THE STRUCTURE OF GROUPS WITH A QUASI-CONVEX HIERARCHY

Abstract. Three lectures on this topic were given by Dani Wise in Montreal on April 19 to 21. During the three lectures Dani Wise gave a sketch of the proof that Haken hyperbolic 3-manifolds are virtually hyperbolic. These notes were taken by Stefan Friedl; all errors, inconsistencies, omissions etc. are due to S.F.

1. Summary of results

Let Γ be a graph. The *graph group* $G(\Gamma)$ associated to Γ is the group

$$G(\Gamma) := \langle x_v : v \in \mathrm{Vertex}(\Gamma) \,|\, [x_u, x_v] = 1 : (u,v) \in \mathrm{Edge}(\Gamma)\rangle.$$

For example, if Γ has no edges, then $G(\Gamma)$ is a free group, if Γ is a complete graph, then $G(\Gamma)$ is a free abelian group. Note that graph groups are often referred to as right angled Artin groups (RAAG).

We say a group G has a *quasi-convex hierarchy* if it can built from trivial groups by a sequence of HNN extensions $A*_{C^t=C'}$ and amalgamated free products $A *_C B$ such that the groups C are finitely generated and embed quasi-isometrically into $A*_{C^t=C'}$ respectively $A *_C B$. Note that every finitely generated graph group has a hierarchy.

Remark. If C embeds quasi-isometrically into A and B then it does not necessarily embed quasi-isometrically into $A *_C B$.

Theorem 1.1. *If G is word hyperbolic and if G has a quasi convex hierarchy, then G is virtually compact special.*

Note that the theorem also holds if G is word hyperbolic relative to abelian subgroups, but the details have not been worked out yet (see Section 4 for details).

Corollary 1.2. *If G is word hyperbolic and if G has a quasi convex hierarchy, then G has a finite index subgroup that embeds in a graph group.*

There are two main applications of this theorem. First, let G be a one-relator group with torsion, i.e. any group with a presentation $\langle a_1, a_2, \dots \,|\, W^n\rangle$ for some $n > 1$. By Newman's spelling theorem the group G is word-hyperbolic. It also has a hierarchy (Magnus hierarchy) which is quasi-convex but which ends in finite groups. (Note that groups with a hierarchy are torsion-free.) But G admits a finite index subgroup which has a quasi-convex hierarchy in the above sense. We can thus apply the theorem and its corollary to conclude that G is residually finite, answering a question posed by Baumslag in 1967.

Second, if M is a (closed or cusped) hyperbolic 3-manifold with a geometrically finite incompressible surface, then M has a quasi convex hierarchy. Since graph groups are (virtually) RFRS it follows from Agol's theorem that Haken hyperbolic 3-manifolds are virtually fibered.

2. Cube complexes and the malnormal quasi-convex hierarchy theorem

2.1. **Cube complexes and hyperplanes.** An n-cube is defined to be $[-1,1]^n$. A *cube complex* is a CW-complex built out of cubes glued along subcubes. A *flag complex* is a simplicial complex such that $n+1$-vertices span an n-simplex if and only if they are pairwise adjacent.

A cube complex X is called *non positively curved* if for each $x \in X^0$ (i.e. each vertex) the link of x is a flag complex (see Figure 1). (Roughly

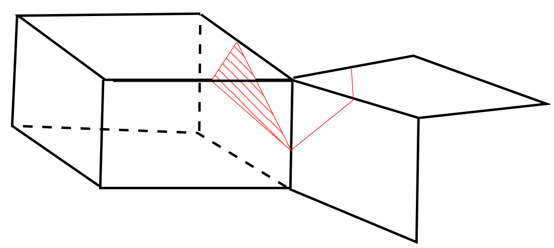


Figure 1. Non-positively curved cube complex with link at a vertex x.

speaking, the link of x is the ε-sphere around x.)

This notion of non positively curved is due to Gromov. Note that if X is a 2-dimensional cube complex, then X is non positively curved if and only if for any $x \in X^0$ the girth of $\text{link}(x)$ is at least 4 (recall that the girth is the infimum of the lengths of closed loops).

Remark. If X is a non positively curved cube complex, then one can equip X with a non positively curved metric such that each n-cube is isometric to the standard cube, in particular its universal cover $\tilde{X}$ can be viewed as a CAT(0) space. At least in these lectures the combinatorial version will be used throughout.

Throughout the lecture a CAT(0) cube complex will be understood to be a non positively curved cube complex which is simply connected. Examples of CAT(0) cube complexes are trees and the tiling of $\mathbb{R}^n$ by

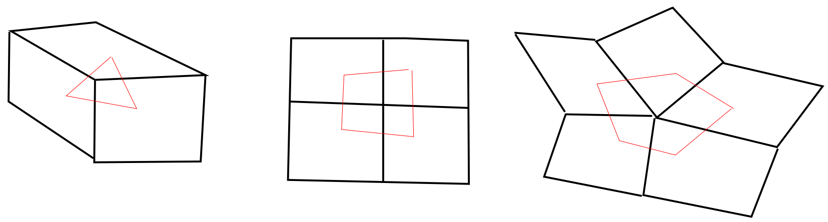


FIGURE 2. 2-dimensional cube complexs.

n-cubes (e.g. for the tiling of $\mathbb{R}^3$ by n-cubes the link at each vertex is an octahedron). Also note that the product of finitely many CAT(0) cube complexes is again a CAT(0) cube complex.

Slogan: CAT(0) cube complexes should be viewed as generalizations of trees. What's special about trees: any vertex separates a tree into two components. We will now see that CAT(0) cube complexes have a similar structure, namely any hyperplane cuts it into two components.

Here a *hyperplane* of a CAT(0) cube complex $\tilde{X}$ is a connected subspace which intersects each cube either in a single midcube (i.e. a subspace of $[-1, 1]^n$ obtained by restricting one coordinate to 0) or in the empty set.

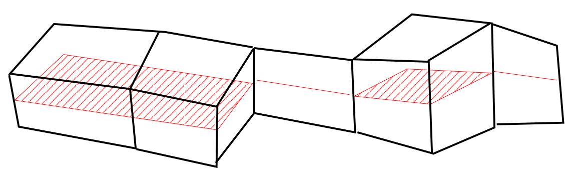


FIGURE 3. Example of a hyperplane.

Sageev proved the following theorem.

Theorem 2.1. *Let $\tilde{X}$ be a non positively curved cube complex.*

(1) *Every midcube in $\tilde{X}$ lies in a unique hyperplane,*
(2) *hyperplanes are themselves CAT(0) cube complexes,*
(3) *hyperplanes separate $\tilde{X}$ into two components,*
(4) *given a hyperplane its cubical neighborhood (i.e. the union of all cubes intersecting the hyperplane) is convex.*

(Note that we can put the usual metric on the 1-skeleton $\tilde{X}^1$ and we say $N \subset \tilde{X}$ is convex *if for any geodesic $\gamma \in \tilde{X}^1$ with endpoints in N^0 we already have $\gamma \subset N$.)*

2.2. **Disk diagrams and shuffling.** A *disk diagram* D is a compact planar 2-complex which is simply connected. The embedding in $\mathbb{R}^2$ determines a boundary cycle $\partial_p D \to D$.

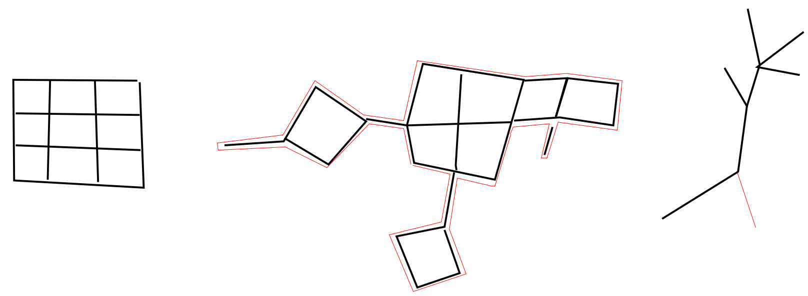


FIGURE 4. Example of disk diagrams with $\partial_p D$ for the middle disk diagram.

If $P \to Y$ is a closed combinatorial graph in a complex Y, then P is null-homotopic if and only if there exists a disk diagram D and a map $f : D \to Y$ which sends i-cells to i-cells such that the restriction of f to $\partial_p D$ equals p.

The hope is that a 'nice' structure on Y pulls back to a nice structure on D. For example, if Y is a non positively curved cube complex, then D is also cube complex, but not necessarily non positively curved.

Figure 2.2 shows the *hexagon move* on disk diagrams. This can

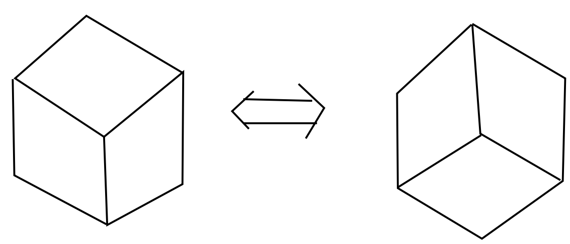


FIGURE 5. Hexagon move.

be described as 'replacing the front of a cube by the back of a cube'. Another move is the removal of bigons shown in Figure 2.2. A sequence of hexagon moves and removing bigons is called *shuffling*. Note that if $f : D \to Y$ is a map to a non positively curved cube complex and D' is obtained from D through shuffling, then there exists a canonical map $f' : D' \to Y$. The idea is to simplify disk diagrams via shuffling.

A *dual curve* in a cubulated disk diagram is defined analogously to a hyperplane above. Examples of pathologies are given in diagram 2.2.

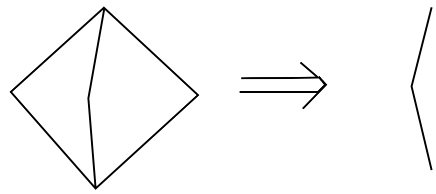


FIGURE 6. Removing bigons.

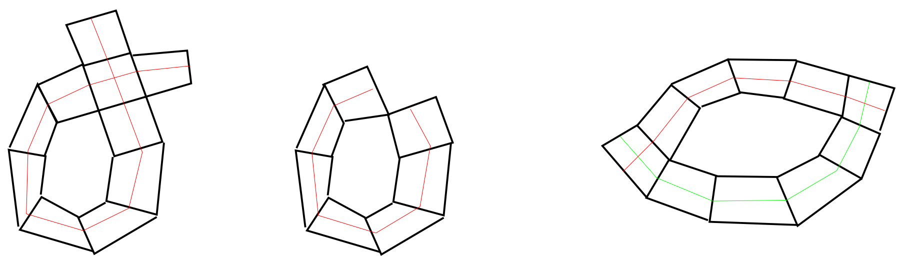


FIGURE 7. Self crossing dual curve, non-gon and bigon of dual curves. (The blank center can be filled in any way.)

Note the if D has a nongon, it also has a bigon of dual curves (see Figure 2.2).

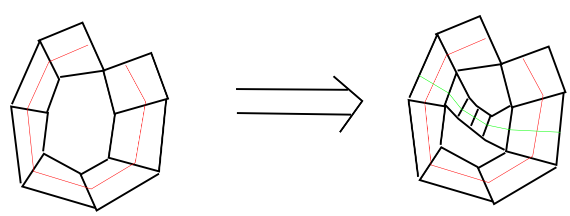


FIGURE 8. A non-gon gives rise to a bigon of dual curves.

For example Casson observed the following.

Lemma 2.2. *Let $P \to X$ be a null-homotopic path in a non positively curved cube complex then there exists a cubulated disk diagram $D \to X$ such that P equals its boundary path and such that D has no self-crossing loops and no bigons of dual curves.*

The idea is to start out with any disk diagram which is a cube complex, and then improve the cube complex by shuffling. In the example we first have three hexagon moves and then we remove a bigon as in Figure 2.2.

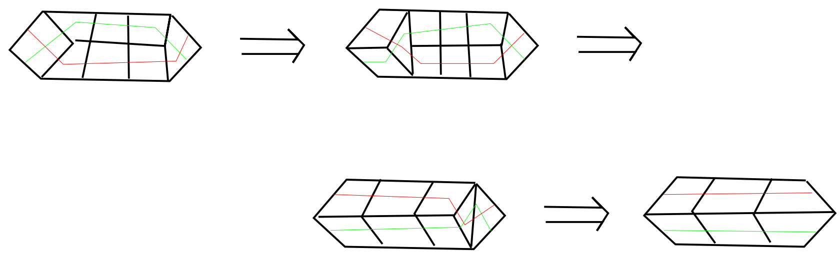


FIGURE 9. Example of shuffling.

2.3. **Special cube complexes.** A non positively curved cube complex X is *special* if its immersed hyperplanes do not have any of the pathologies described in Figures 2.3 and 2.3. The slogan is that hyperplanes

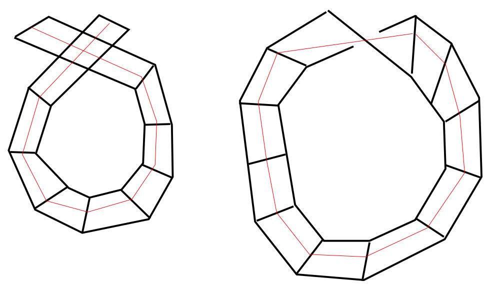


FIGURE 10. Self crossing and one-sidedness.

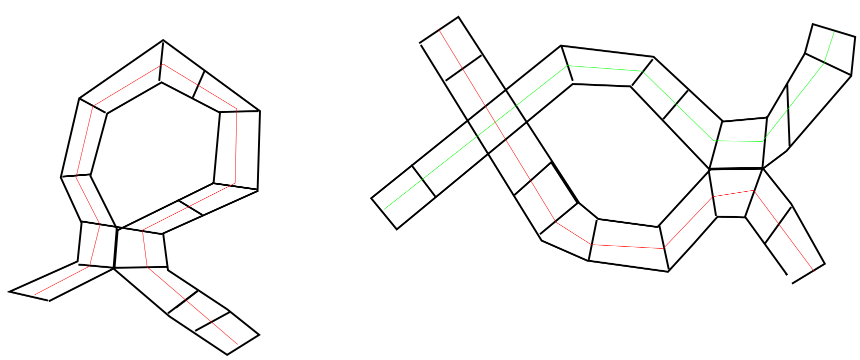


FIGURE 11. Self-osculating and inter osculating.

should behave as they would in a CAT(0) cube complex.

Given a graph Γ we can associate to it a cube complex $R(\Gamma)$ as follows: The 2-skeleton is the usual 2-complex associated to the presentation of the graph group $G(\Gamma)$. We then add an n-cube for any complete graph $K(n)$ with n-vertices in an appropriate way. The resulting cube complex is called the Salvetti complex. It is non positively curved and turns out to be special. For example, if Γ is the complete graph on n vertices, then $R(\Gamma)$ is just the n-torus.

A map $\varphi : X \to R$ between non positively curved cube complexes is a *local isometry* if for any $x \in X^0$ the map $\mathrm{link}(x) \to \mathrm{link}(\varphi(x))$ is an embedding of a full subcomplex. (I.e. if vertices in $\varphi(\mathrm{link}(x))$ span a simplex in $\mathrm{link}(\varphi(x))$ then they already span a simplex in $\mathrm{link}(x)$.) For example the map of Figure 2.3 is *not* a local isometry.

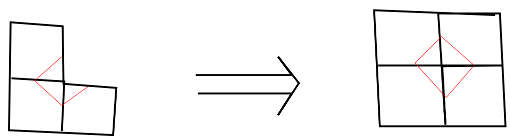


FIGURE 12. Example of a map which is not a local isometry.

By 'shuffling' one can show the following.

Lemma 2.3. *Let $\varphi : Y \to X$ be a local isometry between non positively curved cube complexes. Denote the universal covers by $\tilde{Y}$ and $\tilde{X}$. Then there is an induced embedding $\tilde{Y} \to \tilde{X}$ such that $\tilde{Y} \subset \tilde{X}$ is a convex subcomplex, in particular $\pi_1(Y) \to \pi_1(X)$ is injective. Furthermore, if Y and X are compact, then the embedding $\tilde{Y} \to \tilde{X}$ is in fact a quasi-isometric embedding.*

The following theorem is due to Haglund and Wise:

Theorem 2.4. *Let X be a non positively curved cube complex. Then X is special if and only if there exists a graph Γ and a local isometry $X \to R(\Gamma)$.*

In fact the graph Γ is constructed as follows: the vertices of Γ correspond to hyperplanes of X and two vertices are adjacent in Γ if the corresponding hyperplanes cross. Note that if $X = R(\Gamma)$ for some graph Γ (with no loops at a vertex and at most one edge between two vertices), then the above construction returns the original graph Γ.

The following is a fundamental property of special cube complexes.

Theorem 2.5. *If X is special and if $\varphi : Y \to X$ is a local isometry with Y compact, then there exists a finite cover $\hat{X} \to X$ and a lift $\hat{\varphi} : Y \to \hat{X}$ which admits a retraction $\hat{X} \to Y$.*

The idea of the construction of $\hat{X}$ can be seen in Figure 2.3. The cover $\hat{X}$ is sometimes referred to as the *canonical completion*.

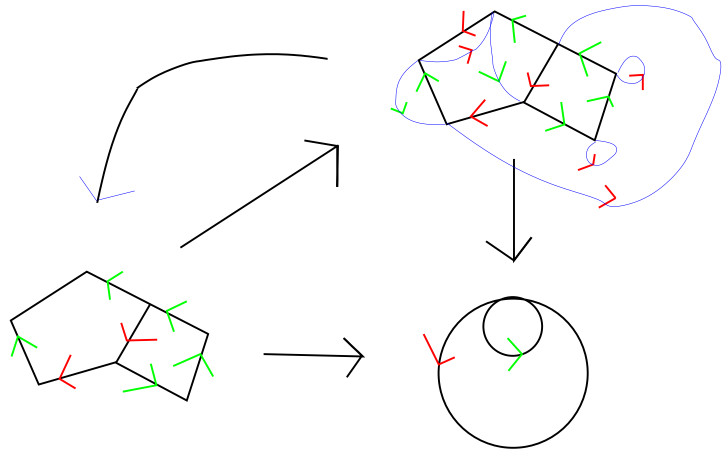


FIGURE 13. Example of a lift $Y \to \hat{X}$.

Corollary 2.6. *Suppose X is compact and special and let $H \subset \pi_1(X)$ be a subgroup. If there exists a compact subspace $Y \subset X$ together with a local isometry $Y \to X$ such that $H = \pi_1(Y)$, then H is separable.*

The corollary follows from the theorem and the fact that retracts of residually finite groups are separable.

The following is the double coset criterion for a cube complex to be virtually special. Note that it gives a completely algebraic criterion for virtual specialness.

Theorem 2.7. *Let X be a non positively curved compact cube complex. Then X is virtually special if and only if double hyperplane cosets are separable. More precisely, if for any hyperplanes A and B the subset $\pi_1(A)\pi_1(B) \subset \pi_1(X)$ is separable.*

It is conjectured that in fact X is virtually special if any single hyperplane subgroup is separable, i.e. if for any hyperplane A the subset $\pi_1(A) \subset \pi_1(X)$ is separable.

For the following recall that a subgroup $M \subset G$ is called *(almost) malnormal* if for any $g \in G \setminus M$ the group $gMg^{-1} \cap M$ is trivial (respectively finite). The following theorem is due to Haglund.

Theorem 2.8. *(Specializing theorem) Suppose X is a non positively curved compact cube complex and $Y \subset X$ an embedded hyperplane. Furthermore suppose that $\pi_1(X)$ is word hyperbolic and $\pi_1(Y) \subset \pi_1(X)$ is malnormal. If the components of $X \setminus N^0(Y)$ are virtually special, then X is virtually special.*

(Note that the special case of $F_2 *_M F_2$ has been done earlier by Wise in about 1998).

The following theorem is due to Hsu and Wise.

Theorem 2.9. *(Cubulating theorem) Suppose that $G = A *_C B$ or $G = A*_{C^t=C'}$, that G is word hyperbolic and that C is malnormal and quasi-convex in G. Furthermore suppose that A and B are virtually special, then there exists a compact non positively curved cube complex X with $G = \pi_1(X)$.*

Putting the specializing theorem and the cubulating theorem together we get the following 'fundamental tool'.

Theorem 2.10. *If G has a malnormal quasi convex hierarchy, then G is virtually special.*

Remark. (1) Note that G is in fact word hyperbolic by a theorem due to Bestvina.
(2) Note that Haken 3-manifolds do in general not admit a malnormal hierarchy, for example there exist Haken 3-manifolds without malnormal surface groups.

For the proof of the theorem we will need the following fact which comes from the double coset criterium. If G is word hyperbolic and $G = \pi_1(X)$ where X is virtually special, then any other non positively curved compact cube complex X' with $G = \pi_1(X')$ is also virtually special.

Proof. We can prove the theorem by induction on the complexity of the hierarchy. So suppose that $G = A *_C B$ where A, B are virtually special by induction. This means that A, B are fundamental groups of compact non positively curved cube complexes. By the cubulation theorem $G = \pi_1(X)$ for some non positively curved compact cube complex X. Moreover we know that X contains an embedded hyperplane Y with $\pi_1(Y) = C$ and such that the fundamental groups of $X \setminus N(Y)$ are A and B. By the aforementioned fact the components of $X \setminus N^0(Y)$ are also virtually special, hence we can apply the specializing theorem. □

3. The special quotient theorem

The main goal now is to reduce Theorem 1.1 to Theorem 2.10. The outline of the reduction argument is similar to Agol-Groves-Manning, but considerably more intricate. The main tool will be the special quotient theorem 3.5, and in order to prove the special quotient theorem we first have to introduce small cancelation theory for disk diagrams.

3.1. **Small cancelation theory for disk diagrams.** We now outline the small cancelation theory for disk diagrams. The following sections are particularly sketchy.

Let $\tilde{X}$ be a simply connected 2-complex. A *piece* in $\tilde{X}$ is subpath on the overlap between two 2-cells (see Figure 3.1).

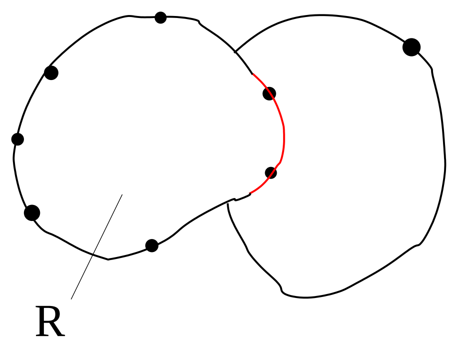


FIGURE 14. Example of a piece.

We say that $\tilde{X}$ satisfies $C'(1/6)$ if for any piece p which sits on the boundary of a 2-cell R we have

$$|p| \leq \frac{1}{6}|\partial R|,$$

where the length of a path is the number of one-cells it intersects (e.g. in Figure 3.1 we have $|p| = 3$ and $|\partial R| = 7$). Note that if $\tilde{X}$ satisfies $C'(1/6)$, then any minimal area disk diagram also satisfies $C'(1/6)$. (Slogan: if a disk diagram satisfies $C'(1/6)$, then it compares favorably to hyperbolic tiling.)

Figure 3.1 should explain what we mean by an i-shell and a spur of a disk diagram. More precisely, an *i-shell* in a disk diagram D is a 2-cell R such that the boundary consists of two components Q and S such that Q is a subpath of the boundary of D and S is internal and a concatenation of i-pieces. This notation goes back to work of McCammond and Wise.

The following lemma can be proved using the combinatorial Gauss-Bonnet Theorem:

Lemma 3.1. *(Greendlinger's lemma and Ladder theorem)*

(1) *If D is a $C'(1/6)$ disk diagram, then it has at least 2π worth of i-shells and spurs. Here an i-shell counts with π for $i = 0, 1$, a 2-shell counts as $2\pi/3$, an i-shell counts as $\pi/3$ for $i > 2$ and a spur counts as π.*

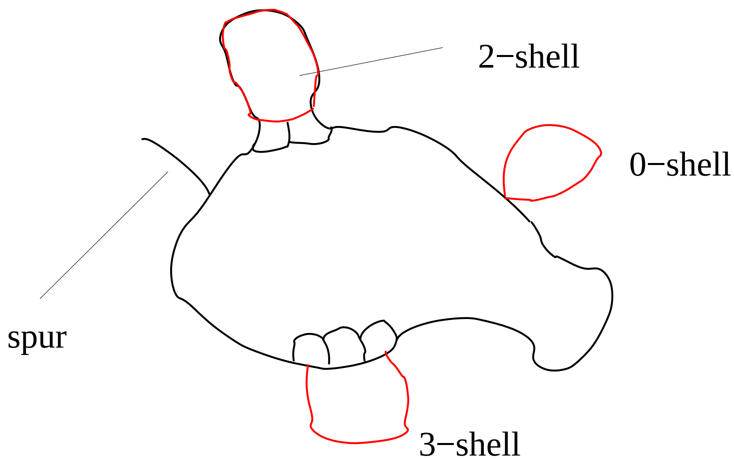


FIGURE 15. i-shells and spurs.

(2) *If D has fewer than three shells and spurs, then D is either a point, a 2-cells or a ladder (see Figure 2).*

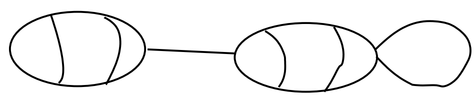


FIGURE 16. Example of a ladder.

3.2. **Small cancelation theory for cube complexes.** A *cubical relation presentation* is given as follows:

$$\langle X \mid Y_1, \ldots, Y_r \rangle$$

where X is a non positively curved cube complex, the Y_i are non positively curved cube complexes together with maps $\varphi_i : Y_i \to X$ which are local isometries. Given such a cubical relation presentation we define

$$X^* := \left(X \cup \bigcup_{i=1}^{r} C(Y_i) \right) / \varphi_i(y_i) \sim (y_i, 1),$$

where $C(Y_i)$ denotes the cone on Y_i. Note that

$$\pi_1(X^*) = \pi_1(X)/\langle\langle \pi_1(Y_1), \ldots, \pi_1(Y_r) \rangle\rangle.$$

Schematically we get the picture of Figure 3.2.

Recall that in small cancelation theory for disk diagrams we want the overlaps between contiguous 2-cells to be small. Now we consider maps

$$(D, \partial D) \to (X^*, X).$$

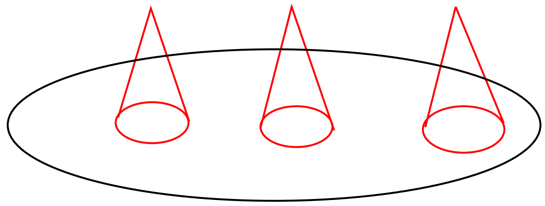


FIGURE 17. Cubical relation presentation.

Note that X is cubulated whereas we view the cones of $X^* \setminus X$ as triangulated. This pulls back to a decomposition of D as a union of squares and triangles (see Figure 3.2). We now number the cones (with

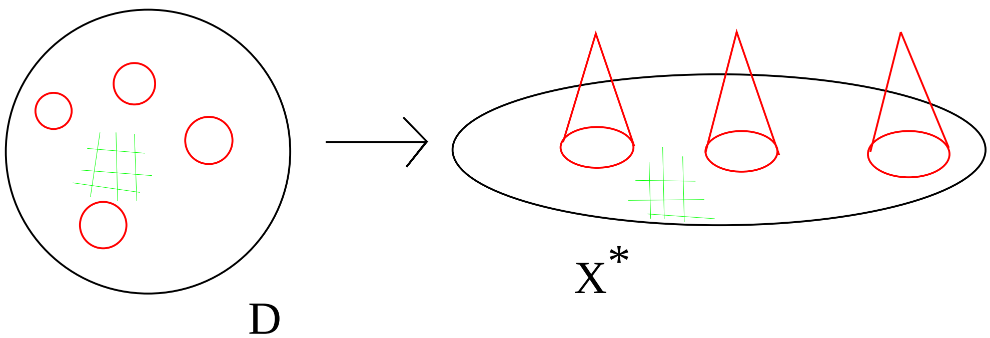


FIGURE 18. Pulling back squares and cones to D.

the cone at infinity counted last) and beginning with the first cone we number the 1-cells on the boundary of each cone. A sequence of squares emanating from the first 1-cells is called an *admitted rectangle*, this sequence ends on one of the cones. Starting with the next 1-cell we again look at a sequence of squares, ending either at a cone or at the previous admitted rectangle. The resulting sequence is again called an admitted rectangle. We continue this way till we decompose the disk diagram into 1-cells, admitted rectangles and the leftovers, which are called shards. This process is illustrated in Figure 3.2. (The definition of shard I give here is a little sloppy). The *pieces* of a cone cell are now overlaps with other cone cells (*contiguous piece*), or the union of boundaries of rectangles which have the same destination (*cone piece* if they go to another cone, *wall piece* if they go to another rectangle). See Figure 3.2.

Given a map $D \to X^*$ we define the area of D to be the pair

$$(\#\text{cone cells}, \#\text{squares}),$$

and we order these pairs lexicographically.

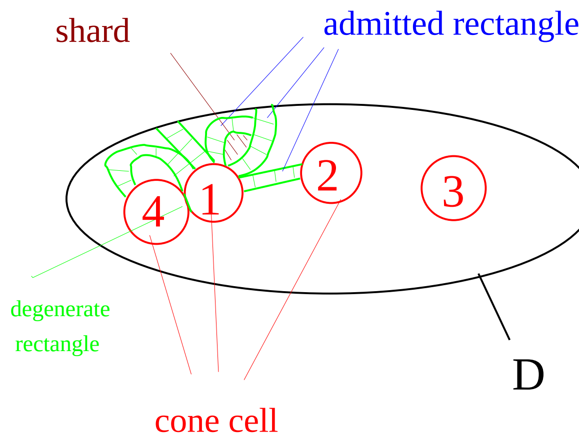


FIGURE 19. Cone cells, admitted rectangles and shards.

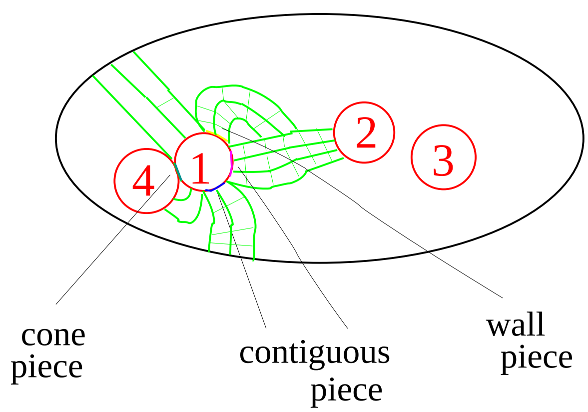


FIGURE 20. Pieces.

In many cases we can use shuffling to replace a map quasi-isometry $D \to X^*$ by another map, also quasi-isometric, which has lower area and which avoids certain pathologies.

The goal will be to formulate a version of Greendlinger's lemma in this context. In order to do this we have to assign angles. We will assign an angle of $\pi/2$ to corners of rectangles and angles of π at the other vertices of rectangles. To vertices on cone cells we assign the angle $\pi/2$ if the corresponding rectangles 'go to different places' (transition vertex), and π otherwise. This is summarized in Figure 3.2. The angles to vertices on shards are assigned such that in the interior of D the angles always add up to 2π and to a vertex on the boundary such that the sum of the angles of a shard is ≤ 0. (The slogan is:

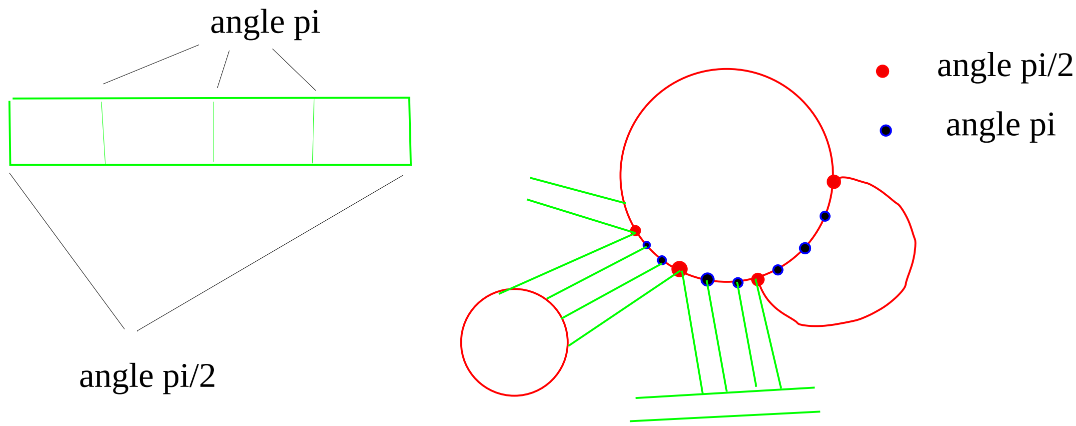


FIGURE 21. Angles assigned to vertices.

shards take care of themselves.) There are also various special cases not mentioned in the above, where angles $2\pi/3$ and $3\pi/4$ get assigned.

We now say X^* is $C'(1/12)$ if for any genuine piece p we have

$$|p| < \frac{1}{12}\mathrm{girth}(Y_i),$$

here the length of p is the defined to be the distance between the end points of the lift $\tilde{p} \subset \tilde{Y}_i$, and the girth of Y_i is defined to be the shortest essential closed path in Y_i. [1] Note that X^* is $C'(1/12)$ if we have many transitions.

We say $\tilde{Y} \to \tilde{X}$ is *superconvex* if $\tilde{Y}$ is convex and if any biinfinite geodesic γ in $\tilde{X}$ which lies within a bounded distance of $\tilde{Y}$ already lies in $\tilde{Y}$.

For example, if X is compact with $\pi_1(X)$ word hyperbolic, and if H is a quasi-convex subgroup of $\pi_1(X)$, then there exists a compact space Y and a local isometry $Y \to X$ such that $\tilde{Y}$ is superconvex and $H = \pi_1(Y)$.

It turns out that if $\tilde{Y}$ is superconvex, then we get upper bounds on the lengths of wall pieces and non-contiguous cone pieces. Furthermore, if the groups $\pi_1(Y_i)$ are malnormal, then we get upper bounds on the lengths of the contiguous pieces p. Hence to obtain the bound $|p| < \frac{1}{12}\mathrm{girth}(Y_i)$ it suffices to increase the girth of the Y_i. If the $\pi_1(Y_i)$ are residually finite, then we can pass to a finite cover $\hat{Y}_i \to Y$ to increase the girth sufficiently and the cubical relation presentation

$$\langle X \mid \hat{Y}_1, \ldots, \hat{Y}_r \rangle$$

[1] This seems to be a definition partly on $D \to X^*$, not sure what's going on.

will satisfy $C'(1/12)$.

3.3. **Greendlinger's lemma for cube complexes.** A *shell* is a cone cell which touches the boundary. Spurs and corners of generalized squares are indicated in Figure 3.3 The now assign a value of π to a

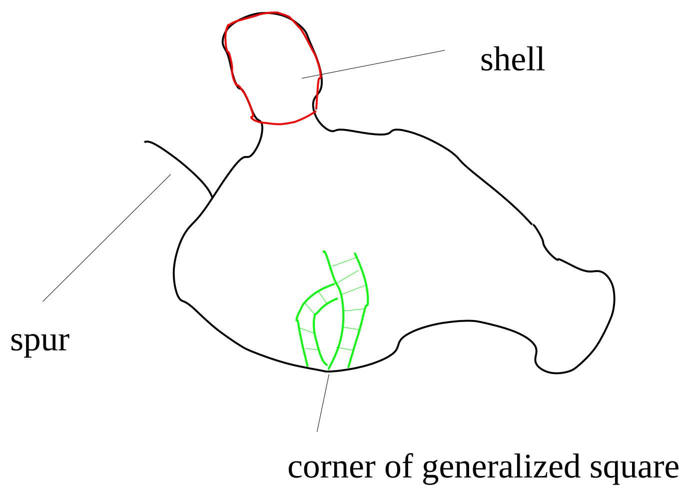


Figure 22. Shells, spurs and corners of generalized squares.

spur, to a shell we assign the value of

$$2\pi - \sum \text{angles at vertices}$$

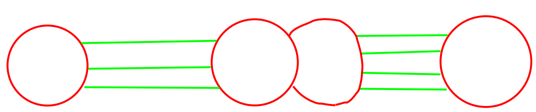
and to a corner of a generalized square we assign the angle as defined earlier. [2]

Theorem 3.2. *(Greendlinger's lemma) Suppose*

$$\langle X \mid Y_1, \ldots, Y_r \rangle$$

is a cubical relation presentation which is $C'(1/12)$. Suppose that $D \to X^*$ has a total of at least 2π worth of shells, spurs and corners of generalized squares, then D has minimal area.

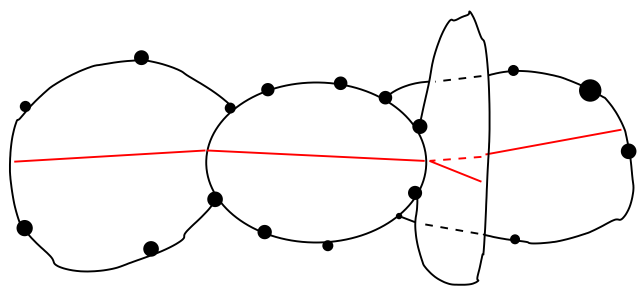
Furthermore, by a generalized ladder theorem, if D has only two features of positive curvature, then D is either a point, a single cone cell or it is a ladder (see Figure 3.3).



[2] At least that's sort of what's going on.

3.4. **Wall spaces.** Haglund and Paulin introduced the notion of a wall space. Loosely speaking, a wall space is a set together will a collection of partitions ('walls'), such that any two elements are separated by at most finitely many partitions.

If $\tilde{X}$ is a C'(1/12) 2-complex, then we will see that $\tilde{X}$ has a natural system of walls. Indeed, first note that we can arrange that any 2-cell has an even number of 1-cells in its boundary. Then starting with a midpoint of a one-cell we connect it to the midpoint of the opposite one-cell and we continue with the adjacent 2-cells (see Figure 3.4). Note that this way we obtain a graph $\tilde{X}$. Since we assumed that $\tilde{X}$



is a C'(1/12) 2-complex it follows from Greendlinger's lemma that we actually obtain trees which are quasi-isometrically embedded.

Consider again a cubical relation presentation

$$\langle X \mid Y_1, \ldots, Y_r \rangle$$

which is $C'(1/12)$. [3] To turn $\tilde{X}^*$ into a wall space we need to hypothesize that each Y_i is a wall space in the following sense: Assume that we partition the hyperplanes of Y_i into equivalence classes such that two hyperplanes in the same equivalence class do not cross each other and such that the union of hyperplanes in an equivalence class separates Y_i (plus a few more technical conditions.) Since we assume the $C'(1/12)$ condition we can use Greendlinger's lemma to show that in the universal cover $\tilde{X}^*$ of X^* the cones of X^* lift to embeddings in $\tilde{X}^*$ and immersed hyperplanes in X^* lift to embedded hyperplanes in $\tilde{X}^*$. Furthermore, we can extend equivalences of hyperplanes for the Y_i across the cones and then across $\tilde{X}^*$ (see Figure 3.4). The resulting hyperplanes are not necessarily separating, but we can count intersection numbers for all hyperplanes to construct a 2^n-fold cover in which the hyperplanes become separating.

[3] The following stuff is particularly opaque.

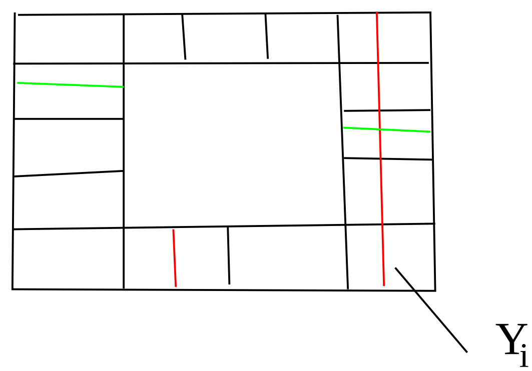
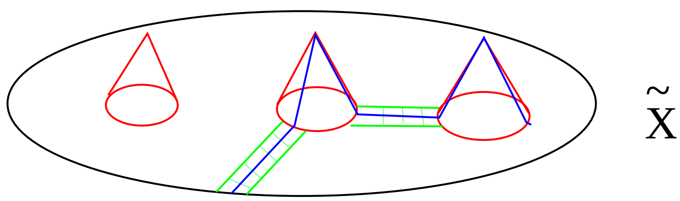


FIGURE 23. Two equivalence classes of hyperplanes in Y_i.



3.5. **Malnormal special quotient theorem.** The following is the second fundamental tool in proving the main theorem.

Theorem 3.3. *(Malnormal special quotient theorem) Suppose that G is word hyperbolic and virtually compact special and suppose that $\{H_1, \dots, H_r\}$ is a collection of malnormal (i.e. $H_i \cap H_j^g$ for any i, j and $g \in G$) quasiconvex subgroups. For any finite index subgroups $A_i \subset H_i$ there exist finite index subgroups $B_i \subset A_i$ such that*

$$G/\langle\langle B_1, \dots, B_r\rangle\rangle$$

is virtually special compact and word hyperbolic.

(The statement regarding word hyperbolic was known before.)

The idea for the proof is to find $B_i = \pi_1(Y_i)$ carefully such that $\langle X \,|\, \hat{Y}_1, \dots, \hat{Y}_r\rangle$ has virtually a malnormal quasi-convex hierarchy. We will sketch the proof in the next section.

The following is an almost immediate corollary:

Theorem 3.4. *Let $\langle a_1, \dots, a_s \,|\, W_1, \dots, W_r\rangle$ be a presentation. Then there exist $n_1, \dots, n_r$ such that $\langle a_1, \dots, a_s \,|\, W_1^{n_1}, \dots, W_r^{n_r}\rangle$ is virtually special, in particular residually finite.*

3.6. **Sketch of proof of malnormal special quotient theorem.** Since G is virtually special there exists a finite index subgroup $J \subset G$ which is torsion-free. Note that J acts properly and cocompactly on the CAT(0)-space $\tilde{X}$ which is the universal cover of the special cube complex. One can show that then G also acts properly and cocompactly on the CAT(0)-space $\tilde{X}$.

For each i let $\tilde{Y}_i$ be a superconvex H_i-cocompact subcomplex of $\tilde{X}$. We define $J_i := H_i \cap J$. We let $X = J \setminus \tilde{X}$ and $Y_i = J_i \setminus \tilde{Y}_i$. We now consider

$$\langle X \,| gY_1, \dots, gY_r : g \in G/J \rangle.$$

Given a finite cover $\hat{X}^* \to X^*$ we denote the induced covers of X and Y_i by $\hat{X}$ and $\hat{Y}_i$. We now pick a finite cover $\hat{X}^* \to X^*$ such that the following hold:

(1) hyperplanes in $\hat{X}$ have 'large collars', this implies in particular that
   (a) all hyperplanes in $\hat{X}$ embed, and
   (b) all hyperplanes in $\hat{X}$ are malnormal,
(2) all $\hat{Y}_i$ have large girth. (By the superconvexity and malnormality of the Y_i this implies that $\langle X | g\hat{Y}_1, \dots, g\hat{Y}_r \rangle$ is $C'(1/12)$.)

We now consider the map $\pi_1(\hat{X}) \to (\mathbb{Z}/2)^{w(\hat{X})}$ given by the $\mathbb{Z}/2$-intersection numbers with the walls of $\hat{X}$ and we denote the induced cover by $\ddot{X}$. We can now choose a wall structure on some finite cover $\ddot{Y}_i \to \hat{Y}_i$. We can now consider the cubical relation presentation

$$\ddot{X}^* = \langle \ddot{X} \mid g\ddot{Y}_1, \dots, g\ddot{Y}_r, g \in G \rangle.$$

This cubical relation presentation can be shown to admit a malnormal quasi convex hierarchy, and we can thus apply Theorem 2.10.

3.7. **Special quotient theorem.**

Theorem 3.5. *(Special quotient theorem) Suppose that G is word hyperbolic and virtually compact special and suppose that $\{H_1, \dots, H_r\}$ is a collection of quasiconvex subgroups (not necessarily malnormal). For any finite index subgroups $A_i \subset H_i$ there exist finite index subgroups $B_i \subset A_i$ such that*

$$G/\langle\langle B_1, \dots, B_r \rangle\rangle$$

is virtually special compact and word hyperbolic.

The statement regarding word hyperbolic was proved by Agol-Groves-Manning.

For the proof we need the notion of width. We say a subgroup $H \subset G$ has width n if there are distinct cosets $Hg_1, \dots, Hg_n$ such that $\cap H^{g_i}$ is infinite, but such that for any $n+1$ such cosets the corresponding intersection is finite. Note that a subgroup has width zero if and only if it is finite, and a subgroup has width one if and only if it is almost malnormal. Similarly one can define the width $\text{width}_G\{H_1,\dots,H_r\}$ for a collection of subgroups.

Gitik, Mitra, Rips and Sageev proved the fundamental result that if $H \subset G$ is quasi-convex, then H has finite width.

The following is now a sketch of the proof of the Special quotient theorem. Suppose we are given $\{G,\{H_1,\dots,H_r\}\}$ of width n. Let K_i be a collection of finitely many classes of n-fold intersections of conjugates of the form $\cap H_{k_i}^{g_{k_i}}$. Then $\{K_1,\dots,K_s\}$ is an almost malnormal collection of subgroups. We can thus apply the malnormal special quotient theorem to $\{G,\{K_1,\dots,K_s\}\}$ to obtain a virtually compact special quotient $\overline{G} = G/\langle\langle K_1',\dots,K_s'\rangle\rangle$ and with

$$\text{width}_{\overline{G}}\{\overline{H}_1,\dots,\overline{H}_r\} < \text{width}_G\{H_1,\dots,H_r\}.$$

We can now apply the induction to conclude the proof.

3.8. **Proof of main theorem.** Recall the statement of the main theorem:

Theorem 3.6. *If G is word hyperbolic and if G has a quasi convex hierarchy, then G is virtually compact special.*

As a warm-up example we first consider the case where $G = A *_C B$ *and where C is separable in G.* It C is malnormal, then we are done by Theorem 2.10. So suppose that C is not malnormal. We now consider a maximal coset of intersecting conjugates $Cg_1,\dots,Cg_m$ with the property that $C \cap C^{g_i}$ is infinite for $i = 1,\dots,m$. (This set is finite by the result of Gitik, Mitra, Rips and Sageev). By our separability assumption we can find an epimorphism $\alpha : G \to Q$ onto a finite group such that $\alpha(g_i) \notin \alpha(C)$ for $i = 1,\dots,m$. We write $G' = \text{Ker}(\alpha)$.

Note that the decomposition $G = A *_C B$ induces a canonical decomposition of G' as the fundamental group of a graph of groups (where the vertex groups are isomorphic to $A\cap G', B\cap G'$ and the edge groups are isomorphic to $C \cap G'$). Note that in the induced splitting of G' all edge groups are malnormal, and we can thus apply Theorem 2.10. This concludes the proof of the warm-up example.

We now turn to the general case where C is not separable. The idea is that we now use the special quotient theorem to obtain a convenient finite index subgroup of G. Instead of the case $G = A *_C B$ we now consider the case where $G = A*_{C^t=C'}$. We again consider a maximal coset of intersecting conjugates $Cg_1,\dots,Cg_m$ with the property that $C \cap C^{g_i}$ is infinite for $i = 1,\dots,m$.

The idea is to find a quotient

$$G = A*_{C^t=C'} \to \overline{A}*_{\overline{C}^t=\overline{C'}} = \overline{G}$$

such that the following hold:

(1) $\overline{g}_i \notin \overline{C}$,
(2) $\overline{C} \subset \overline{G}$ is quasi-convex,
(3) $\overline{A}$ is virtually compact special,
(4) the width of $\overline{C}$ in $\overline{G}$ is less than the width of C in G.

If we can arrange this, then we can argue by induction on the width to conclude that $\overline{G}$ is virtually special. In particular $\overline{G}$ is subgroup separable and we can find an appropriate finite quotient $\overline{G} \to Q$ as above.

How do we find the quotient $G = A*_{C^t=C'} \to \overline{A}*_{\overline{C}^t=\overline{C'}} = \overline{G}$? The first naive idea would be to just apply the special quotient theorem to $(A, \{C, C'\})$. But the result subgroups will not 'be the same on C and C''', i.e. we will not get an induced HNN-extension.

The trick is to apply the special quotient theorem to a more cleverly chosen situation, which will the give the required map.

4. The relative hyperbolic case

A non positively curved cube complex is called *sparse* if it is quasi-isometric to a wedge of finitely many Euclidean half spaces. Note that acting co-sparsely is robust in the following sense: if G' is a finite index subgroup of G and if G' acts co-sparsely on a CAT(0)-cube complex, then G' does so as well. Note that acting cocompactly is *not* robust, e.g. let $G = \langle a, b, c | a^2, b^2, c^2, (ab)^3, (bc)^3, (ca)^3\rangle$, then G does not act properly and cocompactly on a 2-dimensional CAT(0)-cube complex, but $\mathbb{Z}^2 \subset G$ (which is a subgroup of finite index) acts of course properly and cocompactly on $\mathbb{Z}^2$.

Conjecture 4.1. *Let G be word hyperbolic relative to virtually finitely generated abelian groups. Suppose G has a quasi-convex hierarchy, then G is virtually sparse special. (Or possibly virtually compact special.)*

(Note that virtually compact special does give separability, but for the proof we need compactness, i.e. just virtually special does *not* imply separability. But virtually sparse special does give separability.)

At the moment the following is state of the art:

Theorem 4.2. *If G is word hyperbolic relative to a collection of virtually finitely generated abelian groups and if G splits as a graph Γ of groups with quasi-isometrically embedded edge groups and virtually* *compact special word hyperbolic vertex groups, then G is virtually compact special.*

Note that this theorem is strong enough to handle the case of finite volume cusped hyperbolic 3-manifolds. More precisely, let M be any finite volume cusped (non-closed) hyperbolic 3-manifold. By Culler-Shalen there exists a geometrically finite surface which intersects all cusps, this surfaces gives rise to a quasi-convex hierarchy. Alternatively, we can go to a finite cover with $b_1(M') > 1$, take a non-fibered surface and arrange to find such a surface such that it hits any boundary torus (requires some thought).