NON-POSITIVELY CURVED CUBE COMPLEXES

Henry Wilton

Last updated: June 29, 2012

1 Some basics of topological and geometric group theory

1.1 Presentations and complexes

Let Γ be a discrete group, defined by a presentation $\mathcal{P} = \langle a_i \mid r_j \rangle$, say, or as the fundamental group of a connected CW-complex X.

Remark 1.1. Let $X_{\mathcal{P}}$ be the CW-complex with a single 0-cell E^0, one 1-cell E^1_i for each a_i (oriented accordingly), and one 2-cell E^2_j for each r_j, with attaching map $\partial E^2_j \to X^{(1)}_{\mathcal{P}}$ that reads off the word r_j in the generators $\{a_i\}$. Then, by the Seifert–van Kampen Theorem, the fundamental group of $X_{\mathcal{P}}$ is the group presented by $\mathcal{P}$.

Conversely, restricting attention to $X^{(2)}$ and contracting a maximal tree in $X^{(1)}$, we obtain $X_{\mathcal{P}}$ for some presentation $\mathcal{P}$ of Γ.

So we see that these two situations are equivalent. We will call $X_{\mathcal{P}}$ a *presentation complex* for Γ. Here are two typical questions that we might want to be able to answer about Γ.

Question 1.2. *Can we compute the homology $H_*(\Gamma)$ or cohomology $H^*(\Gamma)$?*

Question 1.3. *Can we solve the word problem in Γ? That is, is there an algorithm to determine whether a word in the generators represents the trivial element in Γ?*

1.2 Non-positively curved manifolds

In general, the answer to both of these questions is 'no'. Novikov (1955) and Boone (1959) constructed a finitely presentable group with an unsolvable word problem, and Cameron Gordon proved that the rank of $H_2(\Gamma)$ is not computable. So we would like to find large classes of group in which we can find ways to solve these problems. One such class is provided by geometry.

Theorem 1.4 (Cartan–Hadamard)**.** *The universal cover of a complete m-dimensional Riemannian manifold X of non-positive sectional curvature is diffeomorphic to $\mathbb{R}^m$, in particular is contractible.*

By the long exact sequence of a fibration, such a manifold X is a $K(\Gamma, 1)$ and so we have $H_*(\Gamma) \cong H_*(X)$. The homology of a complex is relatively easy to compute. So if Γ happens to be the fundamental group of a non-positively curved manifold then we can hope to answer Question 1.2.

The limitations of this idea are immediately apparent. The homology of a closed manifold satisfies Poincaré Duality, but there are many groups that do not.[1]

1.3 An aside: the geometry of the word problem

In fact, non-positive curvature also helps with Question 1.3. The material required to explain this is not all directly relevant to the rest of this course. However, it provides a useful opportunity to introduce some of the foundational concepts of modern geometric group theory.

Definition 1.5. Let S be a generating set for Γ. The *Cayley graph* of Γ with respect to S is the 1-skeleton of the universal cover of any presentation complex for Γ with generators S. Equivalently, $\mathrm{Cay}_S(\Gamma)$ is the graph with vertex set Γ and with an edge joining γ_1 and γ_2 for each $s \in S^{\pm 1}$ such that $\gamma_1 = \gamma_2 s$.

The Cayley graph can be given a natural length metric by declaring each edge to be of length one. This metric is called the *word metric* on Γ and denoted d_S. The natural left action of Γ is free, properly discontinuous and by isometries, and if S is finite then it is also cocompact.

This is a natural geometric object associated to a group with a generating set. It would be good if this object was actually an invariant of the group, but changing the generating set will change the Cayley graph. The key observation here is that if S is finite then changing the generating set only changes the metric on Γ by a Lipschitz map. To put this observation into its proper context, we need a definition.

Definition 1.6. A map of metric spaces $f : X \to Y$ is a (λ, ϵ)-*quasi-isometric embedding* if it satisfies

$$\frac{1}{\lambda} d_X(x_1, x_2) - \epsilon \leq d_Y(f(x_1), f(x_2)) \leq \lambda d_X(x_1, x_2) + \epsilon$$

for all $x_1, x_2 \in X$. If it is also *quasi-surjective*, in the sense that for every $y \in Y$ there is an $x \in X$ with $d_Y(y, f(x)) \leq \epsilon$, then f is called a *quasi-isometry* and X and Y are called *quasi-isometric.*

The following result has been called the 'Švarc–Milnor Lemma' and also the 'Fundamental Observation of Geometric Group Theory'. A metric space is called *proper* if closed balls are compact. A *geodesic* in X is just an isometrically embedded closed interval $\gamma : I \to X$ or, more generally, such an embedding after rescaling. The metric space X is called *geodesic* if every pair of points is joined by a geodesic.

[1] Indeed, it is an open question whether every finitely presentable Poincaré Duality group is the fundamental group of an aspherical manifold.

Lemma 1.7. *Let X be a proper (ie closed balls are compact), geodesic metric space and let Γ be a group that acts properly discontinuously and cocompactly on X by isometries. Then Γ is finitely generated and, for any $x_0 \in X$, the map $\gamma \mapsto \gamma x_0$ is a quasi-isometry $(\Gamma, d_S) \to (X, d_X)$.*

Proof. Let C be such that the Γ-translates of $B(x_0, C)$ cover X. It is easy to show that the finite set

$$S = \{\gamma \in \Gamma \mid d_X(x_0, \gamma x_0) < 3C\}$$

generates. Note that $d_X(x_0, \gamma x_0) < 3C d_S(1, \gamma)$ for any $\gamma \in \Gamma$. For the other inequality, consider the geodesic $[x_0, \gamma x_0]$, choose points $x_i \in [x_0, \gamma x_0]$ at distance C apart and let $d_X(x_i, \gamma_i x_0) < C$. Then we see that $d(x_0, \gamma_{i+1}\gamma_i^{-1}x_0) < 3C$ for each i and so $\gamma_{i+1}\gamma_i^{-1} \in S$. As we need at most $\lceil d_X(x_0, \gamma x_0)/C \rceil$ points x_i, we see that

$$d_S(1, \gamma) \leq d_X(x_0, \gamma x_0)/C + 1$$

as required. ◻

The hypothesis that X is geodesic is rather strong. In fact, a similar proof works in the much more general context of length spaces, which we will encounter later.

As a special case, if Γ is the fundamental group of a closed Riemannian manifold (or, more generally, any length space) then Γ is quasi-isometric to the universal cover $\widetilde{X}$.

The Švarc–Milnor Lemma shows that the quasi-isometry class of metric spaces on which a finitely generated group acts nicely is an invariant of the group. One of the key themes of this course is to try to choose a good representative $\widetilde{X}$ on which Γ acts.

What has this to do with the word problem? From now on, we will assume that the presentation $\langle S \mid R \rangle$ we are considering is finite. In order to study the word problem a little more deeply, we need some more definitions.

Definition 1.8. For a word $w \in F(S)$ that represents the trivial element in Γ, define $\mathrm{Area}_{\mathcal{P}}(w)$ to be the minimal N such that

$$w = \prod_{k=1}^{N} g_k r_{j_k} g_k^{-1} \ ;$$

that is, the area is the smallest number of conjugates of relators needed to prove that w represents the trivial element.

Definition 1.9. The *Dehn function* $\delta_{\mathcal{P}} : \mathbb{N} \to \mathbb{N}$ is defined by

$$\delta_{\mathcal{P}}(n) = \max_{l_S(w) \leq n} \mathrm{Area}_{\mathcal{P}}$$

where w only ranges over words of $F(S)$ that represent the trivial element in Γ.

Computationally, the Dehn function can be thought of as quantifying the difficulty of the word problem in $\mathcal{P}$. Indeed, the following lemma is almost obvious.

Lemma 1.10. *The word problem in $\mathcal{P}$ is solvable if and only if $\delta_{\mathcal{P}}$ is computable.*

Once again, we would like this to be an invariant of the group, rather than just the presentation, and to this end we again need a coarse equivalence relation. For functions $f, g : \mathbb{N} \to [0, \infty)$, we write $f \preceq g$ of there is a constant C such that

$$f(n) \leq Cg(Cn + C) + (Cn + C)$$

for all n, and $f \simeq$ if $f \preceq g$ and $g \preceq f$. The $\simeq$-class of $\delta_{\mathcal{P}}$ turns out to be a quasi-isometry invariant, and so it makes sense to write δ_Γ as long as we bear in mind that we are really talking about an equivalence class of functions.

In a Riemannian manifold M, the analogue of the Dehn function is related to a beautiful geometric problem. Let γ be a null-homotopic loop. Then for any disc D bounded by γ we can measure the area, and thereby define a function analogous to the Dehn function.

Definition 1.11. Let M be a Riemannian manifold. For a null-homotopic loop γ, $\mathrm{Area}_M(\gamma)$ is the infimal area of a filling disc for γ. The *filling function* $\mathrm{Fill}_M : [0, \infty) \to [0, \infty)$ is now defined to be

$$\mathrm{Fill}_M(x) = \sup_{l(\gamma) \leq x} \mathrm{Area}_M(\gamma)$$

where γ ranges over all suitable null homotopic loops.

The Filling Theorem, stated by Gromov and proved by various authors, provides an analogue of Lemma 1.7 in this context.

Theorem 1.12 (Gromov, Bridson, Burillo–Taback)**.** *If M is a closed Riemannian manifold then* $\mathrm{Fill}_M \simeq \delta_{\pi_1(M)}$.

The final piece of the jigsaw is the fact that manifolds of non-positive satisfy have a quadratic isoperimetric inequality. This well known fact can be proved, for instance, using the convexity of the metric (see below). Using this, it is not hard to prove that one can cut a polygon into triangles of bounded diameter, and the number of triangles is at worst quadratic in the perimeter of the polygon.

2 CAT(0) spaces and groups

In the last section, we learned that non-positive curvature can have useful group-theoretic applications, but that its applicability is limited. In this section we will describe a more general geometric condition that can be applied to a wide variety of groups.

2.1 The CAT(κ) condition

A CAT(κ) space is, roughly speaking, a metric space in which triangles are at least as thin as triangles in a space of curvature κ. In order to make this idea precise, we need to develop some notation.

We shall denote by M_κ the unique connected, complete, 2-dimensional Riemannian manifold of constant curvature κ. That is,

- $M_{+1} \cong S^2$;
- $M_0 \cong \mathbb{R}^2$;
- $M_{-1} \cong \mathbb{H}^2$.

The remaining M_κ's are just rescaled copies of these three.

Because we need to work in a more general setting than manifolds, we shall work with a complete, geodesic, proper metric space (X, d). We will often abuse notation and denote a geodesic between two points p and q by $[p, q]$, even though *a priori* this geodesic may not be unique.

The definition of CAT(κ) geometry is motivated by the idea that triangles capture most of the interesting information about a geometry. A *triangle with vertices* $\{p, q, r\}$ is the union of three geodesics $[p, q] \cup [q, r] \cup [r, p]$. Again, we will often abuse notation and refer to a triangle with vertices $\{p, q, r\}$ by $\Delta(p, q, r)$, even though such a triangle may not be unique.

Let D_κ be the diameter of M_κ; for instance, $D_\kappa = \infty$ for $\kappa \leq 0$ and $D_1 = \pi$. Let $\Delta = \Delta(x_1, x_2, x_3)$ be a geodesic triangle in X, and suppose that the perimeter of Δ is at most $2D_\kappa$. Then there is, up to isometry, a unique triangle $\overline{\Delta} = \overline{\Delta}(\bar{x}_1, \bar{x}_2, \bar{x}_3) \subseteq M_\kappa$ with $d(x_i, x_j) = d_{M_\kappa}(\bar{x}_i, \bar{x}_j)$, which we will call the *comparison triangle* for Δ. There is a surjection $\overline{\Delta} \to \Delta$ that restricts to an isometry on each edge, so given $y \in [x_i, x_j]$ there is a well defined *comparison point* $\bar{y} \in [\bar{x}_i, \bar{x}_j]$. (Note that the edges of Δ may intersect at points other than the endpoints, so each $y \in \Delta$ may have up to three comparison points in $\overline{\Delta}$.)

We are, at last, able to give the definition of a CAT(κ) space.

Definition 2.1. A complete, geodesic metric space (X, d) is CAT(κ) if, for any geodesic triangle Δ of perimeter at most $2D_\kappa$ and any $p, q \in \Delta$, the comparison points $\bar{p}, \bar{q} \in \overline{\Delta} \subseteq M_\kappa$ satisfy

$$d(p, q) \leq d_{M_\kappa}(\bar{p}, \bar{q}) \ .$$

If X is *locally* CAT(κ) then X is said to be *of curvature at most* κ. In particular, a locally CAT(0) space is called *non-positively curved.*

This definition was first given by Alexandrov in 1951. The terminology was coined by Gromov in 1987, in honour of Cartan, Alexandrov and Toponogov.

It is a straightforward exercise in the hyperbolic, Euclidean and spherical cosine rules to check that a CAT(κ) space is also CAT(λ) for all $\lambda \geq \kappa$. Although CAT(1) spaces will play an important role shortly, we will be most interested in CAT(0) spaces, which are precisely the spaces that exhibit non-positive curvature. A little later on, I will discuss why we are slightly less interested in CAT(-1) spaces.

Example 2.2. Here are some examples of CAT(0) spaces.

- Any inner product space.
- Any simply connected manifold of non-positive sectional curvature.
- In particular, symmetric spaces.
- Any tree is CAT(0) (indeed, CAT(κ) for all κ).
- If X and Y are CAT(0) then $X \times Y$, endowed with the l^2-metric, is CAT(0).
- So the first interesting example of a CAT(0) space is a product of two trees.

The convexity of the metric is an important property of CAT(0) spaces.

Lemma 2.3. *Let X be a* CAT(0) *space. If $\gamma, \delta : [0,1] \to X$ are geodesics then*

$$d(\gamma(t), \delta(t)) \leq (1-t)d(\gamma(0), \delta(0)) + td(\gamma(1), \delta(1)) \ .$$

Proof. If $\gamma(0) = \delta(0)$, this follows immediately from the CAT(0) inequality and the fact that the comparison geodesics satisfy

$$d(\bar{\gamma}(t), \bar{\delta}(t)) = d(\bar{\gamma}(1), \bar{\delta}(1)) \ .$$

For the general case, divide the quadrilateral into two triangles and apply the preceding case. □

As an immediate consequence, CAT(0) spaces are *uniquely geodesic.*

Lemma 2.4. *If X is a* CAT(0) *space then there is a unique geodesic joining each pair of points.*

As a consequence, geodesics vary continuously with their endpoints.

Lemma 2.5. *Let X be a proper, uniquely geodesic space. The geodesics vary continuously in the compact-open topology with their endpoints.*

Proof. Suppose that $x_n \to x$ and $y_n \to y$ as $n \to \infty$, and let $\gamma_n = [x_n, y_n]$ and $\gamma = [x, y]$. For convenience, we will take the domain of each γ_n and γ to be $[0, 1]$. We will start by showing that $\gamma_n \to \gamma$ pointwise. If not, then there is a t_0 and an $\epsilon > 0$ such that $d(\gamma(t_0), \gamma_{n_i}(t_0)) > \epsilon$ for some subsequence γ_{n_i}. All the γ_n are contained in some compact ball B of radius R, and so $d(\gamma_n(s), \gamma_n(t)) < 2R|s-t|$ for all s, t. The γ_n are an equicontinuous family of maps $[0, 1] \to B$ and hence, by the Arzelà–Ascoli Theorem, there is a subsequence of the (γ_{n_i}) that converges uniformly to a geodesic γ_∞ from x to y, a contradiction. Finally, note that the convergence must in fact be uniform. Let $\epsilon > 0$. If $d(\gamma_n(t_0), \gamma(t_0)) < \epsilon/3$ then we also have $d(\gamma_n(t), \gamma(t)) < \epsilon$ whenever $|s - t| < \epsilon/6R$. Using the compactness of $[0, 1]$, it follows that the convergence is uniform. □

Proposition 2.6. *Any* CAT(0) *space X is contractible.*

Proof. For each $y \in X$, let $\gamma(\cdot, y)$ be the unique geodesic from x to y. The map $F : X \times [0,1] \to X$ that sends $(y,t) \to \gamma(1-t, y)$ is a homotopy equivalence from X to $\{x\}$. □

Therefore, if we have a group acting freely and properly discontinuously on a CAT(0) space, we will know that the quotient is a classifying space.

Definition 2.7. A group Γ that acts freely, properly discontinuously and cocompactly by isometries on a proper CAT(0) space is called a CAT(0) *group*. In order to allow the possibility that Γ might have torsion, some authors do not require the action to be free.

Example 2.8. That the following groups are CAT(0) can be easily deduced from the examples of CAT(0) metric spaces given earlier.

- $\mathbb{Z}^n$ for any n.
- The fundamental group of any manifold of non-positive sectional curvature.
- Uniform lattices in semisimple Lie groups.
- Free groups.
- Any direct product of free groups.

By Proposition 2.6, if we are able to exhibit an explicit CAT(0) space on which our group acts then the quotient is a classifying space, from which we can compute the group homology and cohomology. So the CAT(0) condition helps us with Question 1.2. The CAT(0) condition also helps with Question 1.3. Indeed, we have the following proposition.

Proposition 2.9 ([3], Proposition III.Γ.1.6)**.** *If Γ is* CAT(0) *then the Dehn function of Γ is bounded above by a quadratic function. In particular, the word problem for Γ is solvable.*

The proof of this proposition is more or less identical to the proof outlined above for fundamental groups of manifolds of non-positive curvature. Recall that convexity of the metric was the key property.

The CAT(0) condition is restrictive enough to ensure that CAT(0) groups and spaces have many nice properties, but also flexible enough that CAT(0) groups can exhibit some pathological behaviour. For this reason, CAT(0) geometry is a source of many very interesting examples in modern topology and geometric group theory.

2.2 Another aside: hyperbolic spaces and groups

We will take a short while to discuss another famous thin-triangles condition introduced by Gromov, and to compare and contrast this condition with the CAT(κ) conditions. Again, we will work in a geodesic metric space X, and use the same notation as before.

Definition 2.10. A geodesic triangle $\Delta \equiv \Delta(x, y, z) \subseteq X$ is called *δ-slim* if each side is contained in the δ-neighbourhod of the union of the other two sides; that is,

$$[x, y] \subseteq N_\delta([x, z] \cup [y, z])$$

and similarly for $[y, z]$ and $[z, x]$.

Definition 2.11. The metric space X is *δ-hyperbolic* or *(Gromov-)hyperbolic* if every geodesic triangle in X is δ-slim.

Example 2.12. The following spaces are Gromov-hyperbolic.

- Any metric tree.
- The hyperbolic plane. This follows from the fact that semicircles inscribed in triangles are of uniformly bounded size, which in turn follows from the fact that triangles are of uniformly bounded area.

For a negative example, it is easy to see that $\mathbb{R}^2$ is not Gromov-hyperbolic.

Gromov-hyperbolicity and the CAT(-1) condition are both notions of negative curvature for a metric spaces, and indeed the above examples are also CAT(-1). This is not altogether surprising. It is immediate from the definitions and the above example that any CAT(-1) space is Gromov-hyperbolic.

To understand how CAT(-1) and Gromov-hyperbolicity differ, the first thing to notice is that Gromov-hyperbolicity is a coarse condition—any triangle of diameter less than δ is trivially δ-slim, so Gromov-hyperbolicity only places a restriction on the large triangles in a space. By contrast, the CAT(-1) condition (and, indeed, the CAT(0) condition) place non-trivial restrictions even on the small triangles in a space.

To make this observation precise, we shall appeal to a useful theorem. See, for instance, [3] for the proof.

Theorem 2.13. *Let X and Y be geodesic metric spaces. If X is hyperbolic and Y is quasi-isometric to X then Y is also hyperbolic.*

This follows easily from an important result which is sometimes called the Morse Lemma. A (λ, ϵ)-quasi-geodesic in X is a (λ, ϵ)-quasi-isometric embedding of an arc into X.

Lemma 2.14. *Suppose that X is δ-hyperbolic. Any (λ, ϵ)-quasigeodesic from x to y is contained in the R-neighbourhood of a geodesic from x to y, where R is a constant depending only on δ, λ and ϵ.*

From Theorem 2.13, it follows that many metric graphs are hyperbolic. For instance, we have the following.

Example 2.15. If Σ is a closed surface of constant Gaussian curvature -1 then the universal cover of Σ is $\mathbb{H}^2$, so for any finite generating set, the Cayley graph of $\pi_1\Sigma$ is quasi-isometric to the hyperbolic plane and hence is Gromov-hyperbolic.

In contrast, the restrictions on small triangles imposed by the CAT(0) inequality make it difficult for a graph to be CAT(0), let alone CAT(-1).

Exercise 2.16. Show that any CAT(0) metric graph is a tree.

Example 2.17. If Γ acts freely, properly discontinuously and cocompactly by isometries on a tree T then T/Γ is a graph and $\Gamma \cong \pi_1(T/\Gamma)$, from which it follows that Γ is free. In particular, the Cayley graph of $\pi_1(\Sigma)$ (for Σ as above) is not CAT(0).

To summarise, every CAT(-1) space is hyperbolic, but the converse is far from true.

We now turn our attention to groups. By Theorem 2.13, the following definition makes sense.

Definition 2.18. A finitely generated group is called *(word-)hyperbolic* if its Cayley graph is Gromov-hyperbolic.

From what we have seen above, any CAT(-1) group is word-hyperbolic. In fact, the converse is still an open question.

Question 2.19. *Does every word-hyperbolic group act properly discontinuously and cocompactly by isometries on a* CAT(-1) *space?*

This is the reason why we are slightly less interested in CAT(-1) groups: the notion of word-hyperbolicity provides an extremely natural, successful and comprehensive theory of negative curvature for groups. On the other hand, many of the techniques that we will develop for building CAT(0) spaces work just as well for building CAT(-1) spaces, and this can be a useful method of building hyperbolic groups.

Hyperbolic groups enjoy a variety of attractive properties and characterisations that are not shared by CAT(0) groups. We mention two facts as illustrations.

The first such fact is Theorem 2.13 above, which implies that every group quasi-isometric to a hyperbolic group is hyperbolic. The corresponding fact does not hold for CAT(0) groups. The standard example is $\pi_1(U\Sigma)$, where $U\Sigma$ is the unit tangent bundle over a closed negatively curved surface Σ. This group is not CAT(0), but is quasi-isometric to $\pi_1(\Sigma \times S^1)$ which, of course, is CAT(0).

The second fact is a theorem stated by Gromov and proved by Papazoglou (check!) that a group is hyperbolic if and only if its Dehn function is linear. By contrast, by the previous paragraph, $\pi_1(U\Sigma)$ has a quadratic Dehn function but is not CAT(0).

In summary, the notion of a hyperbolic group provides a comprehensive theory of negative curvature in the group-theoretic setting. The theory of CAT(0) groups is less satisfactory. Nevertheless, as we shall see it provides a rich source of interesting examples in group theory and topology. Our next task is to develop machinery that will enable us to write down examples of CAT(0) groups efficiently.

2.3 Alexandrov's Lemma

Lemma 2.20. *Suppose the triangles* $\Delta_1 \equiv \Delta(x, y, z_1)$, $\Delta_2 \equiv \Delta(x, y, z_2)$ *satisfy the* CAT(0) *condition and* $y \in [z_1, z_2]$. *Then* $\Delta \equiv \Delta(x, z_1, z_2)$ *satisfies the* CAT(0) *condition.*

Proof. The first claim is that quadrilateral $\overline{Q}$ obtained by gluing the comparison triangles $\overline{\Delta}_1$ and $\overline{\Delta}_2$ along $[\bar{x}, \bar{y}]$ has a non-acute interior angle at $\bar{y}$. Indeed, if not then there are $\bar{p}_i \in [\bar{y}, \bar{z}_i]$ with $[\bar{p}_1, \bar{p}_2] \cap [\bar{x}, \bar{y}] = \{\bar{q}\}$ where $\bar{q} \neq \bar{y}$. Since $y \in [z_1, z_2]$ we have

$$\begin{aligned} d(p_1, p_2) &= d(p_1, y) + d(y, p_2) \\ &= d(\bar{p}_1, \bar{y}) + d(\bar{y}, \bar{p}_2) \\ &> d(\bar{p}_1, \bar{q}) + d(\bar{q}, \bar{p}_2) \\ &\geq d(p_1, q) + d(q, p_2) \\ &\geq d(p_1, p_2) \end{aligned}$$

a contradiction.

The comparison triangle for Δ is obtained by straightening $[\bar{z}_1, \bar{y}] \cup [\bar{y}, \bar{z}_2]$. The remainder of the proof is a case-by-case analysis of this straightening process. We will deal with one of these cases below and leave the remainder to the reader.

Suppose $p_i \in [x, z_i]$, and let $\bar{p}_i$ be the comparison points in $\overline{Q}$. In the worst case, the Euclidean geodesic $[\bar{p}_1, \bar{p}_2]$ is not contained in $\overline{Q}$. Consider therefore the geodesic $[\bar{p}_1, \bar{y}]$. Notice that, as $[\bar{z}_1, \bar{y}] \cup [\bar{y}, \bar{z}_2]$ is straightened, $d_{\mathbb{R}^2}(\bar{p}_i, \bar{y})$ does not decrease. Therefore, we always have $d(p_i, y) \leq d_{\mathbb{R}^2}(\bar{p}_i, \bar{y})$. Eventually, $[\bar{p}_1, \bar{p}_2]$ lies in the interior of $\overline{Q}$, and so we can reduce to that case. Let $\{\bar{q}\} = [\bar{p}_1, \bar{p}_2] \cap [\bar{x}, \bar{y}]$. Then

$$d(p_1, p_2) \leq d(p_1, q) + d(q, p_2) \leq d_{\mathbb{R}^2}(\bar{p}_1, \bar{q}) + d_{\mathbb{R}^2}(\bar{q}, \bar{p}_2) = d_{\mathbb{R}^2}(\bar{p}_1, \bar{p}_2)\ .$$

As the last quantity only increases during the straightening process, the lemma is proved. □

This enables us to perform gluing constructions. The following proposition is an easy application of the lemma. For the definition of the induced length metric, see the next section.

Proposition 2.21. *Suppose* X_1, X_2 *are locally compact, complete* CAT(0) *spaces and* Y *is isometric to closed, convex subspaces of both* X_1 *and* X_2 *then* $X_1 \cup_Y X_2$, *equipped with the induced length metric, is* CAT(0).

In particular, we have:

Corollary 2.22. *If* Γ_1, Γ_2 *are both* CAT(0) *groups then so is* $\Gamma_1 * \Gamma_2$.

2.4 Length metrics and the Cartan–Hadamard Theorem

In the motivating case of Riemannian geometry, it is easy to see how a (Riemannian) metric on a manifold X induces a metric on its universal cover. In this subsection, we will introduce length metrics, which play an analogous role. It is convenient to allow the possibility that two points are at infinite distance.

Definition 2.23. Let (X, d) be a metric space and let $\gamma : [a, b] \to X$ be a path. The *length* of γ is the quantity

$$l(\gamma) = \sup_{a=t_0<t_1<\ldots<t_k=b} \sum_{i=1}^{k} d(\gamma(t_{i-1}), \gamma(t_i))$$

where the supremum ranges over all finite partitions $t_0 < t_1 < \ldots < t_k$ of $[a, b]$. The path γ is called *rectifiable* if its length is finite.

Definition 2.24. A metric space is called a *length space* if the distance between any pair of points is equal to the infimum of the lengths of the paths between them.

There is a length pseudometric associated to any metric space, defined in the obvious way:

$$\hat{d}(x, y) = \inf_{\gamma(0)=x,\ \gamma(1)=y} l(\gamma)$$

where γ ranges over all paths connecting x_0 to x_1. Of course, any geodesic metric space is automatically a length space. In the context of length spaces, there is a natural metric on any covering space.

Definition 2.25. Let $p : \widetilde{X} \to X$ be a covering map and X a length space. Then there is a unique length metric defined on $\widetilde{X}$ that makes p into a local isometry, defined as follows. Define the length of a path $\tilde{\gamma}$ in $\widetilde{X}$ to be the length of $p \circ \tilde{\gamma}$. The corresponding length metric on $\widetilde{X}$ is called the *induced metric on* $\widetilde{X}$. In particular, if X is complete then so is $\widetilde{X}$.

We leave it as an exercise to prove that this really is a length metric on $\widetilde{X}$, that the covering map becomes a local isometry, and that this is the unique metric on $\widetilde{X}$ with these properties.

Theorem 2.26 (Hopf–Rinow)**.** *Let X be a length space. If X is complete and locally compact then X is proper and geodesic.*

Proof. To prove properness, we need to prove that closed balls $\overline{B}_{x_0}(r)$ are compact. Consider the set of all r for which this is true. Using local compactness, it is easy to argue that this set is open, so it remains to check that it is closed. Let r be a limit point of this set; we need to show that $\overline{B}_{x_0}(r)$ is (sequentially) compact. Let (x_n) be a sequence of points in $\overline{B}_{x_0}(r)$. Choose a path γ_n from x_0 to x_n, and for each $p \in \mathbb{N}$ let y_n^p be a suitable point on the path, with $d(x_0, y_n^p) < r - 1/2p$ and $d(x_n, y_n^p) < 1/p$. Each $y_n^p \in \overline{B}_{x_0}(r - 1/2p)$ which is compact by hypothesis, so (y_n^p) has a convergent subsequence for fixed p. Using a diagonal argument, there is a sequence (n_k) so that $(y_{n_k}^p)$ is convergent, hence Cauchy, for all p. It is now easy to argue that x_{n_k} is Cauchy, and hence convergent.

The fact that X is geodesic is now a straightforward application of the Arzelà–Ascoli Theorem and the easily checked fact that length of paths is lower semicontinuous. Let $\gamma_n : [0,1] \to X$ be a sequence of paths from p to q whose lengths converge to $d(p,q)$. They can all be taken to have length at most $R = d(p,q) + 1$ and to be parametrised by arc length. Then for any $s, t \in [0,1]$, we have

$$d(\gamma_n(s), \gamma_n(t)) < \frac{|s-t|}{R+1}$$

and so the γ_n are uniformly equicontinuous. Therefore some subsequence converges to a path γ from p to q. By lower semicontinuity of length, γ is a geodesic. □

Theorem 2.27 (The Cartan–Hadamard Theorem)**.** *If X is a complete, connected length space of non-positive curvature then the universal cover $\widetilde{X}$, with the induced length metric, is* CAT(0)*.*

This theorem is the first step in our goal to find ways to write down many examples of CAT(0) groups. In particular, the following is immediate.

Corollary 2.28. *A group Γ is* CAT(0) *if and only if Γ is the fundamental group of a compact non-positively curved space.*

We will prove the Cartan–Hadamard Theorem under the additional hypothesis of local compactness, following Ballmann [1]. The theorem is proved in two parts. First, we prove it under the assumption that geodesics in $\widetilde{X}$ are unique.

Lemma 2.29. *If X is proper and non-positively curved, and there is a unique geodesic between each pair of points in X, then X is* CAT(0)*.*

Proof. By Lemma 2.5, the hypothesis that geodesics are unique implies that they vary continuously with their endpoints.

Consider a triangle $\Delta \equiv \Delta(x,y,z)$, contained in a ball $\overline{B} = \overline{B}_x(R)$. Because $\overline{B}$ is compact, there is an $\epsilon > 0$ such that $B_p(\epsilon)$ is CAT(0) for every $p \in \overline{B}$. Let $\alpha = [y,z]$ and, for each t, let γ_t be the geodesic from x to $\alpha(t)$. Because geodesics vary continuously with their endpoints, there is a δ such that $d(\alpha_{t_1}(s), \alpha_{t_2}(s)) < \epsilon$, for all s, whenever $|t_1 - t_2| < \delta$.

To prove the lemma, we now divide Δ into a patchwork of geodesic triangles, each contained in a ball of radius ϵ. By Alexandrov's Lemma, it follows by induction that Δ satisfies the CAT(0) condition. □

The Cartan–Hadamard Theorem is an immediate consequence of this lemma and the next result.

Theorem 2.30. *Let X be a proper length space of non-positive curvature. Each homotopy class of paths from p to q contains exactly one local geodesic.*

The proof of this theorem uses a process called *Birkhoff curve shortening.* Consider a path $\gamma : [0,1] \to X$ from p to q. As before, let γ be contained in a compact ball $\overline{B} = \overline{B}_{x_0}(R)$ and let $\epsilon > 0$ be such that $B_y(2\epsilon)$ is CAT(0) for every $y \in \overline{B}$. Let N be such that $\gamma([\frac{i-1}{2N}, \frac{i+1}{2N}]) \subseteq B_{\gamma(i/2N)}(\frac{\epsilon}{2})$ for each integer i.

The Birkhoff curve-shortening map is defined as follows. For each $s \in [0,1]$, let $\beta^0_s(\gamma)$ be the curve obtained by replacing γ by the unique geodesic on each interval $[\frac{i}{N}, \frac{i+s}{N}]$. Likewise, let $\beta^1_s(\gamma)$ be the curve obtained by replacing γ by the unique geodesic on each interval $[\frac{2i+1}{2N}, \frac{2i+2s+1}{2N}]$. Finally, define $\beta(\gamma) = \beta^1_1 \circ \beta^0_1(\gamma)$. Note that this is a continuous function of γ.

Remark 2.31. For a path γ:

1. γ is homotopic to $\beta(\gamma)$ (respecting endpoints);
2. $\gamma = \beta(\gamma)$ if and only if γ is a local geodesic.
3. Set
$$\lambda(\gamma) = \sum_{i=1}^{2n} d\left(\gamma\left(\frac{i-1}{2n}\right), \gamma\left(\frac{i-1}{2n}\right)\right)^2,$$
which is a continuous function of γ. Note that $\lambda(\beta(\gamma)) \leq \lambda(\gamma)$, with equality if and only if γ is a local geodesic, parametrised by arc length.

Put the supremum metric on the space of continuous paths between p and q.

Lemma 2.32. *Let γ_1 and γ_2 be continuous paths from p to q, and suppose that $d_{\sup}(\gamma_1, \gamma_2) < \epsilon$. Then*
$$d_{\sup}(\beta(\gamma_1), \beta(\gamma_2)) \leq d_{\sup}(\gamma_1, \gamma_2)\ .$$

Proof. Since each component of the straightening happens inside the same ball of radius 2ϵ, the lemma follows from convexity of the metric. □

Proof of Theorem 2.30. First we prove existence. Let γ be as above. Because the distance between any pair of nearby points decreases when β is applied, it follows that $\{\beta^n(\gamma)\}$ is equicontinuous, and so there is a convergent subsequence $\beta^{n_k}(\gamma) \to \gamma_\infty$ by the Arzelà–Ascoli Theorem, where γ_∞ is a curve from p to q in the homotopy class of γ. The sequence $\lambda(\beta^n(\gamma))$ is decreasing, and so
$$\lim_{n\to\infty} \lambda(\beta^n(\gamma)) = \lim_{k\to\infty} \lambda(\beta^{n_k}(\gamma)) = \lambda(\gamma_\infty)\ .$$
Finally, $\lambda(\beta(\gamma_\infty)) = \lim_{n\to\infty} \lambda(\beta^{n+1}(\gamma_\infty)) = \lambda(\gamma_\infty)$, so γ_∞ is a geodesic as required.

We now prove uniqueness. Let $\gamma_0, \gamma_1 : [0,1] \to X$ be local geodesics from p to q, and let γ_s be a homotopy between them. By compactness of the unit square, we can choose R, ϵ and N as above that are suitable for all γ_s. Applying β^n, we obtain a sequence of homotopies $\beta^n(\gamma_s)$ from γ_0 to γ_1. From Lemma 2.32, it follows that $\{(s,t) \mapsto \beta^n(\gamma_s)(t)\}$ is equicontinuous, and so there is a subsequence $\beta^{n_k}(\gamma_s)$ that converges to a limiting homotopy $\hat{\gamma}_s$. By the existence argument $\hat{\gamma}_s$ is a geodesic for each s. We have proved that γ_0 and γ_1 are homotopic through geodesics.

Whenever s_1 and s_2 are sufficiently close together, we have that $d_{\sup}(\hat{\gamma}_{s_1}, \hat{\gamma}_{s_2}) < \epsilon$ and so the function $t \mapsto d(\hat{\gamma}_{s_1}(t), \hat{\gamma}_{s_2}(t))$ is locally convex. But a locally convex function on the reals is globally convex, and so $\hat{\gamma}_{s_1} = \hat{\gamma}_{s_2}$. Therefore $\gamma_0 = \gamma_1$, as required. ☐

2.5 Gromov's Link Condition

By the Cartan–Hadamard Theorem, to exhibit a CAT(0) group it is enough to exhibit a compact metric space with a non-positively curved metric. In this section we will work with Euclidean complexes—that is, CW-complexes in which every cell is isometric to a convex polyhedron embedded in Euclidean space, and in which the attaching maps identify faces with cells isometrically. After subdivision, we may assume that the cells are isometric to Euclidean simplices and that the attaching maps are embeddings. We will also always assume that our complexes are locally finite. Such a complex is naturally endowed with a length metric, which by the Hopf–Rinow Theorem is geodesic. Gromov's Link Condition provides a criterion that ensures that the length metric on such a complex is non-positively curved.

Definition 2.33. Let X be a geodesic space. The *link* of a point $x_0 \in X$, denoted by $\mathrm{Lk}(x_0)$, is the space of unit-speed geodesics $\gamma : [0, a] \to X$ with $\gamma(0) = x_0$ (equipped with the compact-open topology), modulo the equivalence relation that $\gamma_1 \sim \gamma_2$ if and only if γ_1 and γ_2 coincide on some interval $[0, \epsilon)$ for $\epsilon > 0$.

The link of x_0 can be metrised using angle. There is a general definition of angle (due to Alexandrov) that makes sense in metric spaces, but in the case of Euclidean complexes it is easy to define angle on an *ad hoc* basis. For any open simplex Σ with x_0 in its closure, let $\mathrm{Lk}_\Sigma(x_0)$ be the subset of the link that consists of geodesics that point into Σ. Angle in Σ defines a metric on $\mathrm{Lk}_\Sigma(x_0)$, and these metrics can be glued together to give $\mathrm{Lk}(x_0)$ the metric structure of a (spherical) complex. We denote the metric on $\mathrm{Lk}(x_0)$ by $\angle_{x_0}$.

Theorem 2.34 (Gromov's Link Condition)**.** *Any Euclidean complex X is non-positively curved if and only if* $\mathrm{Lk}(x_0)$ *is* CAT(1) *for every $x_0 \in X$.*

In fact, the same theorem holds for complexes with cells of curvature κ, and the proof we will give also works in these cases. We loosely follow Ballmann [1].

Definition 2.35. Let L be a metric space. A *(Euclidean) cone on L*, denoted by $C_\epsilon L$, is the metric space associated to the pseudometric space $L \times [0, \epsilon]$, where we define the pseudometric by

$$d((x,s),(y,t))^2 = s^2 + t^2 - 2st\cos\min\{\pi, d(x,y)\}$$

and note that the two points are at distance zero if and only if $s = t = 0$.

Now the Link Condition follows from the following lemmas.

Lemma 2.36. *Consider two points $x, y \in L$ at distance less than π. For any $s, t > 0$, there is a bijection between the set of geodesic segments joining x to y in L and the set of geodesics joining (x,s) to (y,t) in $C_\epsilon L$.*

Proof. Consider a geodesic $[x,y]$ in L. Together with the cone point, it spans a subcone $C_\epsilon[x,y]$ which is isometric to a Euclidean cone. The unique Euclidean geodesic from (x,s) to (y,t) is then the geodesic corresponding to $[x,y]$.

Conversely, consider a geodesic $[(x,s),(y,t)]$ in $C_\epsilon L$. If the cone point is in the geodesic then the distance between them is $s+t$, which implies that $d(x,y) \geq \pi$. Otherwise, there is a well defined projection of the geodesic to L. Let (z,r) be any point on $[(x,s),(y,t)]$. It is enough to prove that $d(x,y) \geq d(x,z) + d(z,y)$; the converse inequality is obvious, and it then follows that the image is a geodesic. To see this, simply note that the length of the projection of a geodesic is equal to the angle in the corresponding comparison triangle. But the comparison triangle for $[(x,s),(y,t)]$ can be obtained by straightening the comparison triangles for $[(x,s),(z,r)]$ and $[(y,t),(z,r)]$, and when we do so the angle at the cone point increases. □

Note that if $d(x,y) \geq \pi$ then the path through the cone point is a geodesic. In particular, the cone is a geodesic space.

Lemma 2.37. *If $x_0 \in X$ then, for some $\epsilon > 0$, $\overline{B}_{x_0}(\epsilon)$ is convex and isometric to $C_\epsilon \mathrm{Lk}(x_0)$.*

Proof. Let $\{\Sigma_i\}$ be the set of open simplices whose closures contain x_0. Their union U is an open neighbourhood of x_0. For each $y \in U$ there is a well defined geodesic $[x_0, y]$ and so a continuous map $\pi : U \smallsetminus \{x_0\} \to \mathrm{Lk}(x_0)$.

Let $\gamma : [0,1] \to U \smallsetminus \{x_0\}$ be any local geodesic, and suppose that $\pi \circ \gamma$ is non-constant. Consider the triangle $\Delta \equiv \Delta(\gamma(0), x_0, \gamma(1))$. Let $0 = t_0 < t_1 < \ldots < t_n = 1$ be so that $\gamma|[t_k, t_{k+1}]$ is equal to a component of the intersection of the image of γ with the interior of Δ_{i_k}. Then $\Delta(\gamma(t_{k-1}), x_0, \gamma(t_k))$ is isometric to a Euclidean triangle, which we shall denote by $\overline{\Delta}_k$. Arranging the $\overline{\Delta}_k$ side by side in $\mathbb{R}^2$, we obtain a comparison triangle $\overline{\Delta}$ for Δ.

There is a distance-non-increasing continuous map $\overline{\Delta} \to \Delta$. Let 2ϵ be smaller than the minimal distance from x_0 to a simplex not contained in U. Then every point of every geodesic in $\overline{B}_{x_0}(\epsilon)$ is contained in U, and so $\overline{B}_{x_0}(\epsilon)$ is convex.

In particular, the induced metric on $\overline{B}_{x_0}(\epsilon)$ is a length metric. By contruction it agrees with the cone metric on the interior of each cell. As both metrics are length metrics, it follows that they agree. □

Lemma 2.38. *The cone $C_\epsilon L$ is* CAT(0) *if and only if L is* CAT(1).

Proof. Let $\Delta = \Delta(x, y, z)$ be a triangle in $C_\epsilon L$. Note that it suffices to prove the inequality in the case in which one of the comparison points is a vertex.

There are three cases to consider. In the first, x_0 is on the boundary of Δ. Cutting Δ into pieces, we may assume that $x_0 = x$. As in the proof of convexity above, the map $\overline{\Delta} \to \Delta$ does not increase distance and so we are done.

We may therefore assume that x_0 is not contained in an edge of Δ. Let $p = x$ and $q \in [y, z]$. In the second case, $\pi(\Delta)$ has perimeter at least 2π. Consider the comparison triangles $\overline{\Delta}(x_0, x, y)$, $\overline{\Delta}(x_0, y, z)$, $\overline{\Delta}(x_0, z, x)$. Glue them along $[\bar{x}_0, \bar{y}]$ and $[\bar{x}_0, \bar{z}]$ and let $\bar{x}_1$, $\bar{x}_2$ be the unglued vertices of $\overline{\Delta}(x_0, x, y)$ and $\overline{\Delta}(x_0, z, x)$. Let $\bar{x}$ be the point of intersection of the circle based at $\bar{y}$ of radius $d(x, y)$ and the circle based at $\bar{z}$ of radius $d(x, z)$ on the same side of $[\bar{y}, \bar{z}]$ as the triangles. The comparison triangle for Δ is constructed by moving $\bar{x}_1$ to $\bar{x}$ along the first circle and $\bar{x}_2$ to $\bar{x}$ along the second circle. For any $\bar{r} \in [\bar{x}_0, \bar{y}]$, we have $d(\bar{x}, \bar{r}) \geq d(\bar{x}_1, \bar{r})$, and similarly for $\bar{r} \in [\bar{x}_0, \bar{z}]$. After straightening,

$$d(\bar{x}, \bar{q}) = d(\bar{x}, \bar{r}) + d(\bar{r}, \bar{q})$$

for some such $\bar{r}$, and the second term has remained constant during the straightening process, so $d(x, q) \leq d(\bar{x}, \bar{q})$ as required.

In the remaining case, the perimeter of $\pi(\Delta)$ is less than 2π, and so it satisfies the CAT(1) condition. Consider a geodesic $\gamma = [x, q]$. Then $\pi \circ \gamma$ is a geodesic in $\mathrm{Lk}(x_0)$, which we can develop as follows. Let $0 = t_0 < t_1 < \ldots < t_n = 1$ be so that $\pi \circ \gamma | [t_k, t_{k+1}]$ is equal to a component of the intersection of the image of γ with the interior of $\Delta_{i_k} \cap \mathrm{Lk}(x_0)$. Then $\pi \circ \gamma(t_{k-1}), \gamma(t_k))$ defines a set of directions, which are equal to the intersection of Δ_{i_k} with a 2-dimensional Euclidean plane. Let $\widetilde{\Delta}_k$ denote the Euclidean segment spanned by this set of directions. Gluing the $\widetilde{\Delta}_k$ together, we obtain a Euclidean segment $\widetilde{\Delta}$, and we may consider points $\tilde{x}_0$ at the cone point of the segment, $\tilde{x}$ on one edge with $d(\tilde{x}, \tilde{x}_0) = d(x, x_0)$ and $\tilde{q}$ on the other edge with $d(\tilde{q}, \tilde{x}_0) = d(q, x_0)$. To finish the proof, we compare $\widetilde{\Delta}$ with the comparison triangle $\overline{\Delta}$.

The CAT(1) hypothesis applied to $\pi(\Delta)$ tells us that

$$\angle_{\tilde{x}_0}(\tilde{x}, \tilde{q}) \leq \angle_{\bar{x}_0}(\bar{x}, \bar{q}) \ .$$

Therefore

$$\begin{aligned} d(x,q)^2 &= d(\tilde{x}, \tilde{q})^2 \\ &= d(\tilde{x}, \tilde{x}_0)^2 + d(\tilde{q}, \tilde{x}_0)^2 - 2d(\tilde{x}, \tilde{x}_0) d(\tilde{q}, \tilde{x}_0) \cos \angle_{\tilde{x}_0}(\tilde{x}, \tilde{q}) \\ &\leq d(\bar{x}, \bar{x}_0)^2 + d(\bar{q}, \bar{x}_0)^2 - 2d(\bar{x}, \bar{x}_0) d(\bar{q}, \bar{x}_0) \cos \angle_{\bar{x}_0}(\bar{x}, \bar{q}) \\ &= d(\bar{x}, \bar{q})^2 \end{aligned}$$

as required.

For the converse assertion, note that if $\mathrm{Lk}(x_0)$ fails to be CAT(1) then the inequality in the final calculation fails, and so the oringal triangle Δ did not satisfy the CAT(0) condition. □

3 Examples of CAT(0) groups

Armed with the Link Condition, we can easily write down non-positively curved complexes, whose fundamental groups are then CAT(0) by the Cartan–Hadamard Theorem.

3.1 Wise's example

In dimension two, the Link Condition is very easy to check.

Corollary 3.1. *If X is a 2-dimensional Euclidean complex then X is* CAT(0) *if and only if, for every vertex of X, every loop in* $\mathrm{Lk}(x_0)$ *is of length at least* 2π*.*

Example 3.2. Consider the group

$$W \cong \langle a, b, s, t \mid [a, b] = 1,\ a^s = (ab)^2, b^t = (ab)^2 \rangle \ .$$

First, we will show that W is CAT(0). Consider the following combinatorial Euclidean 2-complex X, in which the edges labelled by white triangles have length one and every other edge has length two.

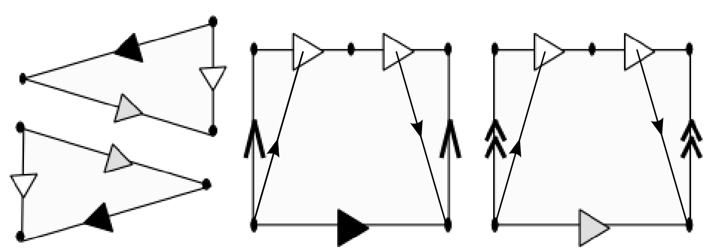

Figure 1: A complex X with fundamental group W. A locally geodesic representative of $sabs^{-1}$ is marked on the left square, and a locally geodesic representative of $tabt^{-1}$ is marked on the right square.

It is straightforward to check that its fundamental group is W: there is one vertex, the black arrow represents a, the grey arrow b, the white arrow ab, the single arrow s and the double arrow t. The link of the unique vertex is illustrated below. One can easily check that the shortest loops are of length exactly 2π.

We will show that W is *non-Hopfian*; that is, there is a surjection $f : W \to W$ which is not an isomorphism. Define $f : W \to W$ by $a \mapsto a^2$, $b \mapsto b^2$, $s \mapsto s$, $t \mapsto t$; it is easy to check that the images of the generators satisfy the relations. Note also that every generator is in the image of f, so f is a surjection. Consider the element

$$g = [(ab)^{s^{-1}}, (ab)^{t^{-1}}] = (sabs^{-1})^{-1}(tabt^{-1})^{-1}(sabs^{-1})(tabt^{-1}) \ .$$

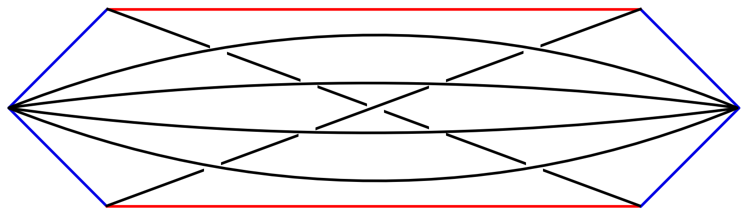

Figure 2: The link of the vertex in the above example. Each black arc is a pair of edges each of length $\pi/2$, each blue arc is an edge of length $\arccos 1/4$ and each red arc is a pair of edges each of length $\arcsin 1/4$.

One can easily check that $f(g)$ is trivial: $f(g) = [a, b] = 1$.

In general, it can be difficult to show that any given element is non-trivial. We will show that $g \neq 1$ using some geometry. The loop g can be represented by a loop γ that, each time it crosses an edge or the vertex, traverses an angle of at least π; γ can be constructed explicitly in the obvious way from the geodesics indicated in Figure 1. It follows that γ is a local geodesic $S^1 \to X$—recall that every geodesic that misses the cone point projects into the link with an angle of less than π. Therefore γ lifts to a local geodesic $\tilde{\gamma}$ in the universal cover $\widetilde{X}$, which is a CAT(0) space. But each based homotopy class in $\widetilde{X}$ contains a unique local geodesic, so the image of $\tilde{\gamma}$ is isometric to a line on which $\langle g \rangle$ acts non-trivially, which means that that g is of infinite order, in particular non-trivial.

That this group is non-Hopfian is interesting because of a pair of theorems of Malt'sev.

Theorem 3.3. *(Malt'sev) Every finitely generated linear group is residually finite.*

Proof for $GL_n(\mathbb{Z})$. Think of $A \in GL_n(\mathbb{Z}) \smallsetminus \{I\}$ as a matrix, and let p be a prime that does not divide any of the entries of $A - I$. Then the reduction homomorphism $GL_n(\mathbb{Z}) \to GL_n(\mathbb{Z}/p)$ does not kill A, as required. □

Theorem 3.4. *(Malt'sev) Every finitely generated residually finite group is Hopfian.*

Consider also the following deep theorem of Sela.

Theorem 3.5. *(Sela) Every hyperbolic group is Hopfian.*

There are known examples of non-linear hyperbolic groups, but it is a major open question whether every hyperbolic group is residually finite.

3.2 Injectivity radius and systole

Before we construct examples in higher dimensions, we need to establish some conditions under which a space of curvature at most κ may fail to be CAT(κ).

Definition 3.6. Let X be a geodesic metric space. The *injectivity radius* of X is the smallest $r \geq 0$ such that there are distinct geodesics in X with common endpoints of length r. The *systole* of X is the length of the smallest isometrically embedded circle in X.

Note that the systole of is clearly at least twice the injectivity radius.

The results of this subsection are needed in the situation of general curvature at most κ. Many of the above results go through in this setting. The major missing fact is that, for $\kappa > 0$, the metric is not convex. But we do have uniqueness of short geodesics.

Proposition 3.7. *Let X be a compact geodesic metric space of curvature at most κ. Then X is not* $\mathrm{CAT}(\kappa)$ *if and only if it contains an isometrically embedded circle of length less than $2D_\kappa$. If it does, then it contains a circle of length equal to twice the injectivity radius of X; in particular, twice the injectivity radius is equal to the systole.*

Proof. Clearly, if X contains an isometrically embedded circle of length less than $2D_\kappa$ then X is not $\mathrm{CAT}(\kappa)$. Conversely, suppose that X is $\mathrm{CAT}(\kappa)$. Let r be the injectivity radius. Note that, by the argument using Alexandrov's Patchwork, any triangle of perimeter less than $2r$ satisfies the $\mathrm{CAT}(\kappa)$ condition. The key idea of the following proof is to divide any small digons into pairs of triangles, which must satisfy the $\mathrm{CAT}(\kappa)$ condition, and then to apply Alexandrov's Lemma (backwards!).

Let $[x_n, y_n]$, $[x_n, y_n]'$ be a sequence of pairs of distinct geodesics with common endpoints, whose lengths tend to r. By compactness and Arzela–Ascoli, we may pass to a subsequence and assume that these converge to a pair of geodesics $[x, y]$, $[x, y]'$ with $d(x, y) = r$. First, we shall show that these two geodesics are distinct. Suppose not. Then for all suitably large n, the midpoints m_n and m'_n of $[x_n, y_n]$ and $[x_n, y_n]'$ respectively are close together, so we may take them to be close enough that the triangles $\Delta(x_n, m_n, m'_n)$ and $\Delta(y_n, m_n, m'_n)$ are both of perimeter less than $2r$, and therefore satisfy the $\mathrm{CAT}(\kappa)$ condition. The comparison triangle $\overline{\Delta}(x_n, m_n, y_n)$ is degenerate. By Alexandrov's Lemma it follows that the comparison triangles $\overline{\Delta}(x_n, m_n, m'_n)$ and $\overline{\Delta}(y_n, m_n, m'_n)$ are also degenerate, so $m_n = m'_n$.

Finally, we want to show that $[x, y] \cup [x, y]'$ is an isometrically embedded circle. To do this, note that we could have applied the same argument as in the previous paragraph to any pair of points z, z' that are at distance r in the digon but distance less than r in X. The result follows. □

3.3 Cube complexes

Definition 3.8. A Euclidean complex is a *cube complex* if every cell is isometric to a cube.

The link of a vertex of a cube is naturally a simplex, so in this case the links are simplicial complexes. Moreover, they are naturally *all-right spherical* simplicial complexes, meaning that every edge has length $\pi/2$.

Definition 3.9. A simplicial complex L is *flag* if, for every $k \geq 2$, whenever $K \subseteq L^{(1)}$ is a subcomplex of the 1-skeleton that is isomorphic to the 1-skeleton of an n-simplex, there is an n-simplex Σ in L with $\Sigma^{(1)} = K$.

Theorem 3.10 (Gromov)**.** *An all-right spherical simplicial complex L is* CAT(1) *if and only if it is flag.*

Therefore, the Link Condition for cube complexes is purely combinatorial. The barycentric subdivision of any simplicial complex is flag so, in particular, there is no topological obstruction to being CAT(1).

Proof. The link of a vertex in an all-right spherical complex is an all-right spherical complex. Furthermore, by the Link Condition (whose proof goes through in the curvature-one case), L is locally CAT(1) if and only if the link of every vertex is CAT(1). These two facts suggest a proof by induction.

First, suppose that L is CAT(1). Links are also CAT(1) and therefore, by induction, can be taken to be flag. Suppose now that $K \subseteq L^{(1)}$ is isomorphic to the $(n-1)$-skeleton of an n-simplex and let v be a vertex of L contained in K. The link $\mathrm{Lk}(v)$ is flag, and $\mathrm{Lk}(v) \cap K$ is an $(n-2)$-simplex, which bounds an $(n-1)$-simplex in $\mathrm{Lk}(v)$. Therefore, K bounds an n-simplex in L.

For the converse, suppose that L is flag. Links of vertices are also flag and so, by induction, are CAT(1). Therefore L is of curvature at most one by the Link Condition. By Proposition 3.7, it remains to show that L has no isometrically embedded, locally geodesic circle of length less than 2π. Suppose therefore that γ is such a locally geodesic circle.

Suppose that $x \in L$ and that γ intersects $B_x(\pi/2)$. As before, fix $\bar{x}$ in S^2 and let $\bar{\gamma}$ be the development of γ into S^2. Then $\bar{\gamma} \cap B_{\bar{x}}(\pi/2)$ has length π, and it follows that this is also the length of the intersection of γ with $B_{x_0}(\pi/2)$.

Let u, v be vertices of L such that γ intersects $B_u(\pi/2)$ and $B_v(\pi/2)$. Because γ is of length less than 2π, it follows from the previous paragraph that some point of γ is contained $B_u(\pi/2) \cap B_v(\pi/2)$, and so u and v are distance less than π apart. Therefore $d(u,v) = \pi/2$. So the set of vertices of every open simplex that γ touches span the 1-skeleton of a simplex and hence span a simplex, because L is flag. So γ is contained in a simplex, which is absurd. □

3.4 Right-angled Artin groups

Definition 3.11. Let N be a simplicial graph. The corresponding *right-angled Artin group*, *graph group*, or *free partially commutative group* is defined by

$$A_N \cong \langle V(N) \mid \{[u,v] \mid (u,v) \in E(N)]\} \rangle \ .$$

Let Σ be the unique flag complex with $\Sigma^{(1)} = N$. It is often convenient to denote A_N by A_Σ instead.

Right-angled Artin groups are the fundamental groups of non-positively curved cube complexes, called *Salvetti complexes*, defined as follows.

Definition 3.12. Consider the cube $[0,1]^{V(\Sigma)}$ and identify Σ with a subcomplex of Lk(0) in the natural way. Let $\pi : [0,1]^{V(\Sigma)} \smallsetminus \{0\} \to \mathrm{Lk}(0)$ be radial projection and let $q : [0,1]^{V(\Sigma)} \to (S^1)^{V(\Sigma)}$ be the natural quotient map obtained by identifying opposite sides. The *Salvetti complex* is defined to be

$$\mathcal{S}_\Sigma = q(\pi^{-1}(\Sigma) \cup \{0\}) \ .$$

It is a union of coordinate tori in $(S^1)^{V(\Sigma)}$, and in particular has the structure of a cube complex. Note that $\mathcal{S}_\Sigma$ one open $(n+1)$-cell for each open n-cell of Σ.

By construction, $\pi_1(\mathcal{S}_\Sigma) \cong A_\Sigma$. To describe the link of the vertex of $\mathcal{S}_\Sigma$, we need another definition.

Definition 3.13. The *double* $D(\Sigma)$ of a simplicial complex L is defined as follows. The 0-skeleton of $D(\Sigma)$ consist of two copies of the 0-skeleton of Σ—for $v \in L^{(\Sigma)}$, the corresponding vertices of $D(\Sigma)$ are denoted v^+ and v^-. A set of vertices $\{v_0^\pm, \ldots, v_n^\pm\}$ of $D(\Sigma)$ spans an n-simplex if and only if $\{v_0, \ldots, v_n\}$ are all distinct and span a simplex in Σ. Note that $D(\Sigma)$ contains two distinguished copies of Σ, namely the full subcomplex spanned by the vertices $\{v^+\}$ and the full subcomplex spanned by the vertices $\{v^-\}$. These are denoted Σ^+ and Σ^-, respectively.

Remark 3.14. It is immediate that $D(\Sigma)$ is flag if and only if Σ is.

Remark 3.15. The subcomplexes Σ^+ and Σ^- are retracts of $D(\Sigma)$.

Lemma 3.16. *The link of the unique vertex x_0 of $\mathcal{S}_\Sigma$ is isomorphic to $D(\Sigma)$.*

Proof. For $v \in \Sigma^{(0)}$, by construction there are precisely two corresponding vertices in $\mathrm{Lk}(x_0)$, which we denote $v^\pm$ according to the orientation of the 1-cell labelled by v. A set of vertices $\{v_0, \ldots, v_n\}$ span a simplex in L if and only if the face of $[0,1]^{\Sigma^{(0)}}$ spanned by the corresponding directions is a cube in $q^{-1}(\mathcal{S}_\Sigma$. Such a cube contributes 2^{n+1} n-cells to $\mathrm{Lk}(x_0)$, namely precisely the simplices spanned by the various combinations $\{v_0^\pm, \ldots, v_n^\pm\}$. □

As the double of a flag complex is flag, it follows by Theorem 3.10 that $\mathcal{S}_\Sigma$ is non-positively curved, and hence A_Σ is CAT(0).

Proposition 3.17. *The Salvetti complex $\mathcal{S}_\Sigma$ is non-positively curved.*

Remark 3.18. Suppose $\Sigma' \subseteq \Sigma$ is a full subcomplex. Then $\mathcal{S}_{\Sigma'}$ is naturally a locally convex subcomplex of $\mathcal{S}_\Sigma$.

Despite their simple definitions, right-angled Artin groups have a remarkably rich subgroup structure, as we shall see.

4 Bestvina–Brady groups

In this section we will see one way in which the CAT(0) geometry of right-angled Artin groups has been used to provide remarkable examples in topology.

4.1 Finiteness properties for groups

Recall how to compute the homology of a group Γ. We construct a presentation complex $X^{(2)}$ for Γ, and then successively add higher dimensional cells to kill higher homotopy groups. The result is an aspherical complex X, and the homology of Γ is the homology of X. Using cellular homology,, it is immediately clear that if Γ is finitely presented then $H_2(\Gamma)$ is of finite rank. This observation raises the subtle topological problem of whether the converse is true.

Question 4.1. *Is there a finitely generated group Γ that is not finitely presentable but such that $H_2(\Gamma)$ is of finite rank?*

To refine this question somewhat, we recall how group homology is calculated from X. Let $\widehat{X}$ be the universal cover of X, which is contractible. There is naturally associated to X the chain complex

$$\cdots \to C_i(\widetilde{X}) \to \cdots \to C_1(\widetilde{X}) \to C_0(\widetilde{X}) \to \mathbb{Z} \to 0 \ ;$$

because Γ acts on $\widetilde{X}$, the terms are naturally $\mathbb{Z}\Gamma$-modules. The sequence is exact, and this defines a free (in particular, projective) resolution of $\mathbb{Z}$, thought of as a trivial $\mathbb{Z}\Gamma$-module. To compute the homology of Γ, we take any projective resolution

$$\cdots \to P_i \to \cdots \to P_1 \to P_0 \to \mathbb{Z} \to 0 \ ,$$

tensor the terms with $\mathbb{Z}$ and then take the homology of the resulting exact sequence.

Finiteness properties of groups measure how much of this procedure can be carried out using finite rank. A group Γ is *of type F_n* if it has an Eilenberg–Mac Lane space with finite n-skeleton; in particular, F_1 is the same as being finitely generated and F_2 is the same as being finitely presentable.

Likewise, a group Γ is said to be of type FP_n if the $\mathbb{Z}\Gamma$-module $\mathbb{Z}$ has a projective resolution which is finitely generated in dimensions up to n. Note that F_n implies FP_n. In turn, FP_n implies that the $H_i(\Gamma)$ is finite dimensional for $i \leq n$. This is often how one tells when a group is not of type FP_n.

The following was a very longstanding open question in topology and group theory.

Question 4.2. *Is there a finitely generated group Γ that is not finitely presentable but is of type FP_2?*

This question belongs to a circle of notoriously difficult problems that concern the topology of 2-dimensional CW-complexes; other examples include the Whitehead Conjecture and the Eilenberg–Ganea Conjecture.

Conjecture 4.3 (Whitehead)**.** *Every connected subcomplex of a 2-dimensional aspherical CW-complex is aspherical.*

Conjecture 4.4 (Eilenberg–Ganea)**.** *If a group has cohomological dimension 2, then it has a 2-dimensional Eilenberg–Mac Lane space.*

4.2 The Bestvina–Brady Theorem

We say that a space X is *n-connected* if and only if $\pi_i(X)$ is trivial for all $i \leq n$, and *n-acyclic* if and only if $H_i(X, \mathbb{Z})$ is isomorphic to H_i of a point for $i \leq n$.

Theorem 4.5 (Bestvina–Brady [2])**.** *Let Σ be a flag complex. Consider the homomorphism $h : A_\Sigma \to \mathbb{Z}$ that sends each generator to 1, and let $H_\Sigma = \ker h$:*

1. *H_Σ is of type FP_{n+1} if and only if Σ is n-acylic;*
2. *H_Σ is finitely presented if and only if Σ is simply connected,*

Note that the second assertion happens in the only dimension for which we can have an interesting difference between the connectivity and ayclicity properties of Σ, by the Hurewicz Theorem.

The power of the Bestvina–Brady Theorem lies in our freedom to choose the topological type of Σ.

Example 4.6. Take Σ to be a flag triangulation of the spine of a homology 3-sphere. Then Σ is 1-acylic but not simply connected, and hence of type FP_2 but not finitely presentable.

Example 4.7 (Bieri–Stallings). Let Σ be a flag triangulation of the n-sphere. Using the obvious model of suspension, this can be done with $\Sigma^{(1)})$ being the complete $(n + 1)$-coloured graph with two vertices of each colour, so $A_\Sigma \cong (F_2)^{n+1}$. Then H_Σ is of type FP_n (indeed, of type F_n) but not of type FP_{n+1}.

Bestvina and Brady were also able to use these techniques to show that either the Whitehead Conjecture or the Eilenberg–Ganea Conjecture is false.

The proof we give of the Bestvina–Brady Theorem follows Bux and Gonzalez [5].

4.3 Brown's criteria

To give a proof of the Bestvina–Brady Theorem, we will need criteria to check whether a group is of type FP_n. As this is not a course in homological algebra, we will not prove the results stated in this section; see [4] for proofs.

We have already seen how to prove that a group is of type FP_n from a very nice (ie free) action on a very nice (ie contractible) complex: if the action is cocompact on the n-skeleton, then we can deduce that our group is of type F_n, and so of type FP_n. For the Bestvina–Brady Theorem the quotient may not have a finite n-skeleton, so we will need to generalise this observation.

A sequence of groups $\cdots \to G_i \to G_{i+1} \to \dots$ is *essentially trivial* if, for all sufficiently large i, the map $G_i \to G_{i+1}$ is trivial.

Theorem 4.8 (Brown [4])**.** *Suppose Γ acts freely and cellularly on contractible X. Suppose furthermore that X has a Γ-equivariant filtration by connected subcomplexes $X_i \subseteq X_{i+1}$ such that X_i/Γ has finite n-skeleton for each i. Then Γ is of type FP_n if and only if the sequence of (reduced) homology groups $\widetilde{H}_n(X_i)$ is essentially trivial.*

This theorem has a homotopical analogue.

Theorem 4.9 (Brown [4])**.** *Suppose Γ acts freely and cellularly on contractible X. Suppose furthermore that X has a Γ-equivariant filtration by connected subcomplexes $X_i \subseteq X_{i+1}$ such that X_i/Γ has finite 2-skeleton for each i. If Γ is finitely generated then Γ is finitely presented if and only if the sequence of fundamental groups $\pi_1(X_i)$ is essentially trivial.*

4.4 Geodesic projection

We work with $\widetilde{\mathcal{S}}_\Sigma$, the universal cover of the Salvetti complex. The homomorphism h can be realised by a continuous map $\mathcal{S}_\Sigma \to S^1$, which lifts to height function, which we also denote by $h : \widetilde{\mathcal{S}}_\Sigma \to \mathbb{R}$. There is a natural choice for h such that each level set $h^{-1}(x)$ is convex and cuts every cube whose interior it intersects into two pieces.

Consider a vertex $x_0 \in \widetilde{\mathcal{S}}_N$, and consider $\mathrm{Lk}(x_0)$. Choose the identification $\mathrm{Lk}(x_0) \equiv D(\Sigma)$ so that the vertices v_i^+ are in directions of increasing height and the vertices v_i^- are in directions of decreasing height. Then Σ^+ is called the *ascending link* of x_0 and is denoted by $\mathrm{Lk}^+(x_0)$; likewise Σ^- is called the *descending link* of x_0 and is denoted by $\mathrm{Lk}^-(x_0)$.

As usual there is a natural projection $\pi : \widetilde{\mathcal{S}}_\Sigma \smallsetminus \{x_0\} \to \mathrm{Lk}(x_0)$. The key observation here is that the germ of a geodesic $[\gamma]$ emanating from x_0 can be canonically extended to an infinite geodesic γ. To be precise, the germ $[\gamma]$ lies in the interior of a unique cube, corresponding to a locally convex torus in $\mathcal{S}_\Sigma$. That torus lifts to a unique convex flat (that is, an isometric copy of Euclidean space) in $\widetilde{\mathcal{S}}_\Sigma$ containing x_0, and in that flat there is a unique geodesic line γ extending $[\gamma]$. If $[\gamma] \in \mathrm{Lk}^+(x_0)$ then $h(\gamma(t))$ is a monotonically increasing.

Let x_0 be a vertex of $\widetilde{\mathcal{S}}_\Sigma$ with $h(x_0) < a$. From this discussion, it follows that there is a well defined map $\rho : \mathrm{Lk}^+(x_0) \to h^{-1}(a)$ satisfying $\pi \circ \rho = \mathrm{id}_{\mathrm{Lk}^+(x_0)}$. If $a < b < c$ and then we have the following commutative diagram.

$$\begin{array}{ccc} h^{-1}[b,c] & \longrightarrow & h^{-1}[a,c] \\ {\scriptstyle\rho}\uparrow & & \downarrow{\scriptstyle\pi} \\ \mathrm{Lk}^+(x_0) & \longrightarrow & \mathrm{Lk}(x_0) \end{array}$$

The following lemmas are immediate, when one recalls that $\mathrm{Lk}^+(x_0)$ is a retract of $\mathrm{Lk}(x_0)$.

Lemma 4.10. *If $\pi_n(\Sigma)$ is non-trivial then the inclusion $h^{-1}[b,c] \hookrightarrow h^{-1}[a,c]$ induces a non-trivial map in π_n.*

Lemma 4.11. *If $\widetilde{H}_n(\Sigma)$ is non-trivial then the inclusion $h^{-1}[b,c] \hookrightarrow h^{-1}[a,c]$ induces a non-trivial map in $\widetilde{H}_n$.*

4.5 Combinatorial Morse Theory

Suppose $a < b$. Denote by $X_{a,b}$ the subcomplex of $\widetilde{\mathbb{S}}_\Sigma$ consisting of all cubes whose heights are contained between a and b.

Lemma 4.12. *The complex $X_{a,b}$ is a deformation retract of $h^{-1}[a,b]$.*

Proof. To obtain $X_{a,b}$ from $h^{-1}[a,b]$ one simply removes all the pieces of the cubes that are cut into two by the level sets $h^{-1}(a)$ and $h^{-1}(b)$. These pieces naturally deformation retract onto their upper or lower faces, and these natural deformation retractions agree on faces, so can be performed simultaneously. □

More generally, we have the following.

Lemma 4.13. *If $a \leq a' < a+1$ and $b-1 < b' \leq b$ then $X_{a,b}$ is homotopy equivalent to the space obtained from $h^{-1}[a',b']$ by coning off the images of the ascending links of the vertices between heights a and a' and the images of the descending links between heights b' and b.*

Proof. Note that the images of the ascending in $h^{-1}(a')$ are disjoint. Likewise, the images of the descending links in $h^{-1}(b')$ are disjoint.

Let C be the space obtained from $h^{-1}[a',b']$ by coning off ascending and descending links. Then $X_{a,b}$ is obtained from C by deleting every open cube that is divided into two by $h^{-1}(a)$ or $h^{-1}(b)$. As in the previous lemma, such cubes can be removed by a deformation retraction. □

Lemma 4.14. *If Σ is n-connected then $X_{a,b}$ is n-connected for all $a < b$.*

Proof. By the previous lemmas, if $a < a' < b' < b$ then the inclusion $X_{a',b'} \hookrightarrow X_{a,b}$ induces an isomorphism on π_i for $i \leq n$. But $\widetilde{\mathbb{S}}_\Sigma$ is contractible and the direct limit of the spaces $X_{a,b}$. □

Likewise, we have the homological analogue.

Lemma 4.15. *If Σ is n-acyclic then $X_{a,b}$ is n-acyclic for all $a < b$.*

Proof of Theorem 4.5. Note that H_Σ acts cocompactly on $X_{-n,n}$ for all n. Suppose that Σ is n-acyclic. By Lemma 4.15, $X_{-n,n}$ is n-acyclic for all n and so, by Theorem 4.8, the group H_Σ is of type FP_n.

Conversely, if $\widetilde{H}_i(\Sigma) \neq 1$ for some $i \leq n$ then by Lemma 4.11 the inclusion $h^{-1}[-n,n] \hookrightarrow h^{-1}[-n-1,n+1]$ induces a non-trivial map on $\widetilde{H}_i$ for all n. By Lemma 4.12, the inclusion $X_{-n,n} \hookrightarrow X_{-n-1,n+1}$ also induces a non-trivial map on $\widetilde{H}_i$ and so, by Theorem 4.8, H_Σ is not of type FP_n.

This completes the proof of part (1). The proof of part (2) is identical, but uses the homotopical versions of the same results. □

5 Special cube complexes

The Bestvina–Brady Theorem shows that right-angled Artin groups have very interesting subgroups. Recent work of Haglund and Wise [8] characterises exactly those cube complexes that correspond to subgroups of right-angled Artin groups. These cube complexes are called *special*.

5.1 Hyperplanes and their pathologies

Let $C = [-1, 1]^n$ be a cube. A *hypercube* in C is the intersection of C with some coordinate plane $\{x_i = 0\}$. For X a cube complex, we define an equivalence relation on the hypercubes of the cubes of X by insisting that, if $M_i \subseteq C_i$ are hypercubes and F is a common face of C_1 and C_2 then $M_1 \sim M_2$ if $M_1 \cap F = M_2 \cap F$.

Given an equivalence class of hypercubes $\{M\}$, the corresponding *hyperplane* can be constructed from the disjoint union of the hypercubes by gluing two faces of M and N if and only if those faces are identified in X. The result is a cube complex X equipped with a natural map $H \to X$. We will often abuse notation and also use H to denote the image of H in X.

The *cubical neighbourhood* U_H of a hyperplane H is the union of the open cubes that contain its image. The pullback

$$\begin{array}{ccc} N_H & \longrightarrow & U_H \\ \downarrow & & \downarrow \\ H & \longrightarrow & X \end{array}$$

is an interval bundle over H, called the *normal bundle* of H. If the normal bundle has trivial monodromy then H is said to be *two-sided*. Otherwise, it is *one-sided*.

For each 1-cell e of X there is a unique dual hyperplane, which we will denote by H_e. Two 1-cells that intersect the same hyperplane are called *parallel*. Hence, hyperplanes correspond exactly to parallelism classes of 1-cells.

The normal bundle has a natural boundary ∂N_H, which double covers H. If H is two-sided then ∂N_H has two components; choosing an orientation for a dual edge to H, we denote these components by $\partial_\pm N_H$.

Clearly a hyperplane is a cube complex, of dimension one lower than the dimension of X. In an effort to minimise confusion, we will denote by $\mathrm{Lk}_Z(z)$ the link of a point z in a space Z.

Lemma 5.1. *Let x_0 be a vertex of X and e and oriented 1-cell emanating from x. Then e determines a point of $\mathrm{Lk}_X(x_0)$, which we also denote by e. Let $y = e \cap H_e$. Then $\mathrm{Lk}_{H_e}(y) = \mathrm{Lk}_{\mathrm{Lk}_X(x_0)}(e)$.*

As links in flag complexes are flag, it follows that hyperplanes in non-positively curved cube complexes are also non-positively curved.

Example 5.2. Let X be a Salvetti complex $\mathcal{S}_\Sigma$. Let a be a 1-cell. By construction, every 1-cell parallel to a is equal to a, so hyperplanes correspond exactly to 1-cells of A_Σ, and hence to vertices of Σ. The 0-skeleton of H is equal to the intersection of H with the set of dual 1-cells. So $H^{(0)}$ is equal to a single point, and therefore H_a is also a Salvetti complex. The 1-cells of H are parallel to 1-cells of X that bound a square with a, so they correspond exactly to $\mathrm{Lk}_\Sigma(a)$.

We now describe some pathologies of hyperplanes. A pair of hyperplanes H_1 and H_2 *intersects* if the natural map $H_1 \cup H_2 \to X$ is not an injection. Likewise, they *osculate* if they do not intersect, but the natural map $\partial H_1 \cup \partial H_2 \to X$ is not an injection.

1. A hyperplane *self-intersects* if H intersects itself. Otherwise, H is called *embedded*.

2. A hyperplane *self-osculates* if H osculates itself. If H is two-sided and a pair of points in the same component of the boundary have the same image in X then H is said to *directly self-osculate*. If a pair of points in different components of the boundary have the same image in X then H is said to *indirectly self-osculate*.

3. A pair of hyperplanes H_1, H_2 *inter-osculates* if they both intersect and osculate.

Definition 5.3. A non-positively curved cube complex is called *special* if:

1. every hyperplane is two-sided;

2. no hyperplane self-intersects;

3. no hyperplane directly self-osculates; and

4. no pair of hyperplanes inter-osculates.

Example 5.4. If X is special then any covering space $\widehat{X}$ of X is also special.

Remark 5.5. In [8], the above defines an *A-special* cube complex, whereas a *special* need not have two-sided hyperplanes. This is not an important distinction: every special (in the sense of [8]) cube complex has an A-special double-cover.

Example 5.6. The surface of genus two, Σ, can be tiled symmetrically by eight right-angled hyperbolic pentagons. The 2-cells of the dual cell complex are squares (because the pentagons were right-angled), and the link of every vertex is a circle of circumference $5\pi/2$. This exhibits Σ as a non-positively curved square complex. The hyperplanes are precisely the curves in the tiling by pentagons, and one can check that this square complex is special. Every orientable surface finitely covers Σ, and so every orientable surface is a special square complex.

Example 5.7. Let $X = \mathcal{S}_\Sigma$. Parallelism in the cubes of X preserves orientations of 1-cells, from which it follows that every hyperplane is two-sided. If a hyperplane H_a were to self-intersect, it would follows that some square has every edge glued to a, which does not occur in the construction of X. If H_a directly self-osculates then it follows that H_a is dual to two distinct edges incident at the same vertex; but each hyperplane is dual to a unique edge. If H_a and H_b inter-osculate then it follows that a and b both bound a square and do not bound a square, a contradiction.

In particular, all subgroups of right-angled Artin groups are fundamental groups of (not necessarily compact) special cube complexes. Haglund and Wise proved a remarkable converse to this result. Given a cube complex X, we define the *hyperplane graph* of X, denoted by $\mathcal{H}(X)$, to be a graph whose vertex set is the set of hyperplanes of X; two hyperplanes are joined by an edge if and only if they intersect.

Let X be a cube complex with all walls two-sided. Fix a choice of transverse orientation on each hyperplane H of X, and for x_0 the unique vertex of $\mathcal{S}_{\mathcal{H}(X)}$, fix an identification $\mathrm{Lk}(x_0) \equiv \mathcal{D}(\mathcal{H}(X)$. We can now define a *characteristic map*

$$\phi : X \to \mathcal{S}_{\mathcal{H}(X)}$$

as follows.

1. The 0-skeleton $X^{(0)}$ is sent to the unique 0-cell of $\mathcal{S}_{\mathcal{H}(X)}$;
2. An oriented 1-cell e of X determines a unique hyperplane H_e with a choice of transverse orientation. The hyperplane H_e corresponds to a vertex in $\mathcal{H}(X)$, which determines a 1-cell $\hat{e}$ of $\mathcal{S}_{\mathcal{H}(X)}$. If the orientation of e coincides with the fixed transverse orientation of H_e then orient $\hat{e}$ pointing from H_e^- to H_e^+; otherwise, give it the reverse orientation. The map ϕ sends e is sent to $\hat{e}$, preserving orientations.
3. A higher dimensional cube C in X, spanned by edges $e_1, \ldots, e_n$, is sent by ϕ to the unique cube in $\mathcal{S}_{\mathcal{H}(X)}$ spanned by $\phi(e_1), \ldots, \phi(e_n)$, preserving orientations. Note that the necessary cube exists because $H_{e_1}, \ldots, H_{e_n}$ pairwise intersect and so span a simplex in $\mathcal{H}(X)$.

Theorem 5.8 (Haglund–Wise [8])**.** *If X is a special cube complex then the characteristic map $\phi_X : X \to \mathcal{S}_{\mathcal{H}(X)}$ is a local isometry.*

Proof. Because every point has a neighbourhood isometric to the cone on the link, it is sufficient to check that ϕ_X induces injection on links of vertices $x \in X$. Indeed, because no hyperplanes self-intersect, no pair of distinct vertices at distance $\pi/2$ in $\mathrm{Lk}_X(x)$ are identified in $\mathrm{Lk}_{\mathcal{S}_{\mathcal{H}(X)}}$; likewise, because no hyperplanes self-osculate, no pair of distinct vertices at distance greater than $\pi/2$ in $\mathrm{Lk}_X(x)$ are identified by ϕ. Because no pairs of hyperplanes inter-osculate, it follows that there is no pair of vertices $u, v \in \mathrm{Lk}_X(x)$ with $d(u, v) > \pi/2$ but $d(\phi(u), \phi(v)) = \pi/2$. It now follows by induction on the length of the shortest non-injective path that ϕ is injective on $\mathrm{Lk}_X(x)$, as required. □

Corollary 5.9. *The characteristic map ϕ lifts to an isometric embedding $\tilde{\phi} : \widetilde{X} \to \widetilde{\mathcal{S}}_{\mathcal{H}(X)}$. In particular, it induces an injective homomorphism $\phi_* : \pi_1(X) \to A_{\mathcal{H}(X)}$.*

Proof. Consider the lift $\tilde{\phi} : \widetilde{X} \to \widetilde{\mathbb{S}}_{\mathcal{H}(X)}$. Let γ be a geodesic path in $\widetilde{X}$ such that $\tilde{\phi} \circ \gamma$ is a loop in $\widetilde{\mathbb{S}}_{\mathcal{H}(X)}$. Then $\tilde{\phi} \circ \gamma$ is a local geodesic, and so must be constant, because $\widetilde{\mathbb{S}}_{\mathcal{H}(X)}$ is CAT(0) and so each based homotopy class contains a unique locally geodesic representative. In particular, the action of $\pi_1(X)$, when pushed forward by ϕ_*, is free on $\widetilde{\mathbb{S}}_{\mathcal{H}(X)}$. □

Corollary 5.10. *A group Γ is a subgroup of a right-angled Artin group if and only if Γ is the fundamental group of a (not necessarily compact) special cube complex.*

Example 5.11. Let us return to the the special square complex structure on the surface of genus two, from Example 5.6. The hyperplane graph Γ can be seen to be two pentagons identified along an adjacent pair of edges. By the previous corollary, we get an injection $\pi_1\Sigma \hookrightarrow \Gamma$. In fact, it is known that $\pi_1\Sigma$ can be embedded into the right-angled Artin group defined by the pentagon graph. Does this embedding come from a special square complex structure on Σ?[2]

Special cube complexes are a subject of active research. Wise has recently announced a proof that a very large class of hyperbolic 3-manifold groups are virtually the fundamental groups of compact special cube complexes. This implies a relation between two very important open questions in 3-manifold topology: a virtually Haken hyperbolic 3-manifold is virtually fibred. Wise also claims that all one-relator groups with torsion are virtually special. It is beginning to become apparent that many interesting classes of groups are, at least virtually, subgroups of right-angled Artin groups.

5.2 Right-angled Coxeter groups

Right-angled Coxeter groups are closely related to right-angled Artin groups, and also have associated CAT(0) cube complexes, but these cube complexes have a slightly richer structure. Coxeter groups can be thought of as generalised reflection groups.

Definition 5.12. Let $M = (m_{ij})$ be a symmetric $n \times n$ matrix with entries in $\mathbb{N} \cup \{\infty\}$. We require that $m_{ii} = 1$ for all i. The associated *Coxeter group* C_M has presentation

$$C_M \cong \langle r_1, \ldots, r_n \mid (r_i r_j)^{m_{ij}} = 1 \rangle \ .$$

(Here, $(r_i r_j)^\infty = 1$ means that the relation should be omitted.)

Example 5.13. The group of isometries of $\mathbb{R}^2$ generated by reflections in the sides of an equilateral triangle is isomorphic to C_M, where

$$M = \begin{pmatrix} 1 & 3 & 3 \\ 3 & 1 & 3 \\ 3 & 3 & 1 \end{pmatrix}$$

[2] Matt Day points out that the answer is 'yes'. The square complex is dual to a tiling of the surface of genus two by two pentagons and a decagon.

The idea that Coxeter groups are generalised reflection groups made concrete by the following theorem.

Theorem 5.14. *Define a bilinear form on* $\mathbb{R}^n$ *by* $\langle e_i, e_j\rangle = -\cos \pi/m_{ij}$*. The assignment*

$$r_i(x) = x - 2\langle e_i, x\rangle e_i$$

defines a faithful representation of C_M*. In particular, Coxeter groups are linear over* $\mathbb{R}$*.*

A Coxeter group C_M is called *right-angled* if every $m_{ij} \in \{2, \infty\}$. (The group of isometries of $\mathbb{R}^2$ generated by reflections in the sides of a square is an example.) In this case, instead of the Coxeter matrix, we can think of the Coxeter group as defined by the (simplicial) *nerve graph* N, which has n vertices corresponding to the generators, and an edge between a pair of generators if and only if they commute. That is,

$$C_N \cong \langle V(N) \mid \{v^2 \mid v \in V(N)]\},\ \{[u,v] \mid (u,v) \in E(N)]\}\rangle\ .$$

Note the similarity with the definition of the right-angled Artin group A_N. Indeed, there is a natural surjection $A_N \to C_N$. (More generally, there is a definition of a (not necessarily right-angled) Artin group A_M that surjects C_M.) Just as in the Artin-group case, we will often denote C_N by C_Σ, where Σ is the unique flag complex with $\Sigma^{(1)} = N$.

Example 5.15. Consider a regular pentagon P in the hyperbolic plane. If P is very small then it is approximately Euclidean, so its internal angles are close to $3\pi/5$. If P is very large then it is close to ideal, in which case the internal angles are approximately zero. Therefore, there is a regular pentagon with all internal angles equal to $\pi/2$. The group of isometries of the hyperbolic plane generated by reflections in the sides of P is isomorphic to the Coxeter group C_Σ, where C_Σ is the pentagon graph. (It is clear that C_Σ surjects the reflection group. We will eventually prove see that this surjection is an isomorphism.) Note that you should think of the pentagon graph as dual to the boundary of P.

Note that, in the right-angled case, every coefficient in the representation of C_Σ as a generalised reflection group is an integer.

Corollary 5.16. *Right-angled Coxeter groups are linear over* $\mathbb{Z}$*.*

Just as right-angled Artin groups are fundamental groups of Salvetti complexes, Coxeter groups also act nicely on cube complexes. However, the fact that they have torsion means that we cannot hope for them to act freely. Indeed, we have the following lemma.

Lemma 5.17. *Any finite-order isometry* γ *of a complete* CAT(0) *space* X *fixes a point.*

Proof. We give a proof under the hypothesis that X is proper. Choose $x \in X$, and let $Y = \langle\gamma\rangle x$, a finite set. It follows easily from the proper hypothesis that there is y_0 such that $\overline{B}_r(y_0)$ is of minimal radius and contains Y. The lemma is an immediate consequence of the claim that y_0 is unique. Suppose note, so $Y \subseteq \overline{B}_r(y_1)$. By the CAT(0) condition, $Y \subseteq \overline{B}_r(y)$ for every $y \in [y_0, y_1]$. Therefore, there is some element of Y, without loss of generality x, such that $r = d(x, y)$ for some y in the interior of $[y_0, y_1]$. But this contradicts the convexity of the metric. □

So we will not be able to write down a finite, non-positively curved complex with C_Σ as its fundamental group. One way around this is to work with *complexes of groups*, as introduced by Haefliger. Roughly, the idea is as follows. Suppose a group Γ acts on a (simplicial, say) complex $\widetilde{X}$, with the 'no rotations' condition that if some $\gamma \in \Gamma$ maps some simplex $\tilde{\sigma}$ to itself then γ fixes $\tilde{\sigma}$ pointwise. (This can always be ensured by subdividing.) We can give the quotient $X = \widetilde{X}/\Gamma$ the structure of a complex of groups by labelling each open simplex σ of X by the (conjugacy class of the) stabiliser of any choice of open simplex $\tilde{\sigma} \subseteq \widetilde{X}$ in the preimage. Furthermore, if τ is a face of σ then we should record the inclusion map of the stabiliser of a suitable choice of representative $\tilde{\tau}$ into the stabiliser of $\tilde{\sigma}$.

You can compute the fundamental group of a complex of groups X using van Kampen's Theorem, if you count the fundamental group of a simplex to be its label.

A complex of groups that arises as above, as the quotient of a complex by a group action, is called *developable*. Not all complexes of groups are developable. Bridson and Haefliger proved that non-positively curved complexes of groups are developable, but we shall exhibit explicit covers of our complexes of groups which are genuine complexes.

We can now define a cube complex of groups associated to a right-angled Coxeter group C_Σ. For a cube C, a codimension-one face C' labelled by $\mathbb{Z}/2$ is called a *mirror*.

Definition 5.18. Let Σ be a flag complex. As in the definition of the Salvetti complex, consider the cube $\mathcal{C} = [0,1]^{V(\Sigma)}$, identify Σ with a subcomplex of $\mathrm{Lk}(0)$ in the natural way, let $\pi : [0,1]^{V(\Sigma)} \smallsetminus \{0\} \to \mathrm{Lk}(0)$ be radial projection and consider the subcomplex $C' = \pi^{-1}(\Sigma) \cup \{0\}$ of C. The *Davis–Moussong complex* of Σ, which we denote by $\mathcal{DM}_\Sigma$, is the complex of groups obtained by setting every codimension-one face of C' that does not adjoin the origin to be a mirror; the lower dimensional faces that do not adjoin the origin are labelled by the direct product of the labels of the adjacent mirrors. It follows from van Kampen's Theorem that $\pi_1 \mathcal{DM}_\Sigma \cong C_\Sigma$.

Remark 5.19. The term Davis–Moussong complex actually usually refers to the universal cover of $\mathcal{DM}_\Sigma$.

Example 5.20. Let Σ be the pentagon graph. Then $\mathcal{DM}_\Sigma$ is a pentagon, each of whose sides is a mirror. This proves the claim that the group generated by reflections in the sides of P is C_Σ.

As mentioned above, one can appeal to a general theorem about non-positive curvature to prove that $\mathcal{DM}_\Sigma$ is developable. Alternatively, one could use the structure theory for elements of Coxeter groups to explicitly build a universal cover for $\mathcal{DM}_\Sigma$. In this course, we will explicitly build a finite complex that finitely covers $\mathcal{DM}_\Sigma$, and deduce that the $\mathcal{DM}_\Sigma$ is developable.

Remark 5.21. There is a natural candidate for a torsion-free subgroup of C_Σ of finite index. By Lemma 5.17, if $\mathcal{DM}_\Sigma$ is developable then every finite-order element of C_Σ is conjugate to the label of some face in $\mathcal{DM}_\Sigma$, in other words to a generator of C_Σ. But each generator of C_Σ survives under the map to the abelianisation, so we have a short exact sequence

$$1 \to [C_\Sigma, C_\Sigma] \to C_\Sigma \to H_1(C_\Sigma) \to 1 \ .$$

in which the commutator subgroup $[C_\Sigma, C_\Sigma]$ is torsion-free.

The abelianisation map $C_\Sigma \to H_1(C_\Sigma)$ can be geometrically represented by the inclusion of $\mathcal{DM}_\Sigma$ into the cube $\mathcal{C}$, where again each face not incident at the origin is mirrored. The universal cover of $\mathcal{C}$ can be thought of as the cube $\widetilde{\mathcal{C}} = [-1,1]^{V(\Sigma)} \subseteq \mathbb{R}^{V(\Sigma)}$, with $\pi_1(\mathcal{C}) = (\mathbb{Z}/2\mathbb{Z})^{V(\Sigma)}$ acting in the natural way by reflection in the coordinate planes. The preimage of $\mathcal{DM}_\Sigma$ in $\mathcal{C}$ is a finite-sheeted cover of $\mathcal{DM}_\Sigma$ is a connected cube complex $\widehat{\mathcal{DM}}_\Sigma$, which covers $\mathcal{DM}_\Sigma$ with covering group $(\mathbb{Z}/2\mathbb{Z})^{V(\Sigma)}$.

$$\begin{array}{ccc} \widehat{\mathcal{DM}}_\Sigma & \longrightarrow & \widetilde{\mathcal{C}} \\ \downarrow & & \downarrow \\ \mathcal{DM}_\Sigma & \longrightarrow & \mathcal{C} \end{array}$$

Remark 5.22. The cube complex structures on $\widehat{\mathcal{DM}}_\Sigma$ and $\widetilde{\mathcal{C}}$ are actually the first subdivisions of coarser cube complex structures. We will use this coarser structure without comment. If we need to make the action of $(\mathbb{Z}/2\mathbb{Z})^{V(\Sigma)}$ cellular, we will need to work with the subdivision.

Lemma 5.23. *The complex* $\widehat{\mathcal{DM}}_\Sigma$ *is non-positively curved.*

Proof. By construction, every vertex of $\widehat{\mathcal{DM}}_\Sigma$ has link Σ, which is flag complex by hypothesis. □

The following theorem is an immediate corollary.

Theorem 5.24 (Davis–Moussong)**.** *The right-angled Coxeter group* C_Σ *acts properly discontinuously and cocompactly by isometries on a* CAT(0) *cube complex* $\widetilde{\mathcal{DM}}_\Sigma$.

Right-angled Coxeter groups give a more flexible construction of non-positively curved cube complexes than right-angled Artin groups. Recall that the link of a vertex in a Salvetti complex is necessarily a double; by contrast, the link of a vertex in $\mathcal{DM}_\Sigma$ is Σ on the nose. Davis was able to use his complex to construct the first examples of closed aspherical manifolds with universal covers not homeomorphic to $\mathbb{R}^n$.

Remark 5.25. Moussong proved in his thesis that every Coxeter group acts properly discontinuously and cocompactly on some CAT(0) complex.

We have already seen that right-angled Coxeter groups are naturally quotients of right-angled Artin groups. In this case, we shall see that they also have other natural relationships.

Proposition 5.26. *The complex $\widehat{\mathcal{DM}}_\Sigma$ is special.*

Proof. For each $v \in V(\Sigma)$, the vertices of $\widehat{\mathcal{DM}}_\Sigma$ can be partitioned into those with coordinate $+1$ in the vth place, and those with coordinate -1. The edges between vertices in the two sets of the partition are precisely the edges dual to the union of hyperplanes that cover the 'mirrored' face of $\mathcal{DM}_\Sigma$ perpendicular to the edge determined by v. It follows that each hyperplane is two-sided. If a hyperplane were to self-intersect, then it would follow that some mirror mirrored face of $\mathcal{DM}_\Sigma$ would intersect itself, which does not happen by definition. Self-osculation does not occur for the same reason. Likewise, if two hyperplanes were to inter-osculate, it would follow that two mirrored faces of $\mathcal{DM}_\Sigma$ both intersect and osculate, which again contradicts the construction of $\mathcal{DM}_\Sigma$. □

A property is said to *virtually* hold of a group Γ if it holds for some subgroup of finite index.

Corollary 5.27. *Every right-angled Coxeter group is virtually a subgroup of a right-angled Artin group.*

This result has a sort of converse.

Theorem 5.28 (Davis–Januszkiewicz [6])**.** *For a flag complex Σ, the homomorphisms*

$$A_\Sigma \to H_1(A_\Sigma, \mathbb{Z}/4\mathbb{Z})$$

and

$$C_{\mathcal{D}(\Sigma)} \to H_1(C_{\mathcal{D}(\Sigma)}, \mathbb{Z}) \equiv H_1(C_{\mathcal{D}(\Sigma)}, \mathbb{Z}/2\mathbb{Z})$$

have isomorphic kernels. In particular, A_Σ and $C_{\mathcal{D}(\Sigma)}$ are commensurable.

Corollary 5.29. *Right-angled Artin groups are linear over $\mathbb{Z}$.*

Proof. By the theorem, A_Σ has a finite-index subgroup A_0 with a faithful representation into $GL_n(\mathbb{Z})$ for some n. The induced representation of A_Σ is then a faithful representation of A_Σ into $GL_{n|A_\Sigma:A_0|}(\mathbb{Z})$. □

Proof of Theorem 5.28. To prove the theorem, we will construct a free action of $\mathbb{Z}/4\mathbb{Z}^{\Sigma^{(0)}}$ on $\widehat{\mathcal{DM}}_{\mathcal{D}(\Sigma)}$ and observe that the quotient is a Salvetti complex.

For each $v \in \Sigma^{(0)}$, let $r_{v^\pm}$ be the reflection of $\mathbb{R}^{\mathcal{D}(\Sigma)^{(0)}}$ in the coordinate plane defined by $v^\pm$. Likewise, let s_v be the transformation that swaps the

v^+ and v^- coordinates. We know that $r_{v^\pm}$ preserves $\widehat{\mathcal{DM}}_{\mathcal{D}(\Sigma)}$ by construction. Because s_v can be realised by an involution of $\mathcal{D}(\Sigma)$, it is also true that s_v preserves $\widehat{\mathcal{DM}}_{\mathcal{D}(\Sigma)}$. The transformation $t_v = r_{v^+}s_v$ is rotation by $\pi/2$ in the plane spanned by the v^+ and v^- axes, and preserves $\widehat{\mathcal{DM}}_{\mathcal{D}(\Sigma)}$.

Consider the action of the group

$$Q = \langle t_v \mid v \in \Sigma^{(0)} \rangle \cong (\mathbb{Z}/4\mathbb{Z})^{\Sigma^{(0)}}$$

on $\widehat{\mathcal{DM}}_{\mathcal{D}(\Sigma)}$. First, note that the action is free; for, a product of rotations only fixes a point if the projection of that point to the plane of rotation is the origin; because v^+ and v^- are not joined by an edge in $\mathcal{D}(\Sigma)$, no point of $\widehat{\mathcal{DM}}_{\mathcal{D}(\Sigma)}$ projects to the origin in the plane spanned by the v^+- and v^--axes.

From the Orbit–Stabilizer Theorem, it follows that $\widehat{\mathcal{DM}}_{\mathcal{D}(\Sigma)}/Q$ has only one vertex, with link $\mathcal{D}(\Sigma)$. For the same reason, the quotient has exactly n edges. To prove that the quotient is a Salvetti complex, it remains only to verify that the quotient has exactly $|\Sigma^{(0)}|$ hyperplanes and that each hyperplane is two-sided.

Each hyperplane of $\widehat{\mathcal{DM}}_{\mathcal{D}(\Sigma)}$ separates vertices that differ in one coordinate $v \in \mathcal{D}(\Sigma)^{(0)}$. We will say that the hyperplanes determined by the v^+ and v^- coordinates are equivalent. This divides the hyperplanes of $\widehat{\mathcal{DM}}_{\mathcal{D}(\Sigma)}$ into $|\Sigma^{(0)}|$ equivalence classes, and these classes are preserved by the action of Q. We therefore have that

$$|\Sigma^{(0)}| = |\{\text{edges of quotient}\}| \geq |\{\text{hyperplanes of quotient}\}| \geq |\Sigma^{(0)}|$$

so the quotient has exaxtly n hyperplanes.

Finally, to show that the quotient is two-sided, orient the edges so that the projection to the (v^+, v^-)-plane is a cycle. This orientation is preserved by the action of Q. □

5.3 Examples of virtually special groups: doubles

In this section, we will begin to see some of the techniques that go into proving that a group is virtually special.

Example 5.30. Let F be a free group of finite rank and let $w \in F \setminus 1$. We consider the double $\Gamma = F *_{\langle w \rangle} F$, which is the fundamental group of a compact cube complex as follows.

Let A be a compact graph of appropriate rank (a rose, say), and fix an isomorphism $F \cong \pi_1(A)$. Equipped with any piecewise Euclidean metric, the resulting metric graph is, of course, non-positively curved. Realise the conjugacy class of w by a local isometry $S^1 \in A$, which we will denote by ∂. Subdivide S^1 to make the map w combinatorial—that is, so that it sends vertices to vertices and edges to edges. Now take two copies of A, denoted by A^+ and A^-, and denote the copy of the map ∂ that maps to $A^\pm$ by $\partial^\pm$. Construct a space

$$X = A^- \sqcup (S^1 \times [-1, 1]) \sqcup A^+ / \sim$$

where the equivalence relation $\sim$ identifies $(\theta, \pm 1) \in S^1 \times [-1, 1]$ with $\partial^\pm(\theta) \in A^\pm$.

The cellular structure on S^1 extends to a square-complex structure on $S^1 \times [-1, +1]$, which in turn is preserved by the gluing maps $\partial^\pm$, so X has the structure of a square complex. Furthermore, because ∂ is a local isometry, it follows that points of $A^\pm$ in X have neighbourhoods that are isometric to a set of half-discs glued to trees; in particular, X is non-positively curved.

We will prove that X is virtually special. To do so, we need to examine the hyperplanes of X, which can be partitioned into two sets. There is a unique *vertical* hyperplane V at the centre of the gluing cylinder and disjoint from $A^+ \cup A^-$. It is isometric to a circle. The *horizontal* hyperplanes H_e are each dual to a unique edge $e^\pm$ of $A^\pm$; they are isometric to graphs with vertices in $A^+ \cup A^-$, with edges that run transversely through the gluing cylinders.

First, note that each hyperplane is two-sided. Indeed, this is clear for V; choose consistent orientations on $e^\pm$ to see that the horizontal hyperplane H_e is two-sided.

Each horizontal hyperplane only intersects the vertical hyperplane, and *vice versa*. From this, it follows that there is no self-intersection, and also no inter-osculation.

The horizontal hyperplanes do not directly self-osculate. Indeed, if H_e directly self-osculated, it would follow that the map $\partial : S^1 \to A$ identified two neighbouring edges of S^1 together with opposite orientations; but this contradicts the hypothesis that ∂ is a local isometry.

Finally, however, the vertical hyperplane V may directly self-osculate. Indeed, this happens whenever a the cyclically reduced representative of w includes the same letter twice. So X is not, in general, special. We will see how to remove this pathology by passing to a finite-sheeted cover of X.

Lemma 5.31. *Consider a combinatorial (ie sending vertices to vertices and edge to edges) local isometry $A' \to A$ of finite metric graphs. Then there is a finite-sheeted covering map $\overline{A} \to A$ and an isometric embedding $A' \hookrightarrow \overline{A}$ such that the map $A' \to A$ factors as the composition of the embedding and the covering map.*

The local isometry condition means that $A' \to A$ is injective on links. To make it a covering map, we need to make it bijective on links. The proof is a nice exercise: simply look for missing edges and then pair them up any way you like. The following lemma is an immediate consequence.

Lemma 5.32. *There is a finite-sheeted cover $\overline{A} \to A$ to which ∂ lifts as an embedding.*

Lemma 5.33. *There is a finite-sheeted cover $\widehat{A} \to A$ to which every component of the pullback of ∂ is an embedding.*

Proof. Take $\pi_1 \widehat{A}$ to be the normal core of $\pi_1(\overline{A})$. □

The special cover of X is now obtained by gluing two copies of $\widehat{A}$ along cylinders, using the pullback of ∂.

5.4 Separability

In the last subsection, we saw how it can be possible to remove pathologies in a finite-index cover, and thus prove that the fundamental group of a cube complex is *virtually* special.

Definition 5.34. Let Γ be a group. A subset $X \subseteq \Gamma$ is *separable* if, for any $\gamma \notin X$, there exists a homomorphism q from Γ to a finite group such that $q(\gamma) \notin q(X)$.

Example 5.35. A group Γ is residually finite if and only if the trivial subgroup is separable.

Remark 5.36. A subgroup $H \subseteq \Gamma$ is separable if and only if, for any $\gamma \notin H$, there exists a subgroup K of finite index in Γ with $\gamma \notin K \supseteq H$. To see this that this follows from the definition, take $K = q^{-1}(q(H))$. Conversely, given K, one can take q to be the natural map $\Gamma \to \mathrm{Sym}(\Gamma/K)$ induced by left-translation.

Lemma 5.37. *A normal subgroup $H \lhd \Gamma$ is separable if and only if Γ/H is residually finite.*

From the existence of non-residually finite groups, we can therefore deduce that free groups have non-separable subgroups. However, these subgroups are never finitely generated. (Indeed, Greenberg's Theorem asserts that a non-trivial normal subgroup of a free group is either of finite index or infinitely generated.)

Definition 5.38. A group Γ is called *extended residually finite (ERF)* if every subgroup is separable, and *locally extended residually finite (LERF)* or *subgroup separable* if every finitely generated subgroup is separable.

We have seen that non-abelian free groups are not ERF. Indeed, it is an open question whether every ERF group is virtually polycyclic. We will see that free groups are LERF.

Example 5.39. Let Q be a finitely presented group and let $q : F \to Q$ be a surjection. Earlier in the course (but not in the notes...), we saw that the fibre product $H \subseteq F \times F$ is finitely generated. It is straightforward to check that the fibre product $H \subseteq F \times F$ is separable if and only if Q is residually finite. So we see that $F \times F$ is not LERF.

Separability is a very useful property in topology, because of the following lemma.

Lemma 5.40. *Let $X' \to X$ be a covering map of CW-complexes. Then $\pi_1 X'$ is separable in $\pi_1 X$ if and only if, for every compact subset $\Delta \subseteq X'$, the covering map factors through a finite-sheeted covering map $\widehat{X} \to X$ so that Δ embeds in $\widehat{X}$.*

Proof. To be added later... □

Combining this with Lemma 5.31, one quickly obtains the following theorem.

Theorem 5.41 (Marshall Hall Jr, 1949)**.** *Free groups are LERF.*

Proof. Exercise. □

Using Lemma 5.40, one can lift self-intersections and osculations away in a finite-sheeted cover. A similar idea works for inter-osculation, although in this case we need to think about double cosets. Moreover, it turns out that the converses of these statements also hold. This gives an algebraic characterisation of virtually special cube complexes.

Definition 5.42. Let X be a non-positively curved cube complex and let $\Gamma = \pi_1 X$. A *hyperplane subgroup* of Γ is a subgroup conjugate to the image of $\pi_1 H$, where H is a hyperplane of X. (Note that because we only care about the conjugacy class, choices of base points do not matter.)

Remark 5.43. The hyperplane H is a locally convex subspace of X, and it follows as usual that the induced map $\pi_1 H \to \Gamma$ is injective.

Theorem 5.44 (Haglund–Wise)**.** *Let X be a compact, non-positively curved cube complex. Then X is virtually special if and only if the following conditions hold.*

1. *Every hyperplane subgroup H is separable.*
2. *For every pair of hyperplane subgroups H_1, H_2 corresponding to intersecting hyperplanes, the double coset $H_1 H_2$ is separable.*

In one direction, the key idea is Lemma 5.40. Consider a hyperplane Y with corresponding hyperplane subgroup H. Let $X^H \to X$ be the covering map corresponding to H. The natural map $\overline{N}(Y) \to X$ is a local isometry, and so lifts to a local isometry $\overline{N}(Y) \to X^H$.

This map is injective. Indeed, if not, then one can find a geodesic γ in $\overline{N}_Y$ that maps to a based locally geodesic loop in X^H. But this loop must then be homotopic (relative to basepoints) to the image of a based loop in $\overline{N}_Y$, which in turn has a locally geodesic representative. This contradicts the fact that each based homotopy class contains a unique local geodesic.

If H is separable then it follows from Lemma 5.40 that there is a finite-sheeted covering space $\widehat{X}_Y$ of X to which the lift of the natural map $\overline{N}_Y \to X$ is an embedding; in particular, the lift of Y is two-sided and neither self-intersects nor self-osculates. The covering space corresponding to any normal subgroup of $\bigcap_Y \pi_1 \widehat{X}_Y$, where Y ranges over all the hyperplanes of X, has these properties for every hyperplane.

A generalisation of Lemma 5.40 uses double-coset separability to deal with inter-osculation in exactly the same way. This proves one direction of the theorem.

To start to prove the converse, it is convenient to think about separability a little more conceptually.

Definition 5.45. The *profinite topology* on a group Γ is the coarsest topology such that every homomorphism from Γ to a finite group (equipped with the discrete topology, of course) is continuous. That is, every coset of a finite-index subgroup in Γ is open.

Exercise 5.46. A subset $X \subseteq \Gamma$ is separable if and only if it is closed in the profinite topology.

As the profinite topology is algebraically defined, any map defined on Γ using its algebraic structure will be continuous in the profinite topology. This is a convenient trick for proving that subsets of Γ are separable.

Lemma 5.47. *Let H, K be subgroups of a residually finite group Γ.*

1. *If there is a retraction $\rho : \Gamma \to H$ then H is separable.*
2. *If, furthermore, $\rho(K) \subseteq K$ and K is separable then the double coset HK is separable.*

Proof. Note that, in fact, 1 is a special case of 2 with $K = 1$, so it suffices to prove 2. The map defined by $\phi(\gamma) = \rho(\gamma)^{-1}\gamma$ is continuous in the profinite topology on Γ. The result now follows from the observation that

$$\phi^{-1}(K) = HK\ .$$

□

Note that the hypothesis that Γ is residually finite cannot be relaxed—the trivial subgroup is always a retract.

Proposition 5.48. *Hyperplane subgroups and hyperplane double cosets in right-angled Artin groups are always separable.*

Proof. Let A_N be a right-angled Artin group and v a vertex of N. The hyperplane subgroup corresponding to the hyperplane dual to v is the natural copy of $A_{\mathrm{Lk}v}$. There is a retraction $A_N \to A_{\mathrm{Lk}v}$ defined by killing every generator not in Lk_v, and this retraction sends $A_{\mathrm{Lk}u}$ into itself for every vertex u. As A_N is linear over $\mathbb{Z}$ and hence residually finite, the proposition follows from the preceding lemma. □

Corollary 5.49. *If X is special then hyperplane subgroups are separable.*

Proof. Observe from the definition of the characteristic map $\phi : X \to \mathcal{S}_{\mathcal{H}(X)}$ that, for any hyperplane H, $\phi(\pi_1 H) = \pi_1 X \cap A_{\mathrm{Lk}H_e}$. □

Unfortunately, I do not know a similarly cheap proof that that the hyperplane double cosets are separable. Haglund and Wise's proof uses a sophisticated completion, called the 'canonical' completion.

5.5 (Quasi-)convex subgroups

Above, we saw that $F \times F$ is not LERF. But it is a right-angled Artin group, so we cannot hope to show that special groups are LERF. The trouble is that the non-separable subgroups of $F \times F$ are 'badly behaved'—indeed, they are not even finitely presentable. We need a notion of a well behaved subgroup of the fundamental group of a cube complex.

Definition 5.50. Let X be a non-positively curved cube complex and write $\Gamma = \pi_1 X$. A subgroup $H \subseteq \Gamma$ is called *combinatorially convex-cocompact* if there is a convex subcomplex $\widetilde{Y}$ of the universal cover $\widetilde{X}$ which is preserved by the action of H, on which H acts cocompactly.

Example 5.51. Consider the action of $\mathbb{R}^2$ with the standard cubulation, and the corresponding action of $\mathbb{Z}^2$. The coordinate subgroups of $\mathbb{Z}^2$ are combinatorially convex-cocompact, but no diagonal cyclic subgroup is.

So we see that, in general, this notion is heavily dependent on the choice of cube complex.

Example 5.52. Let X be a compact special cube complex, and let $\phi : X \to \mathcal{S}_{\mathcal{H}(X)}$ be its characteristic map to a Salvetti complex. This map lifts to an isometric embedding of universal covers

$$\tilde{\phi} : \widetilde{X} \to \widetilde{\mathcal{S}}_{\mathcal{H}(X)} .$$

By uniqueness of geodesics, it follows that $\tilde{\phi}(\widetilde{X})$ is a convex subcomplex of $\widetilde{\mathcal{S}}_{\mathcal{H}(X)}$, and so $\phi_* \pi_1 X$ is a combinatorially convex-cocompact subgroup of $A_{\mathcal{H}(X)}$.

In geometric group theory, one of the standard notions of a well-behaved subgroup is quasiconvexity.

Definition 5.53. Let X be a geodesic metric space. A subspace Y is called *κ-quasiconvex*, for $\kappa \geq 0$, if every geodesic with endpoints in Y is contained in the κ-neighbourhood of Y.

Definition 5.54. Let Γ be a group and S a generating set. A subgroup H is called *word-quasiconvex* if it is a κ-quasiconvex subspace of $\mathrm{Cay}_S(\Gamma)$, for some κ.

In general this notion depends very heavily on the choice of generating set, as the example of a diagonal subgroup of $\mathbb{Z}^2$ once again shows. However, if Γ happens to be word-hyperbolic then it follows from Lemma 2.14 that word-quasiconvexity is independent of the choice of generating set, and in fact is equivalent to the subgroup's being quasi-isometrically embedded. In the case of a right-angled Artin group, which of course is almost never word-hyperbolic, the definition implies a choice of generating set, and we shall always mean this when we talk about word-quasiconvex subgroups.

Haglund related combinatorial convex cocompactness to a condition that relates well to these, more algebraic, notions.

Definition 5.55. Let X be a non-positively curved cube complex with fundamental group Γ. Choose a base vertex x_0 in the universal cover $\widetilde{X}$. Consider the 1-skeleton of the universal cover $\widetilde{X}^{(1)}$, equipped with the induced length metric. A subgroup H of Γ is called *combinatorially quasiconvex* if the orbit Hx_0 is a quasiconvex subspace of $\widetilde{X}^{(1)}$.

Note that, in general, this definition depends on the choice of basepoint. However, Haglund and Wise proved that, in the special case, the definition is independent of the choice of basepoint.

Example 5.56. In the case of a Salvetti complex, the 1-skeleton of the universal cover is isometric to the Cayley graph for the standard generating set. Therefore, a subgroup is combinatorially quasiconvex if and only if it is word-quasiconvex.

Example 5.57. If X is compact and Γ is word-hyperbolic then, by Lemmas 1.7 and 2.14, a subgroup H is combinatorially quasiconvex if and only if it is word-quasiconvex.

Theorem 5.58 (Haglund [7])**.** *Let X be a non-positively curved cube complex. A subgroup H of the fundamental group is combinatorially quasiconvex if and only if it is combinatorially convex-cocompact.*

One direction is clear; the non-trivial direction is to show that a combinatorially quasiconvex subgroup is combinatorially convex-cocompact. In particular, it follows that word-quasiconvex subgroups of word-hyperbolic groups and right-angled Artin groups are combinatorially convex-cocompact.

Scott proved that surface groups are LERF in 1978 [9]. His proof actually made use of the fact that every finitely generated subgroup of a surface group is geometrically finite, which is equivalent to word-quasiconvexity. His proof (apparently at Thurston's suggestion) in fact realises the surface group as a finite-index subgroup of the reflection group of a regular right-angled pentagon. Almost 30 years later, Haglund showed that Scott's argument works for arbitrary right-angled Coxeter groups. Long and Reid observed that, in fact, the argument proves the deeper fact that every quasiconvex subgroup is a virtual retract, ie a retract of a finite-index subgroup.

Theorem 5.59 (Scott [9], Haglund [7])**.** *Let H be a combinatorially convex-cocompact subgroup of a right-angled Coxeter group C_N. There is a subgroup K of finite index in C_N and a retraction $\rho : K \to H$. In particular, H is a separable subgroup of C_N.*

Corollary 5.60. *Let X be a special cube complex with finitely many walls. Every combinatorially convex-cocompact subgroup of $\pi_1 X$ is a virtual retract, and therefore separable.*

We sketch a proof of Theorem 5.59 using the theory of complexes of groups.

Sketch proof of Theorem 5.59. As usual, we think of C_N as the fundamental group of the Davis–Moussong complex $\mathcal{DM}_N$. By hypothesis, the covering space

$\mathcal{DM}'_N = \widetilde{\mathcal{DM}}_N/H$ of $\mathcal{DM}_N$ that corresponds to H has a compact, connected core $Y' = \widetilde{Y}/H$ that carries the fundamental group H.

A hyperplane of $\mathcal{DM}'_N$ is said to neighbour Y' if it is dual to an edge that has one end in Y' and one out of Y'. Let $\widehat{\mathcal{DM}}_N$ be the maximal connected subcomplex of the first subdivision of $\mathcal{DM}'_N$ that contains Y' and that is bounded by hyperplanes that neighbour Y'. This defines a finite-sheeted covering space of $\mathcal{DM}_N$. It remains to see that $\pi_1\widehat{\mathcal{DM}}_N$ retracts onto $\pi_1 Y'$. Indeed, Y' is a deformation retract of $\widehat{\mathcal{DM}}_N$, and the retraction is given by forgetting the mirrors and applying the deformation retraction. □

References

[1] Werner Ballmann. Singular spaces of nonpositive curvature. In *Sur les groupes hyperboliques d'après Mikhael Gromov (Bern, 1988)*, volume 83 of *Progr. Math.*, pages 189–201. Birkhäuser Boston, Boston, MA, 1990.

[2] Mladen Bestvina and Noel Brady. Morse theory and finiteness properties of groups. *Inventiones Mathematicae*, 129:445–470, 1997.

[3] Martin R. Bridson and André Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[4] Kenneth S. Brown. Finiteness properties of groups. In *Proceedings of the Northwestern conference on cohomology of groups (Evanston, Ill., 1985)*, volume 44, pages 45–75, 1987.

[5] Kai-Uwe Bux and Carlos Gonzalez. The Bestvina–Brady construction revisited: geometric computation of σ-invariants for right-angled Artin groups. *Journal of the London Mathematical Society. Second Series*, 60:793–801, 1999.

[6] Michael W Davis and Tadeusz Januszkiewicz. Right-angled artin groups are commensurable with right-angled coxeter groups. *Journal of Pure and Applied Algebra*, 153(3):229–235, 2000.

[7] Frédéric Haglund. Finite index subgroups of graph products. *Geometriae Dedicata*, 135:167–209, 2008.

[8] Frédéric Haglund and Daniel T. Wise. Special cube complexes. *Geometric and Functional Analysis*, 17(5):1551–1620, 2008.

[9] Peter Scott. Subgroups of surface groups are almost geometric. *Journal of the London Mathematical Society. Second Series*, 17(3):555–565, 1978.