cube complexes (walls are embedded) and Haken 3-manifolds.

---

5More generally, a subgroup \( H \) of a hyperbolic group \( G \) is quasi-convex if there is \( R \) so that any geodesic in the Cayley graph of \( G \) with endpoints in \( H \) stays in the \( R \)-neighborhood of \( H \).
4.1. RAAGs. The model special cube complexes come from Right Angled Artin Groups (RAAGs). Let $\Gamma$ be a simplicial graph. Construct a cube complex $X_\Gamma$ as follows. It has only one vertex $v$. There is one edge $e_a$ for every vertex $a$ of $\Gamma$, and there is one 2-cell $Q_{a,b}$ for every edge $[a,b]$ of $\Gamma$: it is obtained by attaching a square via the commutator relation $e_a e_b e_a^{-1} e_b^{-1}$. Thus $Q_{a,b}$ is a 2-torus. We then fill in higher dimensional cells: if the vertices $a,b,c \in \Gamma$ are pairwise joined by edges, then $Q_{a,b} \cup Q_{a,c} \cup Q_{b,c}$ is the 2-skeleton of a 3-torus, and we glue in a 3-cell $Q_{a,b,c}$ by identifying opposite faces of a 3-cube and gluing them to the 2-cells. Proceed similarly in higher dimensions.

For example, if $\Gamma$ is a finite set, $X_\Gamma$ is a wedge of circles. If $\Gamma$ is a complete graph on $n$ vertices, $X_\Gamma$ is the $n$-torus. The fundamental group of $X_\Gamma$ is the RAAG associated to $\Gamma$, given by the presentation
\[
\langle g_v, v \in \Gamma^{(0)} \mid g_v g_w = g_w g_v, [v, w] \subset \Gamma \rangle;
\]
that is, the generators are in one-to-one correspondence with the vertices of $\Gamma$, and two generators commute when the corresponding vertices of $\Gamma$ are joined by an edge. The complex $X_\Gamma$ is a special NPC cube complex.

Now RAAGs are universal special cube complexes, in the sense that any special cube complex $X$ can be locally isometrically immersed in some $X_\Gamma$. The construction of $\Gamma$ is simple: the vertices are the hyperplanes of $X$, and edges represent intersecting hyperplanes. The definition of “special” is designed so that the obvious map that sends each vertex of $X$ to the only vertex of $X_\Gamma$ and each edge of $X$ to the edge of $X_\Gamma$ representing the dual hyperplane (and being careful with orientations) extends to a local isometry $X \to X_\Gamma$.

When $\Gamma$ is finite, the RAAG $\pi_1(X_\Gamma)$ is linear, in fact it can be realized as a subgroup of $SL_N(\mathbb{Z})$ for large $N$. This follows from [DJ00].

**Theorem 4.2** ([HW08]). *If $X,Y$ are finite special cube complexes and $f : X \to Y$ is a locally isometric immersion, there is a finite cover $\hat{Y} \to Y$ to which $f$ lifts and the image of $X$ in $\hat{Y}$ is a retract of $\hat{Y}$.*

The proof is elementary and is modeled on the case of graphs (they are special cube complexes!), which is how Stallings [Sta83] proved the Marshall Hall theorem for free groups. See also [Hag08].

**Sketch of the proof.** First consider the special case when $Y = X_\Gamma$ and $X \to Y$ is constructed above. To construct $\hat{Y}$ look at the components of the preimages of edges. For simplicity, assume that $X$ has no indirect self-osculations as well. Then
these components are circles and arcs, each consisting of a single edge (in general, they will consist of several coherently oriented edges). Leave the circle components alone. To each arc component attach an arc making it a circle and map it to the edge (circle) in \( Y \) by double cover. The retraction takes this new arc to the old arc relative to the endpoints. Now proceed with 2-cells. Each component of the preimage contains one 2-cell, which is a square, an annulus, or a torus. Turn it into a torus by adding 2-cells (the 1-skeleton is already in place), so the resulting torus will have four, two, or one 2-cells and will map to the target 2-cell (also a torus) by a covering map of degree 4, 2, or 1. Proceed analogously in higher dimensions.

Now in general, first immerse \( Y \) in a RAAG complex \( Z \) as above. Then construct \( \hat{Z} \) as in the special case for \( X \to Z \), and then let \( \hat{Y} \) be the pullback of \( \hat{Z} \) to \( Y \).

### 4.2. Consequences of being special.

**Proposition 4.3.** Let \( X \) be a compact virtually special cube complex with \( G = \pi_1(X) \) hyperbolic. Then \( G \) satisfies the following:

1. quasi-convex subgroups are virtual retracts, hence separable;
2. \( G \) is linear, in fact it embeds in \( GL_n(\mathbb{Z}) \);
3. \( G \) is large or virtually abelian;
4. \( G \) is residually finite rationally solvable (RFRS).

**Sketch of proof.** For (1), the key is that a quasi-convex subgroup \( H \subset G \) can be represented by a locally isometric immersion \( Z \to X \) of finite special cube complexes. This follows from Haglund [Hag08] (the proof is a pleasant exercise in hyperbolic geometry: show that the intersection of half-spaces containing a given orbit is contained in a Hausdorff neighborhood of the orbit). Then (1) follows from Theorem 4.2 (it is not hard to see that virtual retracts in residually finite groups are separable).

For (2) use the fact that RAAGs are linear and that local isometric embeddings are \( \pi_1 \)-injective.

For (3), find a quasi-convex free subgroup, using the Tits alternative (plus the fact that solvable subgroups of CAT(0) groups are virtually abelian) and a ping-pong argument, and apply (1).

RFRS is a group-theoretic property satisfied by subgroups of RAAGs and if it holds for a closed hyperbolic 3-manifold \( M \) (or even complete, finite volume), then \( M \) virtually fibers over the circle. This is Agol’s criterion for virtual fibering [Ago08].

### 5. Dehn fillings

Here we discuss the last general prerequisite. The motivation for this theorem is the celebrated theorem of Thurston: *If \( M \) is a hyperbolic manifold with a cusp, then all but finitely many Dehn fillings produce closed hyperbolic manifolds.* Recall that \( M \) is the interior of a compact manifold whose boundary is a torus (assuming orientability). By a Dehn filling along a nontrivial circle \( a \) in the boundary torus, we mean attaching a 2-handle (i.e. a thickened disk) along \( a \) and then attaching a 3-ball to the boundary 2-sphere of the resulting 3-manifold. The curve \( a \) is defined only up to isotopy, and the set of possible choices is naturally \( \mathbb{Q} \cup \{ \infty \} \), i.e. the slope of \( a \), once the coordinates on the torus are chosen.
Another, more simple-minded and 2-dimensional motivation for the theorem below, is the classical small cancellation theory (for a nice exposition, see e.g. [Str90]).

Now let $G$ be a hyperbolic group, and let $H_1, \ldots, H_k$ be a collection of quasi-convex subgroups. We will assume that this collection is almost malnormal, i.e. $gH_ig^{-1} \cap H_j$ infinite implies $i = j$ and $g \in H_i$.

By a Dehn filling of $G$ we will mean passing to the quotient $\overline{G} = G/\langle\langle N_1 \cup N_2 \cup \cdots \cup N_k \rangle\rangle$, where we mod out the normal closure of the union of specified normal subgroups $N_i \leq H_i$. For simplicity we will always have that $N_i$ has finite index in $H_i$.

The following is the main Dehn Filling Theorem. See [Osi07, DGO], and also [GM08] for the torsion-free case. We are stating only a special case relevant to us. For the quasi-convexity statement, see [AGM09] and [MMP10].

**Theorem 5.1.** Let $G$ be a hyperbolic group, and let $H_1, \ldots, H_k$ be a finite almost malnormal collection of quasi-convex subgroups, all contained in a quasi-convex subgroup $H < G$. Let $F$ be a finite subset of $G$. Then for any sufficiently long filling (*) the quotient $\phi : G \to \overline{G}$ satisfies the following:

- $\text{Ker}(\phi|H_i) = N_i$;
- $\overline{G}$ is hyperbolic;
- $\phi|F$ is injective;
- $\phi(F) \cap \phi(H_i) = \phi(F \cap H_i)$;
- the image of $H$ in $\overline{G}$ is quasi-convex.

By a “sufficiently long filling” we mean that there exists a finite bad set $B \subset G - \{1\}$ so that the conclusions hold whenever $N_i \cap B = \emptyset$ for every $i$. Recall that $N_i$ has finite index in $H_i$.

6. Hierarchies

We will say that a hyperbolic group $G$ is virtually special if it acts with finite point stabilizers on a CAT(0) cube complex $\tilde{X}$ so that for a torsion-free finite index subgroup $H < G$ the quotient $\tilde{X}/H$ is a compact special cube complex.

The following two theorems are rather deep. The first combines theorems of Hsu and Wise and of Haglund and Wise. The reader may want to look at the informal discussion of these theorems in [Wis], e.g. the case when $A$ and $B$ are free groups and $C$ is cyclic.

**Theorem 6.1** (The Hsu–Wise Combination Theorem [HsW] and the Haglund—Wise Combination Theorem [HWP12]). Suppose $G$ is hyperbolic, $G = A *_C B$ or $G = A *_C C$, $C$ is quasi-convex and almost malnormal in $G$, and $A, B$ are virtually special. Then $G$ is virtually special.

Following Wise [Wis] we define the class of groups $\mathcal{AMQH}$. These are hyperbolic groups that admit an almost malnormal quasi-convex hierarchy:

- $1 \in \mathcal{AMQH}$;
- if $A, B \in \mathcal{AMQH}$ and $C$ is almost malnormal and quasi-convex in the hyperbolic group $G = A *_C B [G = A *_C C]$, then $G \in \mathcal{AMQH}$.
Thus $G \in \mathcal{AMQH}$ implies $G$ is hyperbolic virtually special and $G$ hyperbolic virtually special implies $G$ is virtually in $\mathcal{AMQH}$ (by taking a cover where wall groups are malnormal).

Thus $\mathcal{AMQH}$ is something like a group-theoretic version of Haken hyperbolic 3-manifolds, but it contains many non-3-manifold groups. This is crucial for the proof of Theorem 6.3 below, since in the course of the argument one passes to quotient groups that are not necessarily 3-manifold groups, even if the original group $G$ is a 3-manifold group. But one stays in the class of virtually special groups, thanks to the following amazing theorem of Wise, which can be regarded as an addendum to the Dehn Filling Theorem.

**Theorem 6.2** (Wise’s Malnormal Special Quotient Theorem [Wis]). Under the assumptions of the Dehn Filling Theorem 5.1 assume that $G$ is also virtually special. Then all sufficiently deep fillings have a virtually special quotient.

Again, the reader is invited to follow the discussion in [Wis] and to think about the special case when $G$ is a free group and $H_i$ are cyclic. By “sufficiently deep” I mean that there are finite index subgroups $H_i' < H_i$ so that if $N_i < H_i'$, then the conclusion follows.

We remark here that it is a major open problem whether hyperbolic groups are residually finite. To apply the Dehn Filling Theorem, one needs to know that the groups $H_i$ are residually finite or else there is no way to choose the normal subgroups $N_i$ to avoid the bad set $B$. The Malnormal Special Quotient Theorem makes inductive arguments possible as it keeps all groups residually finite. The following key ingredient is proved using such an induction.

**Theorem 6.3** (The Agol–Groves–Manning Weak Separation Theorem [Agob]). Let $G$ be a hyperbolic group, let $H$ be a quasi-convex and virtually special subgroup of $G$, and let $g \in G - H$. Then there is a hyperbolic quotient $\pi : G \to \overline{G}$ so that $\pi(H)$ is finite and $\pi(g) \notin \pi(H)$.

Idea of proof. If $H$ is almost malnormal, the statement follows from the Dehn Filling Theorem. Now imagine that $H$ is a surface subgroup in a hyperbolic 3-manifold group $G$. Then $H$ may not be almost malnormal since for example there could be elements $\gamma \in G - H$ such that $H^\gamma \cap H$ is infinite cyclic, corresponding to copies of the universal covers of the surface intersecting in lines (but let’s say these are the only obstructions to malnormality). If each such infinite cyclic subgroup is replaced by the maximal infinite cyclic subgroup containing it and then conjugates are removed, the resulting collection is finite and almost malnormal. Now apply the Dehn Filling Theorem and the Malnormal Special Quotient Theorem to pass to a hyperbolic quotient $G'$ of $G$ so that the image $g'$ of $g$ is outside the image $H'$ of $H$, and so that $H'$ is virtually special. We are now in a similar situation as before, but $H'$ is malnormal so we can finish with another application of the Dehn Filling Theorem.

In general, the proof runs by induction on the height of $H$ in $G$. This is the largest $n$ such that $H^{g_1} \cap \cdots \cap H^{g_n}$ is infinite and the cosets $g_i H$ are all distinct. The Gitik–Mitra–Rips–Sageev theorem [GMRS98] says that the height of a quasi-convex subgroup of a hyperbolic group is always finite. The first step is to apply the Dehn Filling Theorem and the Malnormal Special Quotient Theorem to the collection of subgroups obtained from minimal infinite intersections of conjugates of $H$ by first replacing them by maximal finite index overgroups and then discarding...
conjugates. The quotient groups satisfy all the hypotheses, but the image of $H$ has smaller height than before.

7. Hyperbolic 3-manifolds and NPC cube complexes

The key in this program is that hyperbolic 3-manifold groups act properly on cubings. This fact was established soon after the seminal construction of Kahn and Marković [KM] of nearly totally geodesic immersed surfaces in closed hyperbolic 3-manifolds.

**Theorem 7.1** ([BW]). *Let $M$ be a closed hyperbolic 3-manifold. Then $G = \pi_1(M)$ is the fundamental group of a compact NPC cube complex.*

**Sketch of proof.** Fix a compact fundamental domain $F$ for the action of $G$ on $\mathbb{H}^3$. There is $\epsilon > 0$ so that every geodesic in $\mathbb{H}^3$ that intersects $F$ has endpoints at distance $> \epsilon$ in the boundary $S^2_\infty$ of $\mathbb{H}^3$. Fix a finite collection of round circles in $S^2_\infty$ so that every pair of points at distance $\geq \epsilon$ is separated by at least one circle.

Kahn and Marković [KM] show that for every round circle in $S^2_\infty$, there is a quasi-convex surface subgroup $H$ whose limit set is a fractal circle contained in an arbitrarily small annular neighborhood of it and going around the annulus once. This subgroup is represented by an immersed surface in $M$ which is nearly totally geodesic and its universal cover embeds in $\mathbb{H}^3$. Now for each circle in our collection, choose a Kahn–Marković surface subgroup $H_i$ whose limit set approximates it. Let $\tilde{X}$ be the Sageev complex associated to this finite collection of quasi-convex subgroups, and let $X = \tilde{X}/G$. Then $X$ is compact (see Proposition 3.3). To see that the action is free, observe that for every $g \in G - \{1\}$ some conjugate of $g$ will have its axis in $\mathbb{H}^3$ that intersects $F$, so its endpoints will be separated by one of the round circles, and hence will lie on opposite sides of the plane covering one of the Kahn–Marković surfaces. A high power of $g$ will thus take a half-space properly into itself, and this implies that $g$ acts hyperbolically on $\tilde{X}$. See [HrW].

8. Wise’s conjecture and Agol’s theorem

Let $M$ be a closed hyperbolic 3-manifold. As a consequence of Theorem 7.1, we have $\pi_1(M) = \pi_1(X)$ for a compact NPC cube complex $X$. If we could show that $X$ is virtually special, we would be done by a classical argument. For example, we could use subgroup separability and find a finite cover of $M$ to which a Kahn–Marković surface lifts to an embedding. Alternatively, a finite cover of $X$ has a splitting along a hyperplane, and one could use a construction of Stallings (“pull back and compress”) to find an incompressible surface in the corresponding finite cover of $M$. Thus the following theorem finishes the program.

**Theorem 8.1** (Agol’s theorem (2012) [Agob] and Wise’s conjecture (2011) [Wis]). *Every finite NPC cube complex with hyperbolic $\pi_1$ is virtually special.*

Note that this fails without hyperbolicity. An extreme example is the group constructed by Burger and Mozes [BM97]: It is a simple group acting cocompactly on the product of two trees. Since in particular it is not residually finite, it cannot be virtually special. There is an earlier example of Wise [Wis07] of such a group which also does not have nontrivial subgroups of finite index (but is not simple).

In the remainder of the paper I will outline a proof of Agol’s theorem.
8.1. **Coloring Lemma.** An *n-coloring* of a graph is a function from the vertex set to \([n] := \{1, \ldots, n\}\) that assigns different numbers (colors) to the endpoints of each edge. Note that any countable graph with valencies bounded by \(k\) has a \((k+1)\)-coloring.

**Proposition 8.2.** Let \(G\) act on a countable simplicial graph \(\Gamma\) with valencies bounded by \(k\). Then there is a \(G\)-invariant probability measure on the space of \((k+1)\)-colorings of \(\Gamma\).

**Proof.** The space \([n]^\Gamma\) of all functions \(\Gamma^0 \to \{1, \ldots, n\}\) from the vertex set of \(\Gamma\) is a Cantor set and the space of \(n\)-colorings is a closed subset, and it is nonempty when \(n > k\). Let \(\mu_n\) be the uniform (product) measure on \([n]^\Gamma\) and note that it assigns \(1/n\) to the subset where a particular pair of neighbors are assigned the same color. For \(n > k + 1\), define \([n]^\Gamma \to [n-1]^\Gamma\) by replacing color \(n\) at any vertex \(v\) with the minimal color not represented by any neighbors of \(v\). Compose these to get to \([k+1]^\Gamma\). The feature of this construction is that if an edge has vertices of different colors at the start, it also does at the end. These maps are continuous (since they are locally defined) so we may push the uniform measure to get a \(G\)-invariant measure \(\nu_n\) on \([k+1]^\Gamma\) that still assigns \(\leq 1/n\) to the set where a particular edge gets one color. Take a subsequence that converges to a limiting measure \(\nu\). This is still \(G\)-invariant and assigns 0 to the set that assigns one color to a particular edge. So the support of \(\nu\) is contained in the infinite intersection, one for every edge, where the endpoints get different colors, i.e. it is a measure on the set of colorings of \(\Gamma\). \(\square\)

In the application below \(G\) will act properly (and cocompactly) on \(\Gamma\). If \(G\) is also residually finite, there is a finite index subgroup \(G' < G\) that acts freely with \(\Gamma/G'\) a simplicial graph, so there is a \(G'\)-equivariant coloring of \(\Gamma\) (and hence a \(G\)-invariant measure with support in a finite set of colorings). In fact, with a bit more work, one can arrange that \(G\) is hyperbolic, as follows. A stronger version of the Weak Separation Theorem 6.3 (which can be proved in the same way) states that \(H\) can be separated from any finite collection of elements outside \(H\) (instead of just one element). Then the quotient \(G \to G'\) can be constructed with a single application of this theorem, by taking for \(H\) the free product of suitable conjugates of hyperplane groups and conjugating in the same way the elements that need to be separated from these hyperplane groups. The Coloring Lemma is a way of getting around the issue of residual finiteness of hyperbolic groups.

8.2. **Setting the stage.** Now assume \(X\) is a compact NPC cube complex with \(G = \pi_1(X)\) hyperbolic. The goal is to show that \(X\) is virtually special. The proof is by induction on dimension of \(X\), so we may assume that the wall groups are virtually special. The first step is to construct a (most likely infinite) regular cover \(\tilde{X} \to X\) with deck group \(G\) so that \(\tilde{X}\) is special and its walls are compact. This is accomplished by using the Weak Separation Theorem 6.3. For the subgroup \(H\), one can take the free product of suitable conjugates of wall groups (for one wall in each orbit). There are finitely many elements of \(G\) that need to be separated from \(H\); apply the Weak Separation Theorem for each such element and then intersect the kernels. The cover \(\tilde{X}\) is the one that corresponds to this intersection. The

---

6I should really call them hyperplanes, but at this point there should be no confusion so I will use the shorter word.
As a schematic picture of $\hat{X} \to X$ and an immersed wall in $X$ whose preimage consists of embedded, but intersecting, compact walls.

Figure 9.

A feature of the walls $W$ in $\hat{X}$ is that there are no essential annuli $(S^1 \times [0, 1], S^1 \times \{0, 1\}) \to (\hat{X}, W)$ (i.e. any such map that sends $S^1$ to an infinite order element in $\pi_1(\hat{X})$ is homotopic into $(W, W)$). See Figure 9.

8.3. Decomposing into puzzle pieces. Cutting $\hat{X}$ along walls results in a collection of compact pieces; we will call them “puzzle pieces”. They are simply stars of vertices with respect to the barycentric subdivision of $\hat{X}$ in which an $n$-cube $[-1, 1]^n$ gets subdivided into $2^n$-subcubes of edge-length 1. A face of a puzzle piece is its intersection with a wall (if nonempty). The puzzle pieces will serve as building blocks for the construction of a finite cover of $X$, but we will first color their faces in order to encode gluing information. Let $\Gamma$ be the graph whose vertices are walls in $\hat{X}$ and whose edges correspond to pairs $(W, W')$ of walls such that either $W \cap W' \neq \emptyset$ or there is an essential annulus in $\hat{X}$ with one boundary component in $W$ and the other in $W'$. By [GMRS98] (see the discussion following Proposition 3.3) in the latter case necessarily the distance between $W$ and $W'$ is bounded, so $\Gamma$ has bounded valence, say by $k$, and the deck group $G$ acts on $\Gamma$ cocompactly. We may apply the Coloring Lemma to find a $G$-invariant probability measure $\nu$ on the space of all $(k + 1)$-colorings of $\Gamma$.

Let $c$ be a coloring. Thus each wall is assigned a color in $\{1, 2, \ldots, k + 1\}$. We can cut $\hat{X}$ along walls, first along walls numbered 1, then along walls numbered 2, etc. This gives a particular hierarchy for $\hat{X}$. If $W$ is a wall, the descending set of walls for $(W, c)$ is the following finite collection of walls: it contains $W$, and all walls $W'$ with $(W, W') \in \Gamma$ and $c(W') < c(W)$, and inductively, whenever $V$ is a wall in the collection, add all walls $V'$ with $(V, V') \in \Gamma$ and $c(V') < c(V)$. See Figure 10.

If $c, c'$ are two colorings and $W$ is a wall, we will write $(W, c) \sim (W, c')$ if the descending sets of walls for $(W, c)$ and $(W, c')$ are identical and they receive the same color from both $c$ and $c'$. Likewise, if $P$ is a puzzle piece, we say $(P, c) \sim (P, c')$.
Figure 10. A schematic picture of a possible descending set of the wall $W$ colored 4. For the example, we imagine that there are no essential annuli anywhere.

Figure 11. An example of a coloring of the walls and two of the resulting colored puzzle pieces. The equivalence class $[P, c]$ remembers a bit more than just the colors of the faces—this depends on the ordering of the colors and is suppressed in the figure.

if $(W, c) \sim (W, c')$ for every wall $W$ that intersects $P$. We will write equivalence classes as $[W, c]$ and $[P, c]$.

By a colored puzzle piece we will mean an equivalence class $[P, c]$. Informally, a colored puzzle piece is a cubical polyhedron whose faces are colored with the additional information of the coloring of the descending set of walls for each face of $P$. There are finitely many $G$-orbits of puzzle pieces. See Figure 11.

8.4. Gluing equations. We want to take a finite collection of colored puzzle pieces with appropriate multiplicities, say $\omega$, and glue them together to form a finite cover of $X$. In order to do the gluing, we need the pieces to match up; i.e. $\omega$ must satisfy the gluing equations.

Let $F$ be a common face of two puzzle pieces $P, Q$, and let $c$ be a coloring. Then $F$ is contained in a unique wall $W$ and for simplicity we will write $[F, c]$ and $c(F)$ for $[W, c]$ and $c(W)$. We have a gluing equation

$$\sum_{[P, d] \mid (F, d) \sim (F, c)} \omega([P, d]) = \sum_{[Q, d] \mid (F, d) \sim (F, c)} \omega([Q, d]).$$

Notice that there are finitely many $G$-orbits of equations, so if we impose the requirement that the multiplicities $\omega$ are $G$-invariant, there are only finitely many equations and finitely many unknowns. Also note that $\omega = \nu$ (the measure from the Coloring Lemma) is a nonnegative real solution to all of them (for a fixed $F$ the set of colorings $d$ with $(F, d) \sim (F, c)$ can be written as the disjoint union in
two different ways and the equation is a consequence of the additivity of measures). Thus by linear algebra there is a (nonzero) nonnegative \( G \)-invariant integral solution to all the equations.

8.5. **Virtual regluing.** Now we have a finite collection of colored puzzle pieces with multiplicities satisfying the gluing equations. We will glue them back so they form a finite cover of \( X \). But during the process we may have to pass to finite covers, so the number of colored puzzle pieces at the end of the process will be some multiple of the original.

In the first step, glue the colored puzzle pieces along faces colored \( k + 1 \). That is, to glue \([P, c]\) and \([Q, d]\) along a common (up to the action of \( G \)) face \( F \), we require \((F, c) \sim (F, d)\), \( c(F) = d(F) = k + 1 \) (and that \( F \) receives opposite transverse orientations from \( P \) and \( Q \)). That we can match up all faces of our puzzle pieces colored \( k + 1 \) follows from the fact that multiplicities satisfy the gluing equations. Notice that the gluing also respects the colors of the faces adjacent to \( F \) resulting in bigger colored puzzle pieces, whose faces now carry colors \( 1, 2, \ldots, k \).

The idea is now to continue this process of gluing. There is another issue that comes up; see Figure 12. The faces colored with \( k \) to be glued together are finite covers of the same complex \( Y_k \), but may not be “the same”.

This is resolved by passing to a finite cover which is regular with respect to \( Y_k \). That such a cover can be extended to a cover of the puzzle pieces is Agol’s *Gluing Theorem* and it follows, for example, from Theorem 4.2 after noting that these pieces are equipped with a hierarchy by construction, so they are virtually special. The extra information about cylinders encoded in \( \Gamma \) and about descending set of walls encoded in \([P, c]\) is used to ensure that this hierarchy is almost malnormal.

Proceeding inductively, glue in this way all the colored puzzle pieces, along faces colored from highest to lowest color. The resulting complex is a finite cover of \( X \) and it comes with a hierarchy, so it is virtually special. Note here that the hierarchy is in fact almost malnormal, since by construction there are no essential annuli with both boundary components mapped to the same color. This finishes the proof of Agol’s theorem.

![Figure 12](https://example.com/figure12.png)

**Figure 12.** An example with \( k = 3 \). Imagine having three copies of the puzzle piece on the left and four copies of the piece on the right (both are annuli). The “faces”, colored 2, cover the circle \( Y_2 \) with total degree 0, but no gluing is possible.
Acknowledgements

I would like to thank Michah Sageev for teaching me early on about cube complexes, Dani Wise for explaining some of his ideas before they became theorems, Ian Agol for beautiful lectures on which these notes are partially based, and Jason Manning for patiently explaining even the most trivial points to me and for carefully reading a draft of these notes.

About the author

Mladen Bestvina wrote his thesis in geometric topology in 1984 at the University of Tennessee under the direction of John Walsh. He held a position at UCLA before moving to the University of Utah in 1994.

References


License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use