Notes On Hyperbolic and Automatic Groups

Michael Batty, after Panagiotis Papasoglu

October 21, 2003

Contents

1 Introduction

These notes have grown from an M.Sc. lecture course given by Panagiotis Papasoglu at the University of Warwick in Spring 1994. They were written between 1994 and 1998 and revised in 2003 and are suitable for a course for first-year postgraduates or perhaps final year honours students. In parts, these notes are closely related to the notes in [61]. We assume that the reader is familiar with elementary group theory plus the notion of a free group and a presentation of a group. An elementary knowledge of metric spaces and point-set topology is also assumed.

The first two sections of these notes concerns the theory of *hyperbolic* groups and metric spaces. This makes precise the notion of a group being *negatively curved*. The theory of hyperbolic groups and the philosophy of considering groups as coarse geometric objects was developed by Gromov (see [22],[23] and [24]). The idea of a hyperbolic group generalises on the much earlier work of Dehn on surface groups (see [14] for a good account of Dehn's work) and also parts of the so-called *small cancellation theory* of Tartaskii, Greendlinger and Lyndon-Schupp (REFS?). Dehn asked in 1912 the question of when a curve on a compact orientable surface can be continuously shrunk to a point. Re-phrased in terms of the fundamental group G of the surface, this leads to a formulation of the *word problem* for G, which asks: Given a finite presentation of a group G, does there exist an algorithm which takes as input a word w in the generators and decides whether or not w represents the identity of G? Dehn showed that the answer was yes for a surface group and demonstrated such an algorithm, which is now known as *Dehn's Algorithm*. Gromov generalised this to hyperbolic groups, where it turns out that the existence of a Dehn Algorithm is characterised by Gromov's notion of hyperbolicity.

In a parallel development, Cannon, Epstein, Thurston and others produced the theory of *automatic* and *biautomatic* groups, which is a combination of the theory of regular languages and finite state automata with the geometry of groups. The third section of these notes concerns automatic groups. It will be seen that there is a strong connection with Gromov's theory. In some sense, automatic groups can be thought of as *nonpositively curved*. We give Cannon's proof that hyperbolic groups are examples of automatic groups. We also give Gersten and Short's proof that most nilpotent groups are not biautomatic.

Michael Batty is currently supported by EPSRC MathFIT grant number GR/87406.

2 Hyperbolic Metric Spaces

The essential properties of the hyperbolic plane are abstracted to obtain the notion of a *hyperbolic metric space*, which is due to Gromov [23]. The geometry of such spaces is explored.

2.1 Quasi-Isometries.

Definition 2.1 *Let (X, d_X) and (Y, d_Y) be metric spaces. We say that X and Y are* quasi-isometric *if there are functions $f : X \to Y$ and $g : Y \to X$ and constants $\lambda > 0$ and $c \geqslant 0$ such that*

1. *For all x_1 and x_2 in X, $d_Y(f(x_1), f(x_2)) \leqslant \lambda d_X(x_1, x_2) + c$*
2. *For all y_1 and y_2 in Y, $d_X(g(y_1), g(y_2)) \leqslant \lambda d_Y(y_1, y_2) + c$*
3. *For all $x \in X$, $d_X(g \circ f(x), x) \leqslant c$*
4. *For all $y \in Y$, $d_Y(f \circ g(y), y) \leqslant c$*

The pair of functions (f, g) is called a quasi-isometry*.*

Exercise 2.2 Show that quasi-isometry is an equivalence relation on metric spaces.

Example 2.3

1. if X and Y are both bounded metric spaces then X and Y are quasi-isometric.
2. Example: $\mathbb{Z}^n$ is quasi-isometric to $\mathbb{R}^n$. A quasiisometry is given by (i, j) where i is the inclusion map $\mathbb{Z}^n \to \mathbb{R}^n$ and j is the map from $\mathbb{R}^n$ to $\mathbb{Z}^n$ which is the n-fold cartesian product of the map taking a real number to its integer part. Notice that the maps are not bijections.

A property $\mathcal{P}$ of metric spaces such that whenever a metric space X has property $\mathcal{P}$ then all metric spaces quasi-isometric to X also have property $\mathcal{P}$ is called a *quasi-isometry invariant*.

Note that Gromov has given an equivalent formulation of quasi-isometry via Hausdorff and Lipschitz equivalence (see [24]).

2.2 Geodesic Metric Spaces.

Let (X, d) be a metric space and p be a path $[0, 1] \to X$. If the supremum over all finite partitions $[t_0 = 0, t_1, \ldots, t_{n-1}, t_n = 1]$ of

$$\sum_{i=1}^{n} d(p(t_{i-1}), p(t_i))$$

exists then we say that p is a *rectifiable* path. We denote this supremum by $l(p)$ and call it the *length* of p. Now (X, d) is called a *length space* or a *path metric space* if for all x_1 and x_2 in X,

$$d(x_1, x_2) = \inf\{l(p) \mid p \text{ is a rectifiable path in } X \text{ from } x_1 \text{ to } x_2\}.$$

Definition 2.4 *Let (X, d) be a metric space, and let x_1 and x_2 be two points of X. A* geodesic *between x_1 and x_2 is an isometric embedding γ of the real interval $[0, d(x_1, x_2)]$ into X such that $\gamma(0) = x_1$ and $\gamma(d(x_1, x_2)) = x_2$.*

We also extend the definition of a geodesic to isometric embeddings $[0, \infty) \to X$ (called *geodesic rays*) and $\mathbb{R} \to X$ (called *bi-infinite geodesics*).

Definition 2.5 *A metric space X is* geodesic *if between every two points of X there exists a geodesic.*

Thus a geodesic metric space is a length space in which infima of distances of rectifiable paths are attained. Such a space clearly has to be path connected.

Example 2.6 1. Complete Riemannian manifolds are geodesic metric spaces, by the Hopf-Rinow theorem (every closed bounded subset of a complete manifold is compact).

2. A connected graph is a geodesic metric space. We give each edge length 1 and as above the distance between two vertices v_1 and v_2 is the least number of edges of a path between v_1 and v_2.

3. More generally, if K is an n-dimensional simplicial compex and we metrise each n-simplex of K as a regular Euclidean simplex of side 1, giving the whole complex the metric of infima of path lengths, then K is a geodesic metric space.

2.3 Slim and Thin Triangles.

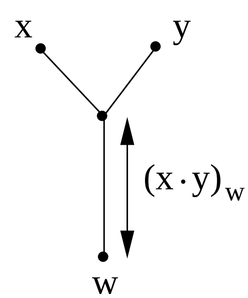
In this section we give various versions of Gromov's hyperbolicity criterion and show that they are all equivalent.

Definition 2.7 *Suppose that* (X, d) *is a metric space with basepoint* w*. Then we define the* Gromov product *on* $X \times X$ *based at* w *by*

$$(x \cdot y)_w = \frac{1}{2}\left(d(x,w) + d(y,w) - d(x,y)\right).$$

This is sometimes called an inner product. (It is not an inner product in the usual sense as it is not necessarily defined on a vector space).

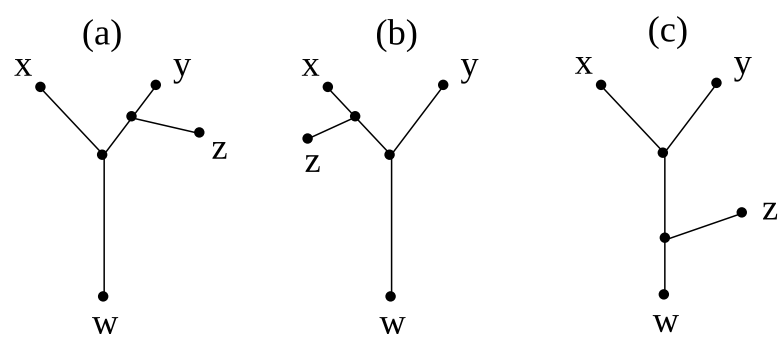
Example 2.8 Suppose x,y and w are points in a tree T, let γ be a geodesic from w to x and let η be a geodesic in T from w to y. Then $(x \cdot y)_w$ is the distance along γ (or η) before the two geodesics separate.



Choose any basepoint w in T. Then the metric d on T satisfies the property that for all x,y and z in X we have

$$(x \cdot z)_w \geqslant \min\{(x \cdot y)_w, (y \cdot z)_w\}.$$

There are three cases as in the following picture.



In case (a) we have

$$(x \cdot z)_w = (x \cdot y)_w = \min\{(x \cdot y)_w, (y \cdot z)_w\},$$

in (b) we have

$$(x \cdot z)_w > (x \cdot y)_w = (y \cdot z)_w$$

and in (c) we have

$$(x \cdot z)_w = (y \cdot z)_w = \min\{(x \cdot y)_w, (y \cdot z)_w\}.$$

If we relax the property satisfied by the metric on a tree in example 2.8 then we obtain the following.

Definition 2.9 *If (X, d) is a metric space with basepoint w and there exists $\delta \geqslant 0$ such that for all x,y and z in X we have*

$$(x \cdot z)_w \geqslant \min\{(x \cdot y)_w, (y \cdot z)_w\} - \delta$$

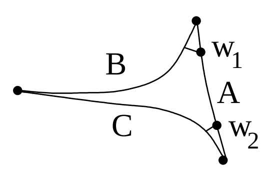
*then we say that the Gromov product based at w is δ-*hyperbolic*. If there exists $\delta \geqslant 0$ such that the Gromov product is δ-hyperbolic, we just say that the Gromov product based at w is* hyperbolic.

We next consider *geodesic triangles* in a geodesic metric space, i.e. triples of points with a specified geodesic (called a *side*) between every pair of points. If (x, y, z) is such a triangle then we denote the geodesic between, say, x and y by $[x, y]$.

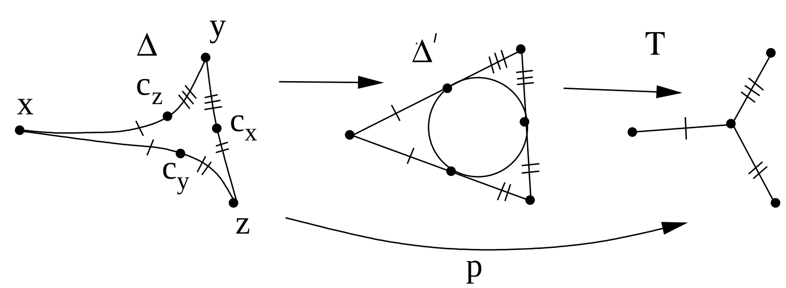
Definition 2.10 *Let (X, d) be a geodesic metric space and let Δ be a geodesic triangle. We say that Δ is δ-*slim *if for each ordering (A, B, C) of its sides, and for any point $w \in A$, we have*

$$\min\{d(w, B), d(w, C)\} \leqslant \delta.$$

For example, suppose that in the following picture we have a δ-slim geodesic triangle. We might have $d(w_1, B) \leqslant \delta$ but $d(w_1, C) > \delta$ and $d(w_2, C) \leqslant \delta$ but $d(w_2, B) > \delta$.



Definition 2.11 *Let X be a geodesic metric space. Given a geodesic triangle Δ in X, we consider the Euclidean triangle Δ' with the same side lengths. Collapse Δ' to a tripod T. Let $p : \Delta \to T$ be the map which so arises. We say that Δ is δ-*thin *if for all $t \in T$ we have diam$(p^{-1}(t)) \leqslant \delta$.*



Note that diam$(p^{-1}(t))$ can be either one, two or three points.

Proposition 2.12 *Let X be a geodesic metric space. Then the following are equivalent.*

1. *There exists a point $w \in X$ such that the Gromov product based at w is hyperbolic.*
2. *For all points $w \in X$ the Gromov product based at w is hyperbolic.*

3. *There exists $\delta \geqslant 0$ such that every geodesic triangle in X is δ-slim.*

4. *There exists $\delta \geqslant 0$ such that every geodesic triangle in X is δ-thin.*

Proof. Clearly the second statement implies the first. We now show that the first statement implies the second. Suppose that the Gromov product based at w in X is δ-hyperbolic, and let t be another point of X. We show that the Gromov product in X based at t is 2δ-hyperbolic. For ease of notation, we write for x and y in X, xy instead of $(x \cdot y)_w$. We now claim that for all x, y and z in X,

$$xy + zt \geqslant \min\{xz + yt, xt + yz\} - 2\delta.$$

By definition of δ-hyperbolicity of the Gromov product based at w we have

$$\begin{aligned} xy + zt &\geqslant \min\{xt, ty\} - \delta + zt \\ &= \min\{xt + zt, ty + zt\} - \delta \\ &\geqslant \min\{xt + \min\{zy, yt\}, ty + \min\{zx, xt\}\} - 2\delta \\ &= \min\{xt + zy, xt + yt, yt + zx\} - 2\delta. \end{aligned}$$

And by a symmetric argument,

$$xy + zt \geqslant \min\{xz + zy, xz + yt, zy + zt\} - 2\delta.$$

So the claim is proved unless the minimum in the first of these expressions is $xt + yt$ and that in the second is $xz + yz$. This can easily be seen to lead to a contradiction.

Secondly we claim that for all x, y and z in X,

$$d(x, y) + d(y, t) \leqslant \max\{d(x, z) + d(y, t), d(x, t) + d(y, z)\} + 4\delta.$$

This follows from the first claim by substituting the definition of the Gromov product and multiplying both sides by -1. Then we can use the identity $-\min\{a, b\} = \max\{-a, -b\}$.

We now prove that X is 2δ-hyperbolic with respect to t, i.e that

$$(x \cdot y)_t \geqslant \min\{(x \cdot z)_t, (y \cdot z)_t\} - 2\delta.$$

This is true if and only if

$$\begin{aligned} &\frac{1}{2}(\min\{d(x, t) + d(z, t) - d(x, z), d(y, t) + d(z, t) - d(y, z)\} \\ &-d(x, t) - d(y, t) + d(x, y)) \leqslant 2\delta. \end{aligned}$$

i.e., regrouping terms, if and only if

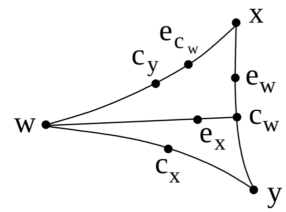
$$\min\{-d(y, t) - d(x, z), -d(x, t) - d(y, z)\} \leqslant -d(x, y) - d(z, t) + 4\delta.$$

But this is true by the second claim above. $\boxed{(1) \Leftrightarrow (2)}$

We now show that the second statement implies the third. More specifically, we show that if for all points w of X the Gromov product based at X is δ-hyperbolic then triangles in X are 3δ-slim. Let $[x, y, w]$ be a geodesic triangle in X and suppose that the Gromov product with respect to w is δ-hyperbolic. Then we claim that

$$d(w, [x, y]) \leqslant (x \cdot y)_w + 2\delta. \tag{1}$$

Let c_x,c_y and c_w be the internal points of $[x, y, w]$ and let e_{c_w},e_w and e_x be those of $[w, x, c_w]$.



Without loss of generality, suppose that $(x \cdot c_w)_w \leqslant (y \cdot c_w)_w$, so that, since the Gromov product based at X is δ-hyperbolic, we must have

$$(x \cdot y)_w \geqslant (x \cdot c_w)_w - \delta. \tag{2}$$

Now $(x \cdot c_w)_w = d(w, c_y)$, which by the last statement is at least $(x \cdot c_w) - \delta$. Hence as $d(w, c_y) + d(c_y, e_{c_w}) = d(w, e_{c_w})$ we have $d(e_{c_w}, c_y) \leqslant \delta$ and since $d(x, c_w) = d(x, c_y)$ and $d(x, e_w) = d(x, c_y)$ we also have $d(c_w, e_w) \leqslant \delta$. Thus

$$d(w, c_w) = d(w, e_x) + d(e_x, c_w) = d(w, e_{c_w}) + d(e_x, c_w) \leqslant (x \cdot y)_w + 2\delta,$$

which proves (1).

Note that we can assume that a general geodesic triangle $[x, y, z]$ in X has one side through w by the assumption that the Gromov product is hyperbolic with respect to all basepoints. Say that w is a general point on $[x, y]$. By δ-hyperbolicity of the Gromov product with respect to w we have

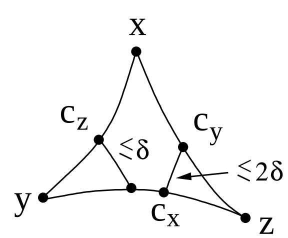
$$(x \cdot y)_w \geqslant \min\{(x \cdot y)_w, (y \cdot z)_w\} - \delta.$$

Assume without loss of generality that $(x \cdot y)_w \leqslant (y \cdot z)_w$. Then as $(x \cdot y)_w = 0$, because w is on a geodesic between x and y, we have $(x \cdot z)_w \leqslant \delta$. Thus by the claim above we have

$$d(w, [x, z]) \leqslant (x \cdot z)_w + 2\delta \leqslant 3\delta,$$

completing the second part of the proof. $\boxed{(2) \Rightarrow (3)}$

To show that the third statement implies the fourth we show that if geodesic triangles are δ-slim then they are 6δ-thin. Suppose $\Delta = [x, y, z]$ is a geodesic triangle in X. Then X is δ-slim. Let c_x, c_y and c_z be the internal points of Δ. First we claim that $\mathrm{diam}(\{c_x, c_y, c_z\}) \leqslant 4\delta$. Since Δ is δ-slim there exists a $t \in [y, z] \cup [x, z]$ such that $d(c_z, t) \leqslant \delta$.



Now we have

$$d(y, t) \geqslant d(y, c_z) \geqslant d(y, c_z) - \delta.$$

Combining this with the triangle inequality for $[y, t, c_z]$ we have

$$d(y, c_z) - \delta \leqslant d(y, t) \leqslant d(y, c_z) + \delta,$$

And by symmetry,

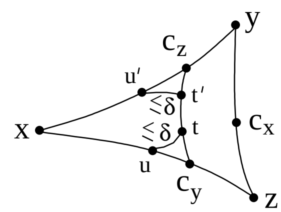
$$d(y, c_x) - \delta \leqslant d(y, t) \leqslant d(y, c_x) + \delta.$$

So these inequalities give

$$d(y, c_x) - d(t, c_x) = d(y, t) \geqslant d(y, c_x) - \delta \Rightarrow d(t, c_x) \leqslant \delta.$$

Thus $d(c_x, c_z) \leqslant d(c_x, t) + d(t, c_z) \leqslant 2\delta$, and similarly we obtain $d(c_y, c_z) \leqslant 2\delta$ and $d(c_y, c_x) \leqslant 2\delta$ from which the claim follows.

We now show that δ-slim implies 6δ-thin. Let $u \in [x, c_y]$ and let $u' \in [x, y]$ with $d(u, x) = d(u', x)$. Choose any geodesic $[c_y, c_z]$.



The triangle $[x, c_y, c_z]$ is δ-slim so there is a $t \in [x, c_z] \cup [c_y, c_z]$ with $d(u, t) \leqslant \delta$. If $t \in [x, c_z]$ we see as in the proof of the above claim that $d(t, u') \leqslant \delta$ so that $d(u, u') \leqslant 2\delta$. Otherwise $t \in [c_y, c_z]$. Using the same reasoning for u' we conclude that there is a $t' \in [c_y, c_z]$ with $d(u', t') \leqslant \delta$. Then by the claim,

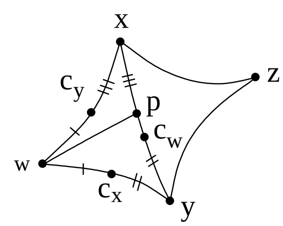
$$d(u, u') \leqslant d(u', t) + d(t, t') + d(u, t') = 6\delta.$$

Thus Δ is 6δ-thin. $\boxed{(3) \Rightarrow (4)}$

We now show that the fourth statement implies the second. Specifically, we show that if triangles in X are δ-thin and $w \in X$ is a basepoint then the inner product based at w is 2δ-hyperbolic. We claim that

$$(x \cdot y)_w \leqslant d(w, [x, y]).$$

To see this, let $p \in [x, y]$ be such that $d(w, [x, y]) = d(w, p)$.



Then we have

$$\begin{aligned}
(x \cdot y)_w &= \frac{1}{2}\left(d(x, w) + d(y, w) - d(x, y)\right) \\
&= \frac{1}{2}\left(d(x, w) + d(y, w) - d(x, p) - d(p, y)\right) \\
&= \frac{1}{2}\left((d(x, w) - d(x, p)) + (d(y, w) - d(p, y))\right) \\
&= \frac{1}{2}\left(d(w, p) + d(w, p)\right) = d(w, [x, y]),
\end{aligned}$$

Which proves the claim.

Now let q be the map which collapses $[w, x, y]$ to a tripod. Let c be the fork of this tripod. Then $d(c, w) = (x \cdot y)_w$ and we have

$$d(w, c_w) \leqslant (x \cdot y)_w + \operatorname{diam}\{q^{-1}(c)\}.$$

Since the triangle $[w, c_w, p]$ is δ-thin and hence (obviously) δ-slim, we have

$$(x \cdot y)_w + 2\delta \geqslant d(w, c_w) + \delta \geqslant \min\{d(w, [x, y]), d(w, [y, z])\}.$$

By the claim, $d(w,[x,y]) \geqslant (x\cdot y)_w$ and by symmetry, $d(w,[y,z]) \geqslant (y\cdot z)_w$. Hence we have

$$(x\cdot y)_w + 2\delta \geqslant \min\{(x\cdot z)_w, (y\cdot z)_w\}$$

and the inner product based at w is 2δ-hyperbolic as required. □

Definition 2.13 *If a geodesic metric space satisfies the properties of the last proposition then we call it a* hyperbolic metric space*.*

Exercises 2.14 1. Show that a geodesic metric space is proper if and only if it is locally compact and complete. Show that this fails when the space is not geodesic (see [19]).

2. Let X be a δ-hyperbolic metric space. Define what it means for a geodesic polygon to be ε-thin. It is easy to see that for $n \geqslant 3$, every n-sided geodesic polygon in X is $4(n-2)\delta$-thin. Show that in fact there exists a function f, which is asymptotic to $\log(n)$, such that an n-sided geodesic polygon in X is $f(n)\delta$-thin (see [13] Chapter 3, Lemma 5).

The most important example of a hyperbolic metric space is the hyperbolic plane $\mathbb{H}^2$. This is modelled by the upper half plane (complex numbers with strictly positive imaginary part) with the Riemannian metric $\mathrm{d}s = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2}$. The *hyperbolic length* of a path $\gamma = \{z(t) = x(t) + iy(t) | t \in [0,1]\}$ is given by

$$h(\gamma) = \int_0^1 \left[\frac{1}{y(t)}\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\right] \mathrm{d}t.$$

$\mathbb{H}^2$ is then made into a metric space via the *hyperbolic distance*

$$\rho(z,w) = \inf_{\{\gamma|\gamma(0)=z,\gamma(1)=w\}} [h(\gamma)]\,.$$

The geodesics in $\mathbb{H}^2$ are circles or lines which intersect $\mathbb{R}$ at right angles. Given distinct points x_1 and x_2 in $\mathbb{H}^2$ there exists a unique geodesic passing through both x_1 and x_2. However, $\mathbb{H}^2$ is *non-euclidean*, which means that given a geodesic γ and a point $z \in \mathbb{H}^2 - \gamma$, there exist more than one geodesic through z and parallel to γ.

Exercise 2.15 Show that hyperbolic distance is given explicitly by the formula

$$\rho(z,w) = \log\left(\frac{|z-\bar{w}| + |z-w|}{|z-\bar{w}| - |z-w|}\right).$$

and that $\mathbb{H}^2$ is a $\log(3)$-hyperbolic metric space

We also have the disc model of the hyperbolic plane, due to Poincaré. This is the unit disc of the complex plane with the Riemannian metric

$$\mathrm{d}s = \frac{2|\mathrm{d}z|}{1-|z|^2}.$$

The geodesics in this model are circles or lines making an angle of 90° with the unit circle and the isometries are transformations of the form

$$\Phi(z) = \frac{a\bar{z}+b}{c\bar{z}+d} \text{ with } ad - bc = 1.$$

If U denotes the upper half plane model of the hyperbolic plane and H the Poincaré disc model, then we have an isometry $U \to D$ given by

$$f(z) = \frac{iz+1}{z+i}.$$

which gives the following formula for hyperbolic distance in the disc model.

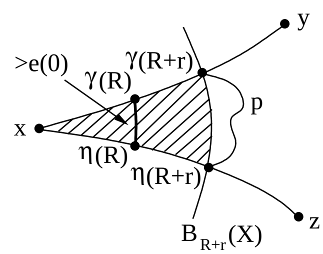
$$\rho(z,w) = \log\left(\frac{|1-z\bar{w}| + |z-w|}{|1-z\bar{w}| - |z-w|}\right).$$

2.4 Divergence Of Geodesics

In this section we prove a useful characterisation of hyperbolicity. Roughly speaking, this is that geodesics in a hyperbolic metric space either stay within a bounded distance of each other or eventually move an exponential distance away from each other. To formulate what is precisely meant by this is somewhat technical.

If X is a metric space and Y is a subset of X then we denote by $B_R(Y)$ the ball of radius R around Y.

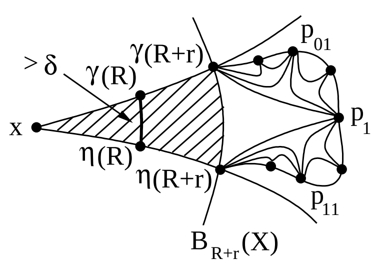
Definition 2.16 *Let X be a geodesic metric space. A function $e : \mathbb{N} \to \mathbb{R}$ is a* divergence function *for X if for all $x \in X$, for all geodesics $\gamma = [x, y]$ and $\eta = [x, z]$ in X to points y and z in X, for all r and R in $\mathbb{R}$ with $R + r \leqslant \min\{d(x, y), d(x, z)\}$ and $d(\gamma(R), \eta(R)) > e(0)$ and for all paths p in $X - B_{R+r}(x)$ between $\gamma(R + r)$ and $\eta(R + r)$ we have $l(p) > e(r)$.*



If there exists a divergence function for X then we say that geodesics *diverge* in X. If there exists an exponential divergence function for X then we say that geodesics *diverge exponentially* in X.

Theorem 2.17 *In a hyperbolic metric space geodesics diverge exponentially.*

Proof. Suppose that triangles are δ-thin in X. We are going to construct an exponential divergence function e. Take $e(0) = \delta$, suppose that γ and η are two geodesics issuing from x and let R be such that $d(\gamma(R), \eta(R)) \geqslant \delta$. Let p be a path in $X - B_{R+r}(x)$ joining $\gamma(R + r)$ to $\eta(R + r)$.



Then we claim that

$$d(\gamma(R), p) < \delta\,(\log_2(l(p)) + 2)\,.$$

To see why this is so, divide p into smaller paths as follows. Let p_1 be the midpoint of p, let p_{01} be the midpoint of the first half of p and let p_{11} be the midpoint of the second half of p. Continue dividing in half in this fashion. After we have subdivided $\log_2(|p|)$ times each resulting segment has length less than 2. By thinness of the triangles $[\gamma(R + r), \eta(R + r), p_1]$, $[\gamma(R + r), p_{01}, p_1]$, $[\eta(R + r), p_1, p_{11}] \ldots$ etc, the claim is proved.

Now if $d(\gamma(R), p) \geqslant r$ then $\delta\,(\log_2(l(p)) + 2) > r$ and we have

$$l(p) > 2^{\left(\frac{r-2}{\delta}\right)}.$$

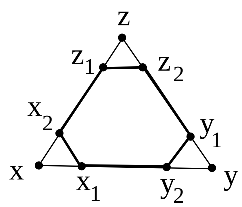
Thus we can take $e(r)$ to be the term on the right hand side, which is exponential. □

We say that geodesics diverge *faster than linearly* in a geodesic metric space X if it has a divergence function e such that

$$\lim \left(\frac{e(r)}{r}\right) = \infty.$$

Theorem 2.18 *If geodesics diverge faster than linearly in a geodesic metric space then it is hyperbolic.*

Proof. Let e be a divergence function for the geodesic metric space X making geodesics diverge faster than linearly. We show that there exists $\delta > 0$ such that triangles in X are δ-slim. Let $[x, y, z]$ be a geodesic triangle in X. We want a bound for $d(w, [x, z] \cup [y, z])$, independent of $[x, y, z]$, where w is a point of $[x, y]$. Let T be maximal such that for all $t \leqslant T$, $d(\alpha_1(t), \alpha_2(t)) \leqslant e(0)$, where α_1 and α_2 are isometric embeddings of the real intervals I_1 and I_2. Let $x_1 = \alpha_1(T)$ and $x_2 = \alpha_2(T)$. Similarly define z_1,z_2,y_1 and y_2. We claim that if $[x, x_1]$ meets $[y_2, y]$ then we can bound $\max\{d(z_2, y_1), d(z_1, x_2)\}$.

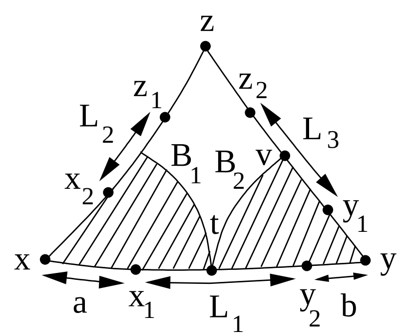


To see this, suppose the segments meet. Then there exist points x_3 of $[x, x_2]$ and y_3 of $[y, y_1]$ such that $d(x_3, y_3) < 2e(0)$. By definition of a divergence function, $[z_1, x_3]$ and $[z_2, y_3]$ are of bounded length, and hence so are $[z_1, x_2]$ and $[z_2, y_1]$. This proves the claim.

So we may assume that no such intersections exist. Let L_1,L_2 and L_3 be the lengths of $[y_2, x_1]$, $[x_2, z_1]$ and $[z_2, y_1]$ respectively. Assume without loss of generality that $L_1 \geqslant L_2 \geqslant L_3$.

We will show that L_1 is bounded, by some constant K. Then $[x, y, z]$ will be $(\frac{K}{2} + e(0))$-thin.

Let t be the midpoint of $[x_1, y_2]$, let $a = d(x, x_1)$ and let $b = d(y, y_1)$. Then if we define B_1 to be $B_{a+\frac{L_1}{2}}(x)$ and B_2 to be $B_{b+\frac{L_1}{2}}(y)$, we have $t \in B_1 \cup B_2$.



We now claim that $[x_2, z]$ does not meet $\mathrm{Int}(B_2)$. Suppose that there exists $s \in [x_2, z] \cap \mathrm{Int}(B_2)$. Then $s \notin B_1$ so it follows that $d(s, x_2) \geqslant \frac{L_1}{2}$. As $L_3 \geqslant L_2$, there exists $u \in [y, z]$ such that $d(u, z) = d(s, z)$. Thus

$$\begin{aligned} \frac{L_1}{2} &\leqslant d(s, x_2) \\ &= d(x_2, z) - d(z, s) \\ &= d(x_2, z_1) + d(z_1, z) - d(z, s) \\ &\leqslant d(z_2, y_1) + d(z_1, z) - d(z, u) \\ &= d(z, y_1) - d(z, u) \\ &= d(u, y_1) \end{aligned}$$

So $u \notin \mathrm{Int}(B_2)$ but $d(z, y) = d(z, u) + d(u, y)$, which is at most $d(z, s) + d(s, y)$. Hence

$$b + \frac{L_1}{2} \leqslant d(u, y) \leqslant d(s, y) \leqslant \frac{L_1}{2} + b,$$

which is a contradition, establishing the second claim.

To complete the proof, we let v be a point on $[y,z]$ such that $d(y,v)=b+\frac{L_1}{2}$. There is a path from t to v in the complement of B_2 of length at most

$$d(t,x_1)+3e(0)+L_3+e(0)+d(z_2,v)\leqslant \frac{L_1}{2}+6e(0)+L_1+\frac{L_1}{2}.$$

Hence $e(\frac{L_1}{2})\leqslant 2L_1+6e(0)$, giving the required bound for L_1. □

The following proof is from [40],[42].

Theorem 2.19 *If geodesics diverge in a geodesic metric space then they diverge exponentially.*

Proof. Let X be a geodesic metric space and f a divergence function for X such that $f(r)\to\infty$ as $r\to\infty$. We define a new divergence function e for X as follows. Let e(r) be the infimum of $l(\alpha)$ over all paths α from $\gamma(R+r)$ to $\eta(R+r)$ in $\overline{X-B_x(R+r)}$ with $d(\gamma(R),\eta(R))\geqslant f(0)$, where γ and η are geodesics with $\gamma(0)=\eta(0)=x$ (we take the infimum over all γ,η,$x\in X$ and $R\in\mathbb{R}$). Note that since $e(r)\geqslant f(r)$ for all r, we also have $e(r)\to\infty$ as $r\to\infty$. Let $N'=\sup\{r|e(r)<9e(0)\}$, let $N=N'+1+3e(0)$ and let $u=\sup\{t|e(t)<4N+2\}$. We show that if $r>u+N$ then $e(r)>\frac{3}{2}e(r-N)$, which implies that e is an exponential divergence function. Let γ and η be geodesics such that $\gamma(0)=\eta(0)$ and $d(\gamma(R),\eta(R))>e(0)$. Then we claim that $d(\gamma(R+N),\eta(R+N))\geqslant 3(e(0))$.
The proof of this claim is by contradiction. Suppose that $d(\gamma(R+N),\eta(R+N))<3e(0)$ and let β' be geodesic joining $\gamma(R+N)$ and $\eta(R+N)$. Consider the arc

$$\beta=[\gamma(R+N-3e(0)),\gamma(R+N)]\cup\beta'\cup[\eta(R+N),\eta(R+N-3e(0))].$$

It is clear that β lies in $\overline{X-B_x(R+N-3e(0))}$. Also, $l(\beta)=6e(0)+l(\beta')<9e(0)$. But $l(\beta)>e(N-3e(0))\geqslant 9e(0)$ since $N-3e(0)=N'+1>N'$, the desired contradiction.
Now let $r>u+N$ and let γ and η be geodesics such that $\gamma(0)=\eta(0)=x$, $d(\gamma(R),\eta(R))>e(0)$ and let α be an arc in $\overline{X-B_x(R+r)}$ such that $\alpha(0)=\gamma(B+r)$ and $\alpha(M)=\eta(R+r)$ where $M=l(\alpha)$ and is less than $e(r)+1$. Put

$$\begin{aligned} t_1 &= \sup\left\{t|\alpha(t)\in B_x(R+r+N), t\leqslant\frac{M}{2}\right\} \text{ and}\\ t_2 &= \sup\left\{t|\alpha(t)\in B_x(R+r+N), t\geqslant\frac{M}{2}\right\}.\end{aligned}$$

If we now let c_1 and c_2 be geodesics from x to $\alpha(t_1)$ and $\alpha(t_2)$ respectively then we have

$$\begin{aligned} d(\gamma(R+N),c_1(R+N)) &+ d(c_1(R+N),c_2(R+N))\\ +d(c_2(R+N),\eta(R+N)) &\geqslant d(\gamma(R+N),\eta(R+N))\\ &\geqslant 3e(0)\end{aligned}$$

One of these three summands is hence greater than e(0).
Suppose that $d(c_1(R+N),c_2(R+N))\geqslant 0$. Then

$$l(\alpha|_{[t_1,t_2]})\leqslant l(\alpha)-2N\leqslant e(r)+1-2N<e(r).$$

So we have a path joining $\alpha(t_1)$ and $\alpha(t_2)$ contained in $\overline{X-B_x(R+r+N)}$ of length less than $e(r)$, which is impossible.
We may assume without loss of generality that $d(\gamma(R+N),c_1(R+N))\geqslant e(0)$. Considering the path $\alpha'=\alpha|_{[0,t_1]}\cup[c_1(R+r),\alpha(t_1)]$ we obtain

$$l(\alpha')\leqslant\frac{M}{2}+N\leqslant\frac{e(r)+1}{2}+N.$$

On the other hand, $l(\alpha) \geqslant e(r-N)$. So $e(r-N) \leqslant \frac{1}{2}e(r) + N + \frac{1}{2}$ which implies that $e(r) \geqslant 2(e(r-N) - N - \frac{1}{2})$. But $e(r-N) \geqslant 4N+2$ which gives $e(r-N) - N - \frac{1}{2} \geqslant \frac{3}{4}e(r-N)$, and thus $e(r) \geqslant \frac{3}{2}e(r-N)$ as required. □

Corollary 2.20 *If geodesics diverge in a geodesic metric space then it is hyperbolic.*

In [42], Papasoglu uses this theorem to show that one can also characterise hyperbolicity of a metric space in terms of thinness of geodesic bigons. This implies that strongly geodesically automatic groups are hyperbolic (see section 5). (One can also characterise hyperbolicity in terms of thinness of quasi-geodesic bigons [45].)

Example 2.21 Why don't geodesics in $\mathbb{R}^2$ diverge? Let γ and η be two lines in $\mathbb{R}^2$ which are at an angle θ to each other. Then (if we assume the lines to be parametrised by length) we have

$$d(\gamma(R+r), \eta(R+r)) = 2(R+r)\sin(\frac{\theta}{2})$$

Now if $f(r) = kr + c$, we have for all γ and η, $k < 2\sin\left(\frac{\theta}{2}\right)$. So $k = 0$. (In general shows f to be bounded.)

2.5 Quasi-Geodesics In Hyperbolic Metric Spaces.

In this section we introduce *quasi-geodesics*. These are paths which approximately behave like geodesics, in that they are not allowed to stay in the same place for too long a time.

If X is a metric space and p is a path in X then we write $i(p)$ for the initial point of p and $t(p)$ for the terminal point of p.

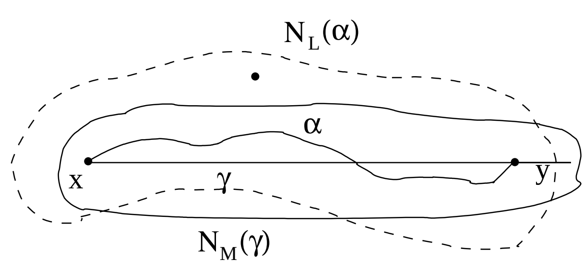
Definition 2.22 *Let (X,d) be a metric space and let x_1 and x_2 be two points in the same path-component of X. Let $\lambda \geqslant 1$ and $\mu \geqslant 0$. Then a* (λ,μ)-quasi-geodesic *from x_1 to x_2 is a rectifiable path p from x_1 to x_2 such that for all subpaths q of p we have*

$$l(q) \leqslant \lambda d(i(q), t(q)) + \mu.$$

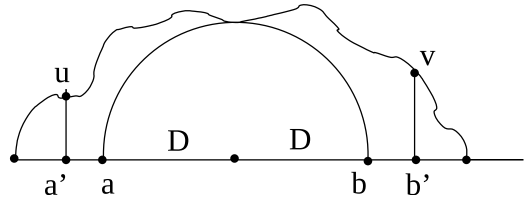
Example 2.23 Let p be the infinite path in $\mathbb{C}$ given by $p(t) = e^{it}$. Then for all λ and μ we can choose $x \in \mathbb{R}$ such that $p|_{[0,x]}$ is not a (λ,μ)-quasi-geodesic.

The most important property of quasi-geodesics in hyperbolic metric spaces is the following theorem, which can be summed up by saying that in a hyperbolic metric space, quasi-geodesics are "close" to geodesics.

Theorem 2.24 *Let x_1 and y_1 be points of a hyperbolic metric space X. If α is a (λ,μ)-quasi-geodesic between x and y then there exist constants $L(\lambda,\mu)$ and $M(\lambda,\mu)$ such that if γ is a geodesic between x and y then $\gamma \subset B_L(\alpha)$ and $\alpha \subset B_M(\gamma)$.*



Proof. First we find L. Let e be an exponential divergence function for X. Let $D = \sup_{x\in\gamma}\{d(x,\alpha)\}$ and let $p \in \gamma$ be such that $d(p,\alpha) = D$. Then $\text{Int}(B_D(P))$ does not meet α. Let a and b be the two points on α at a distance D from p, let a' and b' be points on α at a distance $2D$ from p and let u and v be points on α such that $d(a',u) \leqslant D$ and $d(b',v) \leqslant D$.

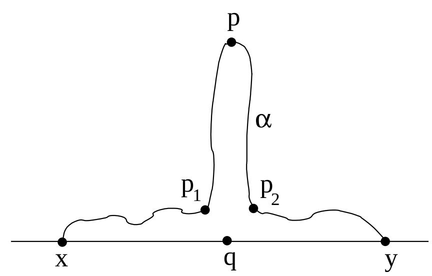


By following a path via a',a,b and b', we see that $d(u,v) \leqslant 6\delta$. Since α is a (λ,μ)-quasi-geodesic, we see that $d_\alpha(u,v) \leqslant 6\lambda D + \mu$. Thus there is an arc connecting a to b lying outside $\mathrm{Int}(B_D(P))$ of length less than $4D + 4D\lambda + \mu$. By the definition of a divergence function we must have

$$e\left(D - \frac{e(0)}{2}\right) < 4D + 4D\lambda + \mu.$$

But e is an exponential function, so it can't be bounded by a linear function. D must hence be bounded, as must L.

Now we must exhibit a M such that $\alpha \subset B_M(L)$. Suppose that α is not contained in $B_L(\gamma)$.



Let p be a point on α such that $d(p,\alpha) > L$. For every point $\gamma(t)$ on γ there exists a point $\alpha(t')$ on α such that $d(\gamma(t), \alpha(t')) \leqslant L$. $\alpha(t)$ lies either before of after p. By continuity, there exists a point $q \in \gamma$ and points p_1 and p_2 on α such that $d(q,p_1) \leqslant L$, $d(q,p_2) \leqslant L$ and p lies between p_1 and p_2 (i.e. if we choose a parametrisation of the arc, the coordinate t of p satisfies $t_1 \leqslant t \leqslant t_2$). Then $d(p_1,p_2) \leqslant 2L$ and hence $d_\alpha(p_1,p_2) \leqslant 2L\lambda + \mu$. We thus have

$$d(p,\gamma) \leqslant L\lambda + \frac{\mu}{2} + L,$$

and can take $M = L\lambda + \frac{\mu}{2} + L$. $\square$

Exercise 2.25 Let X and Y be metric spaces. Show that if $f : X \to Y$ is a (λ, C)-quasi-isometry and p is a (μ,ε)-quasi-geodesic in X then $f(p)$ is a $(\lambda\mu, \lambda\varepsilon + C)$-quasi-geodesic in Y.

Corollary 2.26 *Hyperbolicity of metric spaces is a quasi-isometry invariant.*

Exercise 2.27 A path p is called a K-*local geodesic* if for all x and y in p, $d_p(x,y) \leqslant K$ implies that $d_p(x,y) = d(x,y)$.

1. For all K, give an example of an infinite K-local geodesic in $\mathbb{Z}\oplus\mathbb{Z}$ which is not a quasigeodesic.
2. Let X be a geodesic metric space. Assume triangles in X are δ-thin. Show that if an path p in X is an 8δ-local geodesic then it is a $(2,0)$-quasigeodesic.
3. Show that if every infinite local geodesic in a metric space X is a quasigeodesic then X is hyperbolic.

2.6 The Boundary of a Hyperbolic Metric Space.

In this section we construct a natural compactification of a hyperbolic metric space.

Definition 2.28 *Let X be a geodesic metric space. we say that a sequence (x_n) in X* converges to infinity *if*

$$\lim_{m,n\to\infty}(x_m \cdot x_n) = \infty.$$

Note that this definition is independent of the choice of basepoint, since if w and w' are two basepoints, then we have for all x and y in X,

$$|(x \cdot y)_w - (x \cdot y)_{w'}| \leqslant d(w, w').$$

Note also that if (x_n) converges to infinity, then since

$$(x_m \cdot x_n) \leqslant \min\{d(x_m, w), d(x_n, w)\},$$

we have $d(x_n, w) \to \infty$ as $n \to \infty$.

Exercise 2.29 If a sequence converges to infinity then so do all its subsequences.

We denote by $S_\infty(X)$ the set of all sequences in X converging to infinity, and define a relation on $S_\infty(X)$ via

$$(x_n) \sim (y_n) \text{ if and only if } \lim_{n\to\infty}(x_n \cdot y_n) = \infty.$$

Proposition 2.30 *If X is a hyperbolic metric space then this is an equivalence relation.*

Proof. Suppose that X is δ-hyperbolic. Symmetry and reflexivity are obvious. Transitivity follows since if (x_n), (y_n) and (z_n) are sequences with $(x_n) \sim (y_n)$ and $(y_n) \sim (z_n)$, then $(x_n) \sim (z_n)$ since

$$(x_n \cdot z_n) \geqslant \min\{(x_n \cdot y_n), (y_n \cdot z_n)\} - \delta.$$

□

Example 2.31 In $\mathbb{Z} \oplus \mathbb{Z}$ the relation is not transitive, for let x and y be generators and let the basepoint be the identity. Define the sequences $a_n = x^n$, $b_n = x^n y^n$ and $c_n = y^n$, all of which converge to infinity. Then $(a_n \cdot b_n) = n$ and $(b_n \cdot c_n) = n$ so $a_n \sim b_n$ and $b_n \sim c_n$. However $(a_n \cdot c_n) = 0$ so a_n is not equivalent to c_n.

Definition 2.32 *The* boundary *∂X of a hyperbolic metric space X is the quotient set $S_\infty(X)/\sim$. The* compactification *of X, written $\hat{X}$, is defined to be $X \cup \partial X$.*

We say that $\{a_n\} \in S_\infty(X)$ *converges* to $x \in \partial X$ if $x = [\{a_n\}]$ and write $a_n \to X$.

Example 2.33 The real line is compactified in this way by adding two points, $+\infty$ and $-\infty$, since if $\{a_n\} \in S_\infty(\mathbb{R})$ then either $a_n > 0$ for almost all n or $a_n < 0$ for almost all n. More generally, if T is a tree then we add a point for each geodesic ray in T. We consider two points of this boundary to be close if the rays coincide for a long way before branching. This can be used to define a topology on the boundary of the tree with respect to which the boundary is a Cantor set.

We extend the Gromov product on X to ∂X. If x and y are both in $\hat{X}$, then we define

$$(x \cdot y) = \inf_{x_i \to x, y_i \to y}(\liminf(x_i \cdot y_i)).$$

Example 2.34 Why do we need to use $\liminf$ in this definition? Consider $\mathbb{Z}\oplus\mathbb{Z}_2$, whose boundary consists of two points, $+\infty$ and $-\infty$. Let $a_n = x^n$, $b_n = x^{-n}$, $c_n = yx^n$, $d_n = yx^{-n}$ and let $z_n = x^n$ if n is even and yx^n if n is odd. Then a_n, c_n and z_n tend to $+\infty$ and b_n and d_n tend to $-\infty$. But $(d_n \cdot z_n)$ is 0 if n is even and 1 if n is odd. So the ordinary limit does not exist.

Why do we need to take the infimum over all sequences? In the above, $(a_n \cdot b_n) = 0$ but $(c_n \cdot d_n) = 1$. However, (a_n, b_n) and (c_n, d_n) are representatives of the same pair of boundary points.

Proposition 2.35 (Elementary Properties.) *Let X be a δ-hyperbolic metric space.*

1. *If $x \in X$ and $y \in \hat{X}$ then*
$$(x \cdot y) = \inf_{y_i \to y}\{\liminf_i (x \cdot y_i)\}.$$
2. *The Gromov product restricted to X agrees with the one already defined.*
3. *If x and y are in $\hat{X}$ then $(x \cdot y)$ is infinite if and only if x and y are in ∂X and $x = y$.*
4. *If $x \in \partial X$ and $\{x_i\}$ is any sequence of points in X then $(x_i \cdot x) \to \infty$ iff $\{x_i\} \in S_\infty(X)$ and $x_i \to x$.*
5. *If x and y are in $\hat{X}$ then there exist sequences $x_i \to x$ and $y_i \to y$ with $\lim(x_i \cdot y_i) = (x \cdot y)$. Moreover, if x or y lies in X then the corresponding sequence can be chosen to be constant.*
6. *If x and y are in ∂X, $x_i \to x$ and $y_i \to y$ then*
$$(x \cdot y) \leqslant \liminf_i (x_i \cdot y_i) \leqslant (x \cdot y) + 2\delta.$$
7. *$\hat{X}$ is δ-hyperbolic.*
8. *If x and y are in $\hat{X}$ and $y_i \to y$ then*
$$\liminf_i (x \cdot y_i) \geqslant (x \cdot y)$$
9. *Suppose that $\{x_n\}$ converges to infinity and $(x_n \cdot y_n) \leqslant K$ for all n. Then there exists $N \in \mathbb{N}$ such that for all $m \geqslant N$ and $n \geqslant N$ we have $(x_n \cdot y_m) \leqslant K + \delta$.*

Proof. See [61]. For the last property, note that
$$K \geqslant (x_n \cdot y_n) \geqslant \min\{(x_n \cdot y_m), (y_n \cdot y_m)\} - \delta$$
and $(y_n \cdot y_m) \to \infty$. □

We now construct a topology on $\hat{X}$.

Definition 2.36 *For a hyperbolic metric space X, $\mathcal{O}(\hat{X})$ is the collection of subsets of $\hat{X}$ consisting of*

1. *The open balls $B_r(x)$ with $x \in X$ and $r > 0$.*
2. *The sets of the form $N_K(x) = \{y \in \hat{X} \mid (x \cdot y) > K\}$ where $x \in \partial X$ and $K > 0$.*

Proposition 2.37 *If X is a hyperbolic metric space then $\mathcal{O}(\hat{X})$ is a basis of neighbourhoods for a topology on $\hat{X}$.*

Proof. It is clear that the sets in $\mathcal{O}(\hat{X})$ cover $\hat{X}$. We also need to show that if B_1 and B_2 are in $\mathcal{O}(\hat{X})$ and $Y \in B_1 \cap B_2$ then there exists $B_3 \in B$ such that $Y \in B_3$ and $B_3 \subset B_1 \cap B_2$. First of all, if B_1 and B_2 are both open sets of type 1 then it is obvious (we can use a general arqument which works in any hyperbolic metric space).

Now suppose that B_1 is an open set of type 1 and B_2 is an open set of type 2. Let $B_1 = B_r(e)$ and let $B_2 = N_K(z)$ for some $z \in \partial X$. If we take $y \in B_1 \cap B_2$ then $B_{\varepsilon_1}(y) \subset B_1$ for some $\varepsilon_1 > 0$. Since $y \in B_2$, $(z.y) = K + \varepsilon_2$ for some $\varepsilon_2 > 0$. Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. Then $B_\varepsilon(y) \subset B_r(x)$. We also need $B_\varepsilon(y) \subset N_k(z)$. Let p be in $B_\varepsilon(y)$ and by property 5, pick $z_i \to z$ such that $\lim_{i\to\infty}(z_i \cdot p) = (z \cdot p)$. As in the proof of property 1, we obtain

$$|(z_i \cdot p) - (z_i \cdot y)| \leqslant d(p, y) < \varepsilon \leqslant \varepsilon_2,$$

which in particular gives $(z_i \cdot p) - (z_i \cdot y) \geqslant -\varepsilon_2$. We want to show that $p \in B_2$, i.e. that $(z \cdot p) > K$. But

$$(z \cdot p) = (z \cdot p) - (z \cdot y) + (z \cdot y) > (z \cdot p) - (z \cdot y) + (k + \varepsilon_2),$$

which is at least $-\varepsilon_2 + K + \varepsilon_2 = K$. This completes the second part of the proof.

The third case is where B_1 and B_2 are both of the second type. Say $B_1 = N_K(x)$ and $B_2 = N_L(y)$. Let z be a point of $B_1 \cap B_2$. Then we need B_3 with $y \in B_3 \subset B_1 \cap B_2$. Suppose $z \in X$. Then B_3 is of type 1, and as in the previous case we get ε_1 such that $B_{\varepsilon_1}(y) \subset N_K(x)$ and ε_2 such that $B_{\varepsilon_2}(y) \subset N_L(y)$. Taking $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$ we get $B_\varepsilon(y) \subset N_K(x) \cup N_L(y)$. It is enough to show that if $z \in \partial X \cup N_K(x)$ where $x \in \partial X$ then there exists $n > 0$ such that $N_n(y) \subset N_K(x)$, since if $r_1 > r_2$ then $N_y(r_1) \subset N_y(r_2)$. Suppose not. Then there exists a sequence $\{y_n\}$ in $N_n(y)$ such that $y_n \notin N_K(x)$. Either infinitely many y_n lie in X or only finitely many do. In the first case we can assume by passing to a subsequence that $y_n \in X$ for all n. Then $(y_n \cdot x) \leqslant K$ and $(y_n \cdot y) > n$ for all n by definition. So $y_n \to y$ by property 4. By property 8, we deduce that

$$K < (x \cdot y) < \liminf_m (x \cdot y_m) \leqslant K.$$

Which is a contradiction. So we may assume that only a finite number of y_n lie in X and by passing to a subsequence that all y_n are in ∂X. Now there is a k' such that $k < k' < (x \cdot y) < \infty$ and for all m, $(x \cdot y_m) \leqslant k < k'$. Recall also that $(y_m \cdot y) > m$. By property 5 we may choose for each m, sequences $x_i^{(m)} \to x$ and $y_i^{(m)} \to y_m$ such that

$$\lim_{i\to\infty} (x_i^m \cdot y_i^m) = (x \cdot y_m) \leqslant k < k'.$$

Passing to subsequences we may assume that for each m and for each i, $(x_i^{(m)} \cdot y_i^{(m)}) < k'$. We may also assume by passing to a subsequence that $(x_i^{(m)} \cdot x) > n$ for all n, since $x_i^{(m)} \to x$ as $i \to \infty$. Similarly, as $(y_m \cdot y) > m$ we may assume that for all m and for all i, $y_m^{(i)} > n$. Taking the diagonal subsequences $x_i^{(i)}$ and $y_i^{(i)}$ we see that by property 4 that $x_i^{(i)} \to x$ and $y_i^{(i)} \to y$. But then

$$k' < (x \cdot y) = \inf\{\liminf_i (x_i \cdot y_i)\} \leqslant \liminf_i (x_i^{(i)} \cdot y_i^{(i)}) \leqslant k',$$

which is a contradiction. □

See [19] for details of how to metrize this topology.

2.7 The Rips Complex.

Let X be a metric space and let $D > 0$. The *Rips Complex* $P_D(X)$ is the simplicial complex whose vertices are the points of X and whose simplices are the finite subsets of X whose diameter is at most D.

Example 2.38 For $D \leqslant D'$, $P_D(X)$ is a subcomplex of $P_{D'}(X)$ and $P_\infty(X) = \bigcup_D P_D(X)$ is the standard simplex in $\mathbb{R}^X$. If X is bounded then $P_D(X)$ is the standard simplex in $\mathbb{R}^X$ for $D \geqslant \mathrm{diam}(X)$.

Theorem 2.39 *Let X be a δ-hyperbolic metric space. Then $P_D(X)$ is contractible for all $D \geqslant 4\delta$.*

We say that $P_D(X)$ is *stably contractible.*

Proof. By Whitehead's theorem, the sequence of homotopy groups is a complete invariant of CW complexes, which include simplicial complexes. For such spaces, contractibility is hence equivalent to the vanishing of all homotopy groups. Let x be a basepoint for X. We show that for $D \geqslant 4\delta$, every finite subcomplex K of $P_D(X)$ is homotopic to x, which implies that all homotopy groups must vanish.
Suppose that all the vertices of K are a distance at most $\frac{D}{2}$ from x_0. Then all the vertices of K are a distance at most D from each other and K is contained in a simplex of $P_D(X)$. K is thus homotopic to x.

Otherwise, let y be in K^0 (the 0-skeleton of K) with $|y| = d(y, x)$ maximal (and hence greater than $\frac{D}{2}$). Let $[x, y]$ be a geodesic segment and let z be a point on this segment such that $d(y, z) = D/2$. Let $f : K^0 \to P_D(X)$ be the map which sends y to z and leaves fixed the otherpoints of K^0. This map extends to a simplicial map if for all simplices σ of K, $f(\sigma)$ is a simplex of $P_D(X)$. This condition is true if for all $w \in K^0$ such that $d(w, y) \leqslant D$ we also have $d(w, z) \leqslant D$. Now hyperbolicity of X gives

$$d(y, z) + |y| \leqslant \max\{d(w, y) + |z|, d(w, z) + |w|\} + 2\delta.$$

and as $d(w, y) \leqslant D$ for all $y \in K^0$, we have

$$d(w, z) \leqslant \max\{D + |z| - |y|, d(w, z)\} + 2\delta = \frac{D}{2} + 2\delta,$$

Which is at most D since $d \geqslant 4\delta$. Now a finite sequence of such homotopies allows us to bring all the points of K to a distance less than $\frac{D}{2}$ from the basepoint. As we know that after this sequence the resultant complex is contractible to the basepoint, this completes the proof. $\square$

3 Hyperbolic Groups.

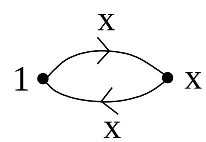
3.1 Cayley Graphs and Hyperbolicity.

Definition 3.1 *Let G be a group with generating set S. Its* Cayley graph *with respect to these generators, $\Gamma_S(G)$ is the graph with vertex set $\{g \mid g \in G\}$ and edge set $\{(g, gs) \mid s \in S, g \in G\}$.*

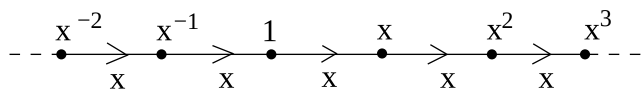
A Cayley graph $\Gamma_S(G)$ has a natural labelling, where the edge (g, gs) is labelled by s. Note that $\Gamma_S(G)$ is directed but we can consider it to be undirected if we take an inverse-closed generating set.

Example 3.2 Here are some examples of Cayley graphs:

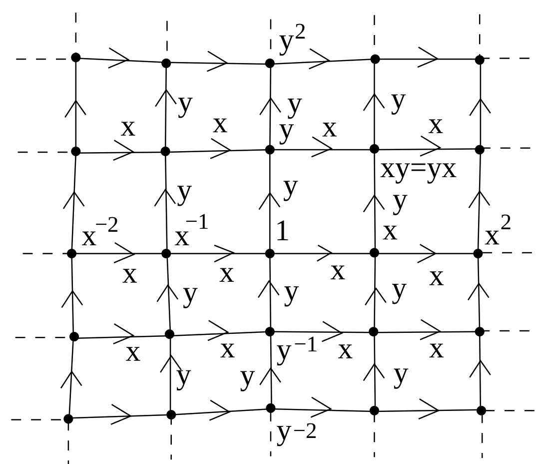
1. $\mathbb{Z}_2 = \langle x | x^2 \rangle$:



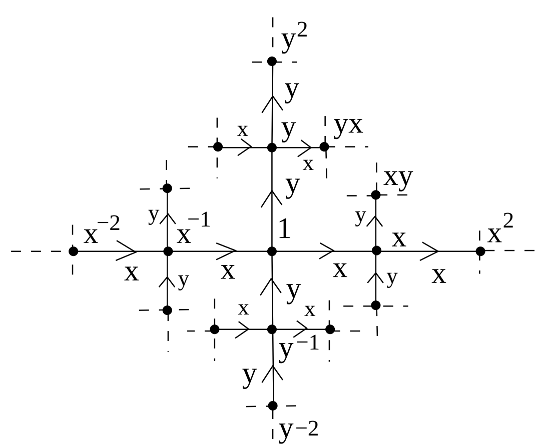
2. $\mathbb{Z} = \langle x | \ \rangle$:



3. $\mathbb{Z} \oplus \mathbb{Z} = \langle x, y | [x, y] \rangle$:



4. $F_2 = \langle x, y | \ \rangle$:

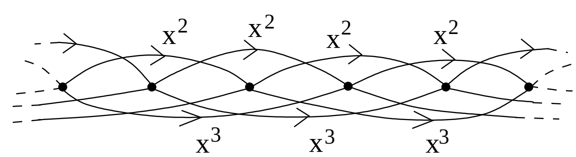


Every word w in $F(S)$ corresponds to a path in $\Gamma_S G$ starting at the identity, where the endpoint of the path is the element w represents in G, denoted $\bar{w}$. In particular, words which represent the identity in G correspond to a loop in the Cayley graph.

We may consider a group G with generating set S as a metric space as follows. We define the *length* $l(w)$ of a reduced word w, in a free group to be the number of letters in it. Then we define the *modulus* of g for $g \in G$ to be $|g|_S = \min_{w \in F(S), w =_G g}\{l(w)\}$. The *word metric* on G is now defined to be the metric given by $d(g,h) = |g^{-1}h|$ for g and h in G. Note that this is the same as the metric on the Cayley graph when we give every edge length 1, and what we have called the modulus is the associated norm when we define the identity to be the basepoint of the Cayley graph.

G acts freely on $\Gamma_S(G)$ by left multiplication, i.e. $h \in G$ sends the edge (g, gs) to the edge (hg, hgs). This is an action by isometries with respect to the word metric. Moreover, the Cayley graph is a homogeneous metric space, i.e. its isometry group acts transitively: If g_1 and g_2 are two vertices of the Cayley graph then an isometry taking g_1 to g_2 is left multiplication by $g_2 g_1^{-1}$.

Note that the Cayley graph of a group depends on the generating set. We have seen the Cayley graph of $\mathbb{Z}$ with generator x in example 3.2. If we take instead the Cayley graph with generators x^2 and x^3 then it is as follows.



However, the following observation due to Gromov allows us to assign a unique geometric object to a group (the quasi-isometry class of its Cayley graph).

Proposition 3.3 *Let S_1 and S_2 be finite generating sets for a group G. Then $\Gamma_{S_1}(G)$ is quasi-isometric to $\Gamma_{S_2}(G)$.*

Proof. A quasi-isometry is given by the identity maps between the groups with different generating sets. Let $S_1 = \{a_1, \ldots, a_n\}$ and $S_2 = \{b_1 \ldots b_m\}$. We can express each a_i as a product of $b_j s$. Say $a_i = \bar{w_i}$ for each $1 \leqslant i \leqslant n$ where each w_i is a word in S_2. Let $\lambda = \max_{1 \leqslant i \leqslant n}(l(w_i))$. Let g_1 and g_2 be in G. Then for some sequence $a_{i_1}, \ldots, a_{i_p}$ of generators in S_1 we have $d_{S_1}(g_1, g_2) = |g_1^{-1} g_2| = |a_{i_1} \cdots a_{i_p}| = p$. But $a_{i_1} \cdots a_{i_p} = w_{i_1} \cdots w_{i_p}$ and hence $d_{S_2}(g_1, g_2) = |g_1^{-1}|_{S_2} \leqslant |w_{i_1} \cdots w_{i_p}| \leqslant \lambda p = \lambda d_{S_1}(g_1, g_2)$. Similarly there exists μ such that $d_{S_1}(g_1, g_2) \leqslant \mu d_{s_2}(g_1, g_2)$. Thus if we let $\nu = \max\{\lambda, \mu\}$ then $\Gamma_{S_1}(G)$ is $(\nu, 0)$-quasi-isometric to $\Gamma_{S_2}(G)$. $\square$

In the light of the last proposition it makes sense to talk about a *quasi-isometry invariant* of a finitely generated group, since a quasi-isometry invariant of the Cayley graph does not depend upon the chosen generating set. The study of quasi-isometry invariants of groups is an important part of geometric group theory.

Exercise 3.4 Show that the property of being finitely presented is a quasi-isometry invariant.

Exercise 3.5 Show that if H is a finite index subgroup of a group G then G and H are quasi-isometric.

If P is a property of groups then we say that a group G is *virtually* P if G has a finite index subgroup which satisfies the property P. For example, a virtually free group is a group with a free subgroup of finite index.

Example 3.6

1. Virtual freedom is a quasi-isometry invariant (see[19]).
2. The property of being virtually abelian is a quasi-isometry invariant by Bieberbach's theorem (see [18],p.88), which states that a finitely generated group is virtually abelian if and only if it it is a discrete subgroup of the group of isometries of $\mathbb{R}^n$ for some n (see also exercise 3.7).

3. By Gromov's growth theorem (see theorem 6.1 in section 6), virtual nilpotency is a quasi-isometry invariant. Dioubina has, however, shown in [16] that virtual solvability is not, and that neither is the property of being virtually torsion free.

4. The property of splitting over a finite subgroup as a free product with amalgamation or HNN extension is a quasi-isometry invariant, by Stallings' ends theorem (see [63],[53] or [15]). We define the *number of ends* of a group as follows. Let G be a finitely generated group and let Γ be the Cayley graph of G with respect to some fixed finite generating set S. Let B_n be the ball of radius n around the identity in Γ and let e_n be the number of infinite connected components of $\Gamma - B_n$. Then $e(G)$, which we define to be $\lim_{n\to\infty} e_n$, always exists and is equal to either 0,1,2 or ∞. It is known to be a quasi-isometry invariant of groups [9]. Stallings' ends theorem states that $e(G) > 1$ if and only if G splits as a free product with amalgamation or an HNN extension over a finite subgroup.

Given a finitely presented group $G = \langle S, R\rangle$ we can construct a 2-dimensional CW Complex of which the Cayley graph is the 1-skeleton. This is the *Cayley complex* $C_{S,R}(G)$, where we attach a 2-cell to each loop in the Cayley graph labelled by a relator. We can also construct a finite 2-dimensional CW Complex $K_{S,R}(G)$ such that $\Pi_1(K_{S,R}(G)) = G$ by taking a wedge of $|S|$ circles and attaching for each relator a 2-cell via the map given by the relator. Then the Cayley complex is the universal covering space of this complex. $K_{S,R}(G)$ has the property (by general CW complex theory) that $\Pi_1(K_{S,R}(G)) \cong G$.

Exercise 3.7 Let G act on the metric space X. If for each compact subset K of X, the set $\{g \in G \mid K \cap gK \neq \emptyset\}$ is finite then we say that G acts *properly discontinuously.* We say that G acts *cocompactly* on X if the quotient X/G is compact. Prove the following theorem, due to Milnor: If a group G acts properly discontinuously and cocompactly by isometries on a locally compact geodesic metric space X then the Cayley graph of G is quasi-isometric to X.

Definition 3.8 *Let G be a group generated by the finite set S. We say that G is* hyperbolic *if $\Gamma_S(G)$ is a hyperbolic metric space.*

We have shown that hyperbolicity of metric spaces is a quasi-isometry invariant. Thus if $\Gamma_S(G)$ is hyperbolic and S' is another finite generating set for G, then $\Gamma_{S'}(G)$ is also hyperbolic. Hyperbolicity is therefore a well defined property of groups.

We shall see later that in fact a hyperbolic group must be finitely presented.

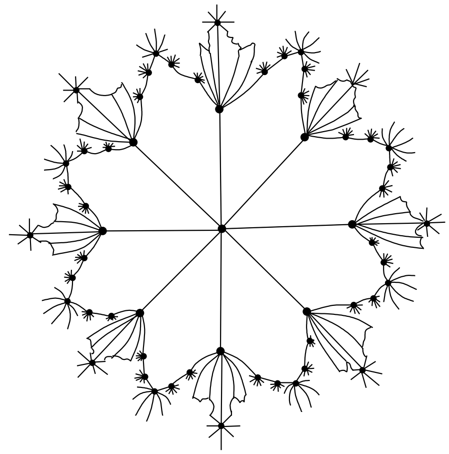
Example 3.9 1. Free groups are hyperbolic as their Cayley graphs with respect to the standard presentations are trees. So virtually free groups are also hyperbolic.

2. Finite groups are hyperbolic.

3. Cocompact Fuchsian groups are hyperbolic by Milnor's theorem (see exercise 3.7), as are any groups acting properly discontinuously and cocompactly on $\mathbb{H}^n$. An example of these is a *surface group*, the fundamental group of a compact surface, of genus 2 or more. The standard presentation for an orientable surface group of genus g is

$$S_g = \langle a_1, b_1, \dots a_g, b_g \mid \prod_{i=1}^{g} [a_i, b_i]\rangle.$$

For example, the Cayley graph of S_2 with these generators is as follows. It is planar, and each face is an octagon. There are eight octagons meeting at each vertex.



The standard presentation for a non-orientable surface group of genus g is

$$T_g = \langle t_1, \dots t_g \mid \prod_{i=1}^{g} t_i^2 \rangle.$$

Non-orientable surfaces have double covers which are orientable so for all g, T_g has an orientable surface subgroup of index 2.

4. The free abelian group $\mathbb{Z} \oplus \mathbb{Z}$ on two generators is *not* hyperbolic. In fact it can never even be a subgroup of a hyperbolic group (we prove this later).

Exercise 3.10 1. Show that if G and H are hyperbolic groups then their free product $G * H$ is hyperbolic.

2. Generalise to the free product of two hyperbolic groups with amalgamation over a finite subgroup.

3. Show that if $G * H$ is hyperbolic then so are G and H.

4. If A and B are torsion free hyperbolic groups and C is a maximal cyclic subgroup of both A and B then $A *_C B$ is hyperbolic (see [23])

3.2 Diagrams and Area in Groups.

If $G = \langle S|R \rangle$ is a finitely presented group and w is a word in the free group $F(S)$, then whenever $w =_G 1$, we can write

$$w = \prod_{i=1}^{n} u_i r_i^{\pm 1} u_i^{-1}$$

for some $u_i \in F(S)$ and $r_i \in R$. (This is just the definition of the group presented by $\langle S|R \rangle$. Recall that we quotient out the normal closure of R in $F(S)$.)

Definition 3.11 *Let $G = \langle S, R \rangle$ be a finitely presented group. Let w be a word in S with $\bar{w} = e$ in G. We define the* area *of w, $A(w)$ to be* $\min\{n \mid w = \prod_{i=1}^{n} u_i r_i^{\pm 1} u_i^{-1}\}$

We say that $f : \mathbb{N} \to \mathbb{N}$ is a *Dehn function* or *isoperimetric function* for G if for all $w \in F(S)$ with $\bar{w} = e$ we have $A(w) \leqslant f(|w|)$. We say that G satisfies a *linear*, *quadratic*, *polynomial* or *exponential isoperimetric inequality* if it has a Dehn function with the corresponding property.

Exercise 3.12 Show that $\mathbb{Z} \oplus \mathbb{Z}$ satisfies a quadratic isoperimetric inequality but not a linear one.

Let f and g be functions from $\mathbb{N}$ to $\mathbb{N}$. Write $f \preceq g$ if there exist constants A,B and C such that $f(n) \leqslant Ag(Bn) + Cn$. We say that $f \sim g$ if $f \preceq g$ and $g \preceq f$. This is an equivalence relation. For example all polynomials of the same degree are equivalent, while those of different degree are not.

Write $f \sim g$ if there exist strictly positive A, B and C such that

$$f(n) \leqslant Ag(Bn) + Cn \text{ and } g(n) \leqslant Af(Bn) + Cn.$$

Alonso has shown the following in [1].

Theorem 3.13 *If G_1 and G_2 are two groups with Dehn functions f_1 and f_2 and G_1 is quasi-isometric to G_2 then $f_1 \sim f_2$.*

In particular, the equivalence class of the Dehn function is a group invariant (i.e. is independent of presentation). We now give a geometric interpretation of area.

Definition 3.14 A map *is a finite, planar, oriented, connected and simply connected simplicial 2-complex. We say that a map D is a* diagram *over a paired alphabet S if every edge e of D has a label $\phi(e) \in S$ such that $\phi(e^{-1}) = (\phi(e))^{-1}$.*

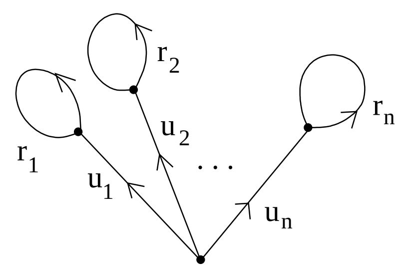
In this way a (simplicial) path along the edges of a diagram is labelled by a word in S and a path which doesn't backtrack over itself is labelled by a reduced word in S.

Define the boundary of a map (hypotheses on "diagram" ensure that the boundary is well defined).

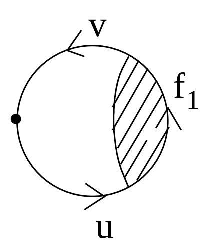
Definition 3.15 A van Kampen diagram *over a group $G = \langle S, R\rangle$ is a diagram D over S such that for all faces f of D the label of the boundary path of f is labelled by some $r^{\pm 1}$ with $r \in R$. The* area *of such a diagram is the number of its faces.*

Proposition 3.16 (van Kampen's Lemma) *Let $G = \langle S, R\rangle$ be a finitely presented group and let w be a word in $F(S)$. Then $\bar{w} = e$ if and only if there exists a van Kampen diagram D over G with boundary labelled by w.*

Proof. Suppose $\overline{w} = 1$. Then in $F(S)$, where each r_i is in R. Fold adjacent edges on the boundary with inverse labels until there are no such edges. This corresponds to cancellation of inverses in $F(s)$ and the resulting diagram has boundary label w.



Conversely suppose that w is the label of the boundary of a van Kampen diagram D. We claim that $\overline{w} = 1$. This is proved by induction on the number k of faces of D. For $k = 1$ it is obvious. Assume true if D has k faces, and then suppose D has $k+1$ faces. Now there exists a face F containing an edge f_1 on ∂D.



Then for some words u and v on ∂D, $w = uf_1v = uf_1f_2f_2^{-1}v$ where ∂F is labelled f_1f_2. But this is equal to $(uf_1f_2u^{-1})(uf_2^{-1}v)$ and by induction, as $D - F$ only has k faces, we know that $\overline{uf_2^{-1}v} = 1$. But also $\overline{uf_1f_2u^{-1}} = 1$ as $f_1f_2 \in R$. Hence $\overline{w} = 1$. □

Definition 3.17 *A* minimal van Kampen diagram *for a word w is a van Kampen diagram for w with the minimum number of faces.*

Thus if D is a minimal van Kampen diagram for w then the area of D is equal to the area of w and the length of ∂D is equal to $|w|$.

If D is a van-Kampen diagram for a word representing the identity in G with respect to a generating set S, then there is an obvious map from the 1-skeleton of D to the Cayley graph of G with respect to S. If we send a point of ∂D to the identity then the rest is determined by the labels of the edges. This map is distance non-increasing and extends to a map to the Cayley complex $C_{S,R}(G)$. A word w equal to the identity in G corresponds to a closed path in the Cayley complex. As the Cayley complex is simply connected, there is a map $f : (D, \partial D) \to C_{S,R}(G)$ with $f(\partial D) = w$. After a homotopy of f the cell decomposition of $C_{S,R}(G)$ induces a cell decomposition on D. This is the van Kampen diagram for w.

3.3 Algorithms In Group Theory.

Try the following exercise. Hopefully it illustrates how working with presentations is largely an *ad hoc* process.

Exercise 3.18 Let G be the group with presentation

$$\langle a, b \mid ab^2 = b^3a, ba^2 = a^3b \rangle.$$

Show that G is trivial.

The following problems were raised by Max Dehn in 1911. By an *algorithm* we shall mean a process (which can ultimately be reduced to the operation of a Turing machine) having a set of input data and terminating after a finite number of steps, giving a set of output data. Let C be a class of finitely presented groups.

- *The Word Problem for C.* Is it true that for all groups $G \in C$ there exists an algorithm which takes as its input a word in G and whose output is either "yes" if the word is equal to the identity or "no" if it is not?

- *The Conjugacy Problem for C.* Is it true that for all groups $G \in C$ there exists an algorithm which takes as its input a pair of words in G and whose output is either "yes" if the two words represent conjugate elements in G or "no" if they don't?

- *The Isomorphism Problem for C.* Does there exist an algorithm which takes as input two finite presentations of groups G_1 and G_2 in C and gives as output either "yes" if the groups are isomorphic or "no" if they are not?

It is a fundamental result of Novikov and Boone that there exist finitely presented groups for which the word problem is not solvable (see [52] for a good account). This means that for $n \geqslant 5$ we cannot hope for an algorithmic classification of n-manifolds, since every finitely presented group can appear as the fundamental group of such a manifold. In fact, Novikov and Boone proved unsolvability of the conjugacy problem in general first. (Solvability of the conjugacy problem implies that of the word problem since we can take one of our words to be the empty word, and if a group element is conjugate to the identity then it is clearly equal to the identity.) the solvability of the word problem is a quasi-isometry invariant of groups. The result then follows from the next exercise.

Exercise 3.19 Let $G = \langle S, R \rangle$ be a finitely presented group and let m be the maximum length of a relator. Suppose that $\bar{w}$ is the identity in G. Show using van-Kampen diagrams on n that there are $u_i \in F(S)$ such that $w = \prod_{i=1}^{n} u_i^{-1} r^{\pm 1} u_1$ where each $r_i \in R$ and

$$\max_{1 \leqslant i \leqslant n} \{|u_i|\} \leqslant mA(w) + |w|.$$

Use this to show that the word problem is solvable in G if and only if G has a recursive isoperimetric inequality (see [18]). Deduce from theorem 3.13 that the solvability of the word problem is a quasi-isometry invariant of groups.

Definition 3.20 *A* Dehn presentation *for a group G is a presentation such that every word in $F(G)$ which represents the identity in G contains more than half of a relator.*

For example, the standard presentations of free groups and surface groups (of genus at least 2) are Dehn presentations.

Theorem 3.21 *If a group G has a Dehn presentation then the word problem is solvable for G*

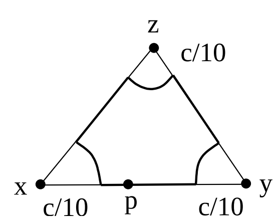
Proof. Let $\langle S, R \rangle$ be a Dehn presentation for G. Suppose that w is a word in $F(G)$. Either it doesn't contain more than half of a relator (in which case we know it is not equal to the identity) or it does, i.e. $w = ar_1b$ where there exists a word r_2 such that $r_1r_2 = 1$ and $|r_1| > |r_2|$. Then $w = ar_2^{-1}b$ in G, which is a word of strictly shorter length. We repeat this process. Either we eventually reach the identity in which case the word is trivial in G, or we reduce w to a word which does not contain more than half a relator and the word is not trivial in G. Thus we have an algorithm terminating after a finite number of steps which tells us if a given word in G is trivial or not. Hence the word problem is solvable in G. $\square$

3.4 The Linear Isoperimetric Inequality.

We show in this section that the linear isoperimetric inequality characterises hyperbolicity of groups. As corollaries we show that the word problem is solvable for hyperbolic groups and that such groups are finitely presented.

Theorem 3.22 *If a finitely presented group $G = \langle S, R \rangle$ satisfies a linear isoperimetric inequality then it is hyperbolic.*

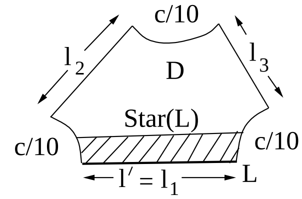
Proof. Suppose that for all $w \in G$ with $\bar{w} = 1$ we have $A(w) \leqslant K|w|$, and G is not hyperbolic. Then triangles in G are not δ-slim, i.e. for all $c > 0$ there is a geodesic triangle in the Cayley graph of G which is c-thick (which means that there is a point of the triangle whose distance from the union of the other two sides is at least c). Choose such a triangle $[x, y, z]$ and truncate it to get a hexagon H whose vertices are the six points of $[x, y, z]$ which are at a distance of $\frac{c}{10}$ from x,y and z.



Suppose p is the point for which

$$d(p, [x, y] \cup [y, z]) = \max_{q \in [a,b]} \{d(q, [x, z] \cup [y, z])\}.$$

Let D be a minimal van-Kampen diagram for H. Suppose that l_1,l_2 and l_3 are the lengths of the sides of H which are truncated sides of $[x, y, z]$ and let $l' = max\{l_1, l_2, l_3\}$ be realised by the line L (in the following picture we have assumed that this is equal to l'.)



Then $l(\partial D) < 6l'$ and by the isoperimetric inequality, $A(D) < 6Kl'$. If we let ρ be the maximum length of a relator of G then we have

$$A(\mathrm{Star}(L)) \geqslant \frac{l'}{\rho}.$$

and if we put $L_1 = \partial(\mathrm{Star}(L)) - \partial H$, then

$$l(L_1) \geqslant l' - 2\rho$$

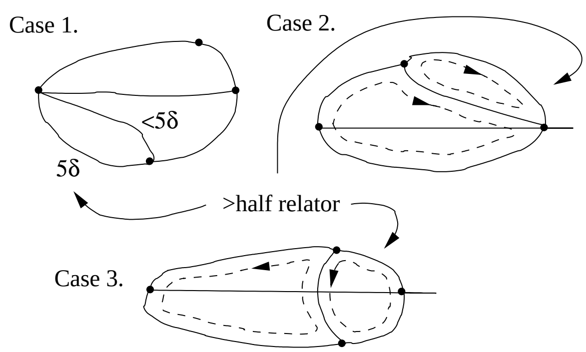
or we contradict the fact that L is a geodesic. By repeating $12K$ times we get

$$\begin{aligned} A(\mathrm{Star}^{12K}(L)) &\geqslant \frac{l'}{\rho} + \frac{l'-2\rho}{\rho} + \cdots + \frac{l'-12k\rho}{\rho} \\ &\geqslant 12K\rho\frac{(l'-12K\rho)}{\rho} \\ &= 12K(l'-12K\rho) \\ &> 6Kl' \end{aligned}$$

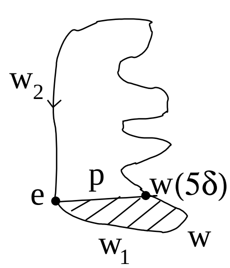
Which violates the isoperimetric inequality. Note that we can pick c large enough so that $l' - 12k\rho \geqslant \frac{l'}{2}$ since $l' > \frac{c}{10}$. Thus the proof by contradiction is complete. □

Theorem 3.23 *Hyperbolic groups satisfy a linear isoperimetric inequality.*

Proof. Let $G = \langle S|R\rangle$ be a hyperbolic group and suppose that triangles in its Cayley graph are δ-thin, where we take δ to be an integer. Let $K = \max\{A(w) \mid |w| \leqslant 10\delta\}$. We shall show by induction on $|w|$ that $A(w) \leqslant K|w|$ for all $w \in F(S)$ with $\overline{w} = 1$. If $|w| \leqslant 10\delta$ then this is obvious. Suppose that it is true for $|w| < n$ where $n \geqslant 10\delta$. Let $|w| = n+1$. There are three cases.



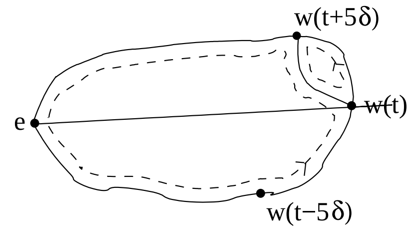
Case 1: Suppose that for all vertices $w(i)$ of w we have $d(w(i), e) < 5\delta$. We may take a shortest path p joining e to $w(5\delta)$.



Then $w = w_1 p^{-1} p w_2$ and we have by our inductive hypothesis,

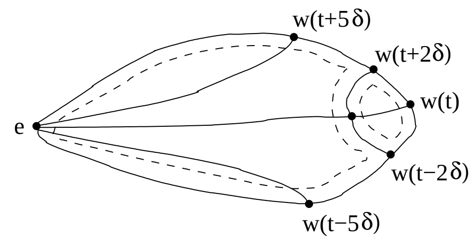
$$A(w) \leqslant A(w_1 p^{-1}) + A(p w_2) \leqslant K + Kn = K(n+1) = K|w|.$$

Case 2: On the other hand, the above may not hold. Let $w(t)$ be the vertex of w furthest from e.



If $d(w(t), w(t-5\delta)) < 5\delta$ or $d(w(t), (w(t+5\delta)) < 5\delta$ then we are done as in the previous case. This is because if we let γ be a geodesic from $w(t-5\delta)$ to $w(t)$ (or $w(t)$ to $w(t+5\delta)$), let the loop w' be equal to $w|_{[t-5\delta,t]}\gamma$ (or $w|_{[t,t+5\delta]}\gamma$) and let the loop w'' be equal to $w|_{[e,t-5\delta]}\gamma w|_{[t,n]}$ (or $w|_{[e,t]}\gamma w|_{[t+5\delta,n]}$) then we have $A(w) = A(w') + A(w'')$. $A(w'') \leqslant K|n|$ by the inductive hypothesis and $A(w'') \leqslant K$ by definition of K.

Case 3:



Otherwise neither of the two cases above may hold. In this case if we condider the geodesic triangles $[e, w(t-5\delta), w(t)]$ and $[e, w(t), w(t+5\delta)]$, it is easy to see, using thinness of these triangles that

$$d(w(t-2\delta), w(t+2\delta)) \leqslant 2\delta,$$

And we are done again, similarly to the last two cases. □

By exercise 3.12, we see that $\mathbb{Z} \oplus \mathbb{Z}$ is not hyperbolic.

Corollary 3.24 *The word problem is solvable for hyperbolic groups.*

Proof. We claim that for a hyperbolic group G generated by a finite set S, $\langle S|R\rangle$ is a Dehn presentation, where R is equal to

$$\{w \in F(S) \mid |w| \leqslant 10\delta, \overline{w} = 1\}.$$

Let w be a word equal to the identity in G. We have to show that w contains more than half a relator. There are three cases as in the previous proof. In each case, we do in fact show that w contains more than half a relator. □

In fact the existence of a Dehn presentation implies hyperbolicity (see [61]). There are also notions of generalised Dehn presentations, which can be found for non-hyperbolic groups such as nilpotent groups [20].

Corollary 3.25 *Hyperbolic groups are finitely presented.*

Proof. This follows from the proof of corollary 3.24. □

An interesting fact about isoperimetric inequalities is the following theorem, which is usually referred to as the "gap in the isoperimetric spectrum". Proofs appear in [41] and [39]. A function $f : \mathbb{R} \to \mathbb{R}$ is called *subquadratic* if $\lim_{x\to\infty} \frac{f(x)}{x^2} = 0$. For example, the function $x \log x$ is subquadratic.

Theorem 3.26 *If a group satisfies a subquadratic isoperimetric inequality then it satisfies a linear one (and is hence hyperbolic).*

By considering the action of a hyperbolic group on the Rips complex, it is possible to show that hyperbolic groups have finite cohomological dimension (over $\mathbb{Q}$) and satisfy the higher finiteness conditions (see [19]).

Bowditch has axiomatized the notion of area in a metric space in [5]. Within this framework, he shows that metric spaces (with an area function satisfying his axioms) are hyperbolic if and only if they satisfy a linear isoperimetric inequality. He also shows in [6] that there is an analogue of theorem 3.26 for metric spaces.

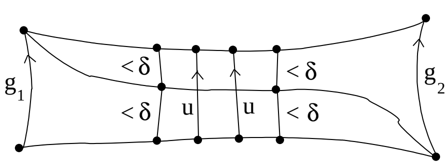
3.5 Conjugacy In Hyperbolic Groups.

Theorem 3.27 *In a hyperbolic group there are only finitely many conjugacy classes of finite order elements.*

Proof. Let G be a hyperbolic group and let $\langle S|R\rangle$ be a Dehn presentation of G. Let $g \in G$ be a finite order element and let w be a word of minimal length in the conjugacy class of g. Then $w^n = e$ for some $n \in \mathbb{N}$. Therefore since we are dealing with a Dehn presentation, w^n contains more than half a relator r, where $r = r_1 r_2$ and $|r_1| > |r_2|$. It is not possible that $r_1 \subset w$ because this contradicts the fact that w has minimal length in the conjugacy class of g. Now suppose $w = utv$ where $vu = r_1$. Then $u^{-1}wu = tvu$ and this is in the conjugacy class of g. Hence $tvu = tr_1 = tr_2^{-1}$ and $|tr_2|^{-1} < |w|$, which is a contradiction. Therefore $|w| < |r_1| < |r|$ and every conjugacy class has a representative of length less than the maximum length of a relator in our presentation. There are thus finitely many conjugacy classes of finite order elements in G. □

Theorem 3.28 *The conjugacy problem is solvable for hyperbolic groups.*

Proof. Let $\langle S, R\rangle$ be a Dehn presentation for a hyperbolic group G and assume that triangles in the corresponding Cayley graph are δ-thin. We claim that if $g_1 \in G$ is conjugate to $g_2 \in G$ then there is an element x of G such that $\overline{g_1} = \overline{xg_2x^{-1}}$ and $|x| \leqslant |S|^{2\delta} + |g_1| + |g_2| + 1$. To see this, let x be a word of minimal length such that $\overline{g_1} = \overline{xg_2x^{-1}}$. Let x_i be the i^{th} prefix of x, i.e. if x is equal to $y_1 \cdots y_n$ with $y_i \in S$ for all $1 \leqslant i \leqslant n$ then $x_i = y_1 \cdots y_i$.



By using 2δ-thinness of the rectangle $[e, \overline{x_i^{-1}}, \overline{x_i^{-1}g_1}, \overline{x_i^{-1}g_1x_i}]$, we see that for $|g_1| \leqslant i \leqslant n - |g_2|$, $\overline{x_i^{-1}g_1x_i}$ has a representative word of length less than 2δ. Thus since we have used every possible combination of elements of S to make words of length less than 2δ, there must be i and j with $i < j$ such that $x_i^{-1}g_1x_i = x_j^{-1}g_1x_j$. But this contradicts the minimality of x.

Now the conjugacy problem is solvable for G because to check if g_1 and g_2 are conjugate, we just check whether or not $\overline{g}_1 = \overline{xg_2x^{-1}}$ for any word x of length at most $|S|^{2\delta} + |g_1| + |g_2| + 1$. □

3.6 Small Cancellation Conditions.

Let F be the free group on a finite set S. A reduced word $w = s_1 \cdots s_n$ in F is called *cyclically reduced* if $s_1 \neq s_n^{-1}$. If there is no cancellation in forming the product $z = w_1 \cdots w_n$ then we write $z \equiv w_1 \cdots w_n$. A subset R of F is called *symmetrised* if all elements of R are cyclically reduced and for all $r \in R$ all cyclically reduced conjugates of r and r^-1 belong to R. Suppose R is symmetrised with r_1 and r_2 in R. If there exist words b, c_1 and c_2 in F such that $r_1 \neq r_2$, $r_1 \equiv bc_1$ and $r_2 \equiv bc_2$ then we call b a *piece*.

Definition 3.29 *The* small cancellation conditions *for a group $G = \langle S, R\rangle$ are as follows*

Condition $C'(\lambda)$ *Whenever $r \in R$ with $r = bc$ where b is a piece, then $|b| < \lambda|r|$.*

Condition $C(p)$ *No $r \in R$ is a product of fewer than p pieces*

Note that $C'(\lambda) \Rightarrow C(p)$ if $\lambda \leqslant \frac{1}{p-1}$.

Exercise 3.30 Show that surface groups of genus at least 2 satisfy small cancellation conditions.

Theorem 3.31 *Let $\langle S \mid R\rangle$ be a symmetrized finite presentation of a group G that satisfies the small cancellation condition $C(7)$ (or $C'(6)$). Then G is hyperbolic.*

Proof. We show that G satisfies a linear isoperimetric inequality. Let w be a word equal to the identity in G and let D be a van Kampen diagram for w. Let $d(V)$ denote the valency (degree) of a vertex in D and delete all vertices of valency 2, giving the new edges the length equal to the number of edges they contained before deletion of such vertices. We write V for the number of vertices of D, V° for the number of those in the interior of D and $V^\bullet$ for the number of those on ∂D. Similarly for edges and faces we define E,E°,$E^\bullet$ and F,F°,$F^\bullet$, where we consider f to be a boundary face if $f \cap \partial D$ contains at least one edge.

Now if we sum over all vertices, $\sum d(v) \geqslant 3v$ but on the other hand $\sum d(v) = 2E$. Hence

$$V \leqslant \frac{2}{3}E.$$

Also if we let $i(f)$ denote the number of edges of a face, then if f is an interior face, the small cancellation condition $C(7)$ implies that $i(f) \geqslant 7$ since each edge in the van Kampen diagram corresponds to a piece. So if we sum over all interior edges f then we have $\sum i(f) \geqslant 7F^\circ$ while we also have $\sum i(f) = 2E^\circ$ because each interior edge is contained in exactly two interior faces. This gives us the following inequality.

$$F^\circ \leqslant \frac{2}{7}E^\circ.$$

Now using Euler's formula $V + F = E + 2$ we obtain

$$E + 2 \leqslant \frac{2}{3}E + \frac{2}{7}E^\circ + F^\bullet,$$

Which then gives

$$\frac{2}{21}E^\circ \leqslant F^\bullet - \frac{1}{3}E^\bullet - 2.$$

Now suppose that m is the largest length of a relator of G. Then we have

$$F^\bullet \leqslant l(\partial D) \text{ and } E^\circ \geqslant \frac{A(D)}{2m}.$$

So combined with the previous inequality we have

$$A(D) \leqslant \frac{21m}{2}l(\partial D),$$

which is a linear isoperimetric inequality as required. □

Exercise 3.32 The contents of this exercise is known as the *Rips Construction* (see [47]). Let G be a group with finite presentation $\langle s_1, \ldots, s_m \mid R_1, \ldots R_n\rangle$ and let H be the group with generators $a_1, \ldots, a_m, b_1, b_2$ and relators

$$R_i b_1 b_2^{r_i} b_1 b_2^{r_i+1} \cdots b_1 b_2^{s_i} \text{ (for } i = 1, \ldots, n),$$
$$a_i^{-1} b_j a_i b_1 b_2^{p_{ij}} b_1 b_2^{p_{ij}+1} \cdots b_1 b_2^{q_{ij}} \text{ (for } i = 1, \ldots, m,\ j = 1, 2) \text{ and}$$
$$a_i b_j a_i^{-1} b_1 b_2^{u_{ij}} b_1 b_2^{u_{ij}+1} \cdots b_1 b_2^{v_{ij}} \text{ (for } i = 1, \ldots, m,\ j = 1, 2).$$

Show that for all $\lambda > 0$ we can choose the integers r_i, s_i, p_{ij}, q_{ij}, u_{ij} and v_{ij} such that H satisfies the small cancellation condition $C'(\lambda)$. Define a homomorphism $\phi : H \to G$ by $\phi(a_i) = a_i$ for each i and show that $\langle b_1, b_2 \rangle$ is a normal subgroup of H and hence equal to $\ker(\phi)$. Moreover show that if L is a subgroup of G generated by r elements then $\phi^{-1}(L)$ is generated by $r + 2$ elements. Now show that there exist groups G such that the following are true.

1. There exists a subgroup of H which is finitely generated but not finitely presented. (Hence there exist finitely generated subgroups of hyperbolic groups which are not hyperbolic.)

2. There are finitely generated subgroups P_1 and P_2 of H such that $P_1 \cap P_2$ is not finitely generated.

3. The *generalised word problem* for a group G (given by a presentation) asks: Is there an algorithm which given input (as words in the generators of G)

   (a) a finite set $\{h_1, \ldots, h_k\}$ of elements of G

   (b) an element g of G

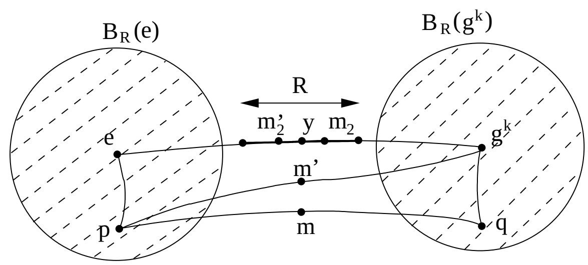
   will terminate after finitely many steps and tell us whether or not g lies in the subgroup generated by $\{h_1, \ldots, h_k\}$?

   Show that the generalised word problem is not solvable in H.

3.7 Cyclic Subgroups of Hyperbolic Groups.

Theorem 3.33 *Let G be a hyperbolic group with Cayley graph X and let g be an infinite order element of G. Then the image of $\langle g \rangle$ in X is a quasi-geodesic.*

Proof. Suppose that $R \geqslant 0$ is given. Fix a presentation of G and take the Cayley graph X with respect to this presentation. Assume that triangles are δ-slim in X and let K be such that $d(g^K, 1) > 8R + 2\delta$. Let β be the geodesic from 1 to g^K, let y be the midpoint of β and let I be the subinterval of β of length R centred at y. In what follows, by a *midpoint* we mean a vertex at distance at most $\frac{1}{2}$ from the actual midpoint.



First, we claim that if $p \in B_R(1)$ and $q \in B_R(g^K)$ then the midpoint m of $[p, q]$ is in $B_{2\delta}(I)$. Let m_1 be the midpoint of $[p, g^K]$. Then since $|d(p, g^K) - d(p, q)| < R$, we know that $|d(p, m) - d(p, m_1)| < R/2$. Using the thinness of the triangle $[p, q, g^K]$ there exists $m_1' \in [p, g^K]$ such that $d(m, p) = d(m_1', p)$, $|d(p, m_1') - d(p, m_1)| < R/2$ and $d(m_1, m_1') < \delta$. Similarly using the thinness of the triangle $[1, p, g^K]$ there are points m_2' and m_1'' on $[1, g^K]$ such that $d(m_2', g^K) = d(m_2, g^K)$, $d(m_1'', g^K) = d(m_1', g^K)$,$d(m_2', m_2) \leqslant \delta$, $d(m_1'', m_1') \leqslant \delta$, $d(m_1'', m_2') = d(m_1', m_2)$ and $d(m_2', y) < R/2$. It follows that $d(m_1'', m_1) \leqslant 2\delta$ and $d(y, m_1'') < R$. Thus we have proved the claim.

Let N be the number of vertices in the ball of radius 2δ about the identity. The 2δ-neighbourhood of I then contains at most RN vertices. Now consider the translates of the arc $B = [1, g^K]$ by each of the elements $1, g, \ldots, g^N R$. Their midpoints are all distinct (otherwise some power of g would have a fixed point and hence have finite order) and there are $1 + NR$ of them. Hence there is a number $P(R) \leqslant NR$ such that $g^{P(R)} \notin B_R(1)$, and so $g^{K+P(R)} \notin B_R(g^K)$. Note that we have

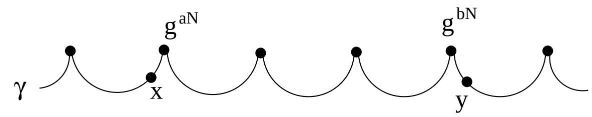
$P(R) \geqslant R/|g|$.

Next, we claim that for all R, $|g^{NR}| \geqslant R$. Suppose on the contrary that for some R_0 and $\varepsilon > 0$ we have $|g^{NR_0}| \leqslant R_0 - \varepsilon$. Then for all $s > NR_0$ let $s = nNR_0 + R_1$ with $0 \leqslant R_1 \leqslant NR_0$ and $n \in \mathbb{Z}$. We have

$$\begin{aligned} |g^s| &\leqslant |g^{nNR_0}| + |g^{R_1}| \\ &\leqslant n(R_0 - \varepsilon) + |g^{R_1}| \\ &< nR_0 \\ &< \frac{S}{N}, \end{aligned}$$

When $n\varepsilon > |g^{R_1}|$. Therefore for every s large enough, $|g^s| < s/N$. If we choose a value of R such that $p(R) > NR_0$, then by the first claim, $|g^{P(R)}| > R$. But the above says that $|g^{P(R)}| \leqslant P(R)/N \leqslant R$, which is a contradiction. This proves the second claim.

Finally, let γ be the image of $\langle g \rangle$ in X.



Let x and y be two points on γ. Now there exist a and b such that $d(x, g^{aN}) \leqslant N|g|$ and $d(y, g^{bN}) \leqslant N|g|$. So

$$d_\gamma(x, y) \leqslant N|b - a||g| + 2N|g| = N(|b - a| + 2)|g|.$$

On the other hand, $d(g^{aN}, g^{bN}) = d(g^{b-a}N, 1) \geqslant |b - a|$ from our second claim. This tells us that $d(x, y) \geqslant |b - a| - 2|g|N$, and so

$$d_\gamma(x, y) \leqslant N|g|d(x, y) + 2N^2|g|^2 + 2N|g|.$$

If we then let $N|g|$ be the multiplicative constant in the definition of a quasi-geodesic, and let $2N^2|g|^2 + 2N|g|$ be the additive constant, then we are done. □

Exercise 3.34

1. Consider the *Baumslag-Solitar group*

$$G = \langle x, y \mid xyx^{-1} = y^2 \rangle.$$

   Show that G has an exponential isoperimetric inequality. Thus G is not hyperbolic.

2. Show that in G we have $x^k y x^{-k} = y^{2^k}$ and that $\langle y \rangle$ is infinite but not a quasi-geodesic. Thus G cannot be a subgroup of a hyperbolic group. Generalise to

$$G_{mn} = \langle x, y \mid xy^m x^{-1} = y^n \rangle \ m \neq n.$$

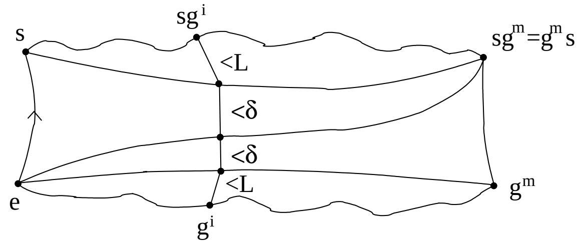
3. Use part 2 to show that if $\langle S \mid R \rangle$ is a Dehn presentation for G and w is a word in S where w represents an infinite order element then there exists $n > 0$ such that w is not conjugate to any element in G represented by a word of length less than $m = \max\{|r| \mid r \in E\}$. Use this to find an algorithm to decide whether or not a given word in S represents a word of infinite order in G.

3.8 Abelian Subgroups of Hyperbolic Groups.

Note that if g is an element of infinite order in a hyperbolic group then there exists $L > 0$ such that for all i and j and for any point X on the geodesic $[g^i, g^j]$ there exists k such that $d(x, g^k) < L$. (This is from the corresponding theorem about hyperbolic metric spaces.)

Theorem 3.35 *Let G be a hyperbolic group and suppose that $g \in G$ has infinite order. Then the centraliser $C(g)$ of g is a finite extension of $\langle g \rangle$, i.e. it is virtually cyclic.*

Proof. Let $\Gamma_S(G)$ be the Cayley graph of G with respect to the finite generating set S. Suppose that triangles in $\Gamma_S(G)$ are δ-thin. The image of $\langle g \rangle$ is a quasi-geodesic in $\Gamma_S(G)$. Let L be such that for all n, the geodesic segment $[1, g^n]$ lies in the L-neighbourhood of $\{1, g, \ldots, g^n\}$. Let s be in $C(g)$ and let m be such that $d(1, g^m) > 2|s| + 2\delta$.



Consider the rectangle $[1, g^m, sg^m, s]$. We split this lengthwise into two δ-thin triangles. There is a $p \in [1, g^m]$ such that $d(p, [s, sg^m]) \leqslant 2\delta$. Therefore there are powers g^i and g^j of g such that $d(g^i, sg^j) \leqslant 2L + 2\delta$, i.e. $g^i u = sg^j$ for some u with $|u| \leqslant 2L + 2\delta$. So $g^{i-j}u = s$, i.e. every coset $\langle g \rangle s$ has a representative of length at most $2L + 2\delta$. Thus $[C(g) : \langle g \rangle]$ is a finite extension. $\square$

Corollary 3.36 *Let G be a hyperbolic group and H be a subgroup of G containing at least one infinite order element. Then H is of the form $\mathbb{Z} \oplus A$ where A is a finite abelian group.*

Proof. In the abelian case, the centraliser of the infinite order element is the whole of H. $\square$

Corollary 3.37 *Abelian subgroups of hyperbolic groups are either finite or virtually cyclic.*

Thus $\mathbb{Z} \oplus \mathbb{Z}$ is not a subgroup of any hyperbolic group. We sometimes say that $\mathbb{Z} \oplus \mathbb{Z}$ is a *poison subgroup* for hyperbolicity.

3.9 Quasi-Convexity.

Definition 3.38 *A subset Y of a geodesic metric space X is called* ε-quasi-convex *if for all geodesics $[y_1, y_2]$ between any two points y_1 and y_2 of Y we have $[y_1, y_2] \subset B_\varepsilon^X(Y)$. We say that a subgroup H of a group G is* quasi-convex *with respect to a finite generating set S if there exists $\varepsilon > 0$ such that H is an ε-quasi-convex subset of $\Gamma_S(G)$.*

Example 3.39

1. Finite index or finite subgroups of any group are clearly quasi-convex.
2. Infinite cyclic subgroups of hyperbolic groups are quasi-convex by theorem 3.33.
3. If $K < H < G$ and K has finite index in H then K is quasi-convex in G if and only if H is quasi-convex in G. Thus virtually cyclic subgroups of hyperbolic groups are quasi-convex since they are finite extensions of quasi-convex subgroups.

Quasi-convexity of a subgroup of a hyperbolic group G does not depend on the finite generating set of G. Suppose that S and T are both finite generating sets of G and that H is a subgroup of G which is quasi-convex in $\Gamma_S G$. Then $\mathrm{id} : G \to G$ induces a quasi-isometry $\Gamma_S G \to \Gamma_T G$ by proposition 3.3. Thus this map sends geodesics to quasi-geodesics by exercise 2.25. Because quasi-geodesics are δ-close to geodesics in the Cayley graph for some δ (theorem 2.24), any geodesic in $\Gamma_S G$ joining two points in H will lie in $B_{\varepsilon+\delta}(H)$ in $\Gamma_T G$.

In contrast, for a general group G, quasi-convexity of a given subgroup depends on the set of generators of G.

Example 3.40 In $\mathbb{Z} \oplus \mathbb{Z}$ with the standard generating set a and b, the subgroup generated by $a+b$ is not quasi-convex with resepct to this generating set, as geodesics between two points of this subgroup can stray arbitrarily far outside the subgroup. However, if we add the generator $c = a+b$ then every geodesic between two points of $\langle c \rangle$ is unique and travels in the direction of c. Thus with respect to this generating set the subgroup *is* quasi-convex.

Example 3.41 Suppose that A is a finitely generated subgroup of a free group F where S is a finite generating set for F. Then A is quasi-convex. For let $T = \{a_1, \dots, a_k\}$ be a finite generating set for A, and for each i let α_i be a word in S representing a_i. Let h be in A and let $w(h)$ be the geodesic in the Cayley graph of $F(S)$ from e to h. Let $v = b_1, \dots, b_m$ be a reduced word in T (i.e. each b_j is some $a_p^{\pm 1}$) representing h. Then $\alpha(v) = \alpha(b_1) \cdots \alpha(b_m)$ also represents h and $\overline{w(h)} \subset \alpha(v)$ as omitting a point from the geodesic disconnects the tree. Thus each point c in the path w lies in the image $\alpha(a_j)$ of some generator a_j and so c is at a distance at most $\frac{1}{2}$ $max_j\{l(\alpha(a_j))\}$ from some vertex in A. As A is finitely generated this number is finite.

Example 3.42 Let S_1 and S_2 be finite generating sets for G_1 and G_2. Then G_1 and G_2 are quasi-convex subgroups of the free product $G_1 * G_2$ with respect to the finite generating set $S_1 \cup S_2$.

Proposition 3.43 *Let G be a finitely generated group. Then every quasi-convex subgroup G is finitely generated.*

Proof. Suppose that H is a ε-quasi-convex subgroup of a group G with respect to a finite generating set S of G. Let $w = a_1 \cdots a_n$ be a shortest representative word of an element of H, where $a_i \in S \cup S^{-1}$. By quasi-convexity of H, for all $i = 1, \dots, n$ there exists $v_i \in F(S)$ with $|v_i| < \varepsilon$ such that $\overline{a_1 \cdots a_i v_i} \in H$. Then $w = \prod_{i=1}^{n} v_i^{-1} a_i v_i$, where v_0 and v_n are the empty word. $v_i^{-1} a_i v_i$ is an element of H for each i and $|v_{i-1} a_i v_i| < 2\varepsilon + 1$. Thus the ball of radius $2\varepsilon + 1$ in the Cayley graph of G with respect to S is a finite generating set for H. □

Note that not all subgroups of hyperbolic groups are hyperbolic. First of all, they don't have to be finitely generated. For example, the commutator subgroup of the free group of rank 2 is the free group of countably infinite rank. The Rips construction (see exercise 3.32) shows that there are finitely generated subgroups of hyperbolic groups which are not hyperbolic. Furthermore, N. Brady [8] has given an example of a finitely *presented* non-hyperbolic subgroup H of a hyperbolic group G. In particular, H is a finitely presented group which is not hyperbolic yet has no Baumslag-Solitar subgroups, because G can't have any. So, even more specifically, while hyperbolic groups contain no $\mathbb{Z} \oplus \mathbb{Z}$- subgroups, the condition of not containing any $\mathbb{Z} \oplus \mathbb{Z}$-subgroups does not force a finitely presented group to be hyperbolic.

Theorem 3.44 *Quasi-convex subgroups of hyperbolic groups are hyperbolic.*

Proof. Suppose H is an ε-quasi-convex subgroup of the hyperbolic group G. Let S' be a finite set of generators for H, and consider the Cayley graph of G with respect to the generators $S \cup S'$. Consider the inclusion map $i : H \to G$. Then for h_1 and h_2 in H we have

$$d_{S \cup S'}(h_1, h_2) \leqslant d_{S'}(h_1, h_2) \leqslant (2\varepsilon + 1) d_{S \cup S'}(h_1, h_2).$$

i.e. this inclusion map is a quasi-isometry. H is then hyperbolic since hyperbolicity is a quasi-isometry invariant. □

Proposition 3.45 *Quasi-convex normal subgroups of hyperbolic groups have finite index.*

Proof. Let G be a hyperbolic group with S a finite set of generators of G. Suppose that N is an ε-quasi-convex normal subgroup of G. If G is finite then there is nothing to prove. On the other hand, If G is infinite then N must also be infinite. Let cN be a coset of N and let

$u \in N$ be such that $|u| > 2\delta + 2|c|$. Consider the geodesic rectangle $[1, c, u, uc]$. There are vertices $p \in [1, u]$ and $q \in [c, uc]$ with $d(p, q) \leqslant 2\delta$ by thinness of triangles in G (we split the rectangle into two such triangles). Since N is quasi-convex, there are n_1 and n_2 in N such that $d(p, n_1) \leqslant \varepsilon$ and $d(q, cn_2) \leqslant \varepsilon$. So there is a word w on S with $|w| \leqslant 2\varepsilon + 2\delta$ such that $n_1\overline{w} = cn_2$. Thus $n_1\overline{w} = (cn_2c^{-1})c$. Since N is normal, $cn_2c^{-1} \in N$ and we have $Nc = N\overline{w}$. Every coset hence has a representative of size at most $2\varepsilon + 2\delta$. G/N is hence finite. □

Example 3.46 If G is a finitely generated free group then all of its finitely generated subgroups are quasi-convex. Hence the above theorem tells us that a non-trivial finitely generated normal subgroup of a finitely generated free group is of finite index, a fact originally due to Schreier (see [LS] for a generalisation to arbitrary free groups).

It is often possible to use geometric methods to shorten proofs in combinatorial group theory. Howson proved the following theorem in [31].

Theorem 3.47 *If F is a finitely generated free group and H and K are finitely generated subgroups of F then $H \cap K$ is finitely generated.*

We now give a proof of Howson's theorem using quasi-convexity. This is due to Short [62].

Definition 3.48 *A group G satisfies the* Howson property *if given any two finitely generated subgroups of G their intersection is also finitely generated.*

Example 3.49 $F_2 \times \mathbb{Z} = \langle a, b\rangle \times \langle z\rangle$ does not satisfy the Howson property. Let $A = \langle a, b\rangle$ and $B = \langle a, bz\rangle$. Then as z commutes with a and b, a word is in the intersection if and only if the exponent sum of z in w (total number of occurrences of z counted according to sign) is 0. The exponent sum of z in a word is the image of the element represented under the map $G \to \mathbb{Z}$ which sends a and b to 0. But in a word written in terms of generators of B, the exponent sum of z is the same as that of b. Thus $A \cup B = \langle b^n a b^{-n}\rangle$ which is infinitely generated.

Proposition 3.50 *Let G be a group generated by the finite set S and let A and B be subgroups which are quasi-convex with respect to S. Then $A \cap B$ is quasi-convex with respect to S.*

Proof. Let $w = a_1 \cdots a_n$ be a geodesic word for an element $H \in A \cap B$. Let K_A and K_B be quasi-convexity constants for A and B respectively. Then for each j there exist words γ_j and γ_j' such that $l(\gamma_j) \leqslant K_A$, $l(\gamma_j') \leqslant K_B$, $a_1 \cdots a_j\gamma_j$ represents an element of A and $a_1 \cdots a_j\gamma_j'$ represents an element of B. Let N be the number of different such pairs $(\gamma_j, \gamma_j') \in G \times G$. If $l(w) > N$ then for some $1 \leqslant i < j \leqslant N$ we have $\gamma_i = \gamma_j$ and $\gamma_i' = \gamma_j'$ In this case,

$$\begin{aligned}(a_1 \cdots a_i\gamma_i)(\gamma_j^{-1}a_{j+1} \cdots a_n) &= a_1 \cdots a_i a_{j+1} \cdots a_n \\ &= (a_1 \cdots a_i\gamma_i')(\gamma_j'^{-1}a_{j+1} \cdots a_n)\end{aligned}$$

and this is an element of $A \cap B$. Continuing in this way, we eventually obtain a word of length less than N. Thus $A \cap B$ is N-quasi-convex. □

Theorem 3.47 now follows immediately from proposition 3.50, proposition 3.43 and example 3.41.

Hanna Neumann refined Howson's theorem further, to prove the following [38].

Theorem 3.51 *If A and B are finitely generated subgroups of a free group then*

$$\mathit{rank}(A \cap B) - 1 \leqslant 2(\mathit{rank}(A) - 1)(\mathit{rank}(B) - 1).$$

It was conjectured (and is still not proved) that

$$\mathrm{rank}(A \cap B) - 1 \leqslant (\mathrm{rank}(A) - 1)(\mathrm{rank}(B) - 1).$$

However, Burns [11] has refined Hanna Neumann's bound to

$$\mathrm{rank}(A \cap B) - 1 \leqslant 2(\mathrm{rank}(A) - 1)(\mathrm{rank}(B) - 1) - \min\{\mathrm{rank}(A) - 1, \mathrm{rank}(B) - 1\}.$$

There is an extensive literature on this problem.

3.10 The Boundary of a Hyperbolic Group.

Definition 3.52 *If G is a hyperbolic group, then the* boundary *of G, ∂G is defined to be ∂X, where X is a Cayley graph of G.*

This is well defined because of the following result.

Proposition 3.53 *If X and Y are quasi-isometric hyperbolic metric spaces then ∂X is homeomorphic to ∂Y.*

Lemma 3.54 *Given a point $x \in \partial G$ we can represent x by a geodesic sequence of points.*

Proof. Suppose X is a Cayley graph of G and $x \in \partial X$ and let $\{x_n\}$ be any sequence in X with $x_n \to x$. We associate to each x_n a geodesic $\gamma_n = [e, x_n]$. As X is a locally finite graph and $x_n \to \infty$, a subsequence of γ_n converges to an infinite geodesic. More explicitly, let $\gamma_n(k)$ be the k^{th} vertex of γ_n. Then there exists infinitely many γ_n with the same $\gamma_n(1)$. We can therefore pass to a subsequence $\{\gamma_n^{(1)}\}$ such that every value of $\gamma_n^{(1)}(1)$ is the same. Similarly we define a subsequence $\{\gamma_n^{(2)}\}$ such that every value of $\gamma_n(2)^{(2)}$ is the same. We continue this process to obtain sequences $\gamma_n^{(m)}$, each a subsequence of the previous one, such that every value of $\gamma_n^{(n)}$ is the same. The diagonal sequence $\{\gamma_n^{(n)}\}$ then converges pointwise to an infinite geodesic γ. We now claim that $\lim_{i\to\infty}(\gamma(i)) = x$. This requires that $\lim_{i,j\to\infty}((x_i \cdot \gamma(j)) = \infty$. But $\gamma(i) \in [e, x_j]$ for some j, and

$$(x_j \cdot \gamma(i)) = \frac{1}{2}\left(d(x_j, e) + d(\gamma(i), e) - d(x_j, \gamma(i)\right) = d(e, \gamma(i)) = i.$$

So $\gamma(i) \to x$. $\square$

Theorem 3.55 *If G is a hyperbolic group with Cayley graph X then $\hat{X}$ is compact.*

Proof. $\hat{X}$ is regular (recall that a topological space Y is called *regular* if for all $y \in Y$ and for all closed subsets K of $Y - \{y\}$ there exists an open neigbourhood of U not containing y.) and it also has a locally countable base (check) so it is metrisable. Thus to show that it is compact we show that every sequence has a convergent subsequence.

Let $\{x_n\}$ be a sequence in $\hat{X}$. If $\{x_n\}$ has infinitely many points in X we pass to a subsequence, which we again denote by $\{x_n\}$, such that $x_n \in X$ for all n. In this case, if $\{x_n\}$ is bounded then it obviously has a convergent subsequence so we may assume that $\{x_n\}$ has a subsequence which tends to infinity. Then as in the proof of the lemma we consider $[e, x_n]$ and find a subsequence converging to the infinite geodesic γ.

Otherwise $\{x_n\}$ has only finitely many points in X. We can hence assume by passage to a subsequence that $x_n \in \partial X$ for all n. For all n, pick a geodesic γ_n such that $\lim_i(\gamma_n(i)) = x_n$. By passing to a subsequence we may assume that γ_n tends to a limit, γ say, and that $\gamma_n(i) = \gamma(i)$ for all $i \leqslant n$. We claim that $\lim(x_n) = \lim_i(\gamma(i))$, which we call x. Now by property 6 of the basic properties of the boundary of a hyperbolic metric space, since $\gamma_n(i) \to x_n$ and x is equal by definition to $\lim_i(\gamma(i))$, we have

$$(x_n \cdot x) \geqslant \liminf_i \left(\gamma_n(i) \cdot \gamma(i)\right) - 2\delta,$$

which is greater than or equal to $n - 2\delta$. Hence $\lim_n((x_n \cdot x)) = \infty$. We have now shown that $\hat{X}$ is compact. $\square$

Example 3.56 The boundary of a virtually cyclic group is a pair of points. If G is virtually (free of rank at least two) then ∂G is a Cantor set. If G is a lattice in a semisimple Lie group then ∂G is a sphere.

If a group G acts on a hyperbolic metric space X then G acts on sequences in X as follows. Let g be in G and (x_n) be a sequence in X. Then we define $g((x_n)) = (g(x_n))$. By definition, if (x_n) converges to infinity then so does $g(x_n)$. Also if (x_n) and (y_n) are sequences in X with $(x_n \cdot y_n) \to \infty$ as $n \to \infty$ then $(g(x_n) \cdot g(y_n)) \to \infty$. Thus the above gives a well defined action of G on ∂G. Since the action of a group G on a hyperbolic metric space X extends to an action by isometries of G on ∂X, a hyperbolic group G acts on its boundary by isometries. Note that if g is an infinite order element of G then we have two sequences $\{g^n | n > 0\}$ and $\{g^n | n < 0\}$. Suppose that $g^n \to a^+$ and $g^{-n} \to a^-$ in ∂G. Then a^+ and a^- are fixed points of g. Since the image of $\{g^n | n \in \mathbb{Z}\}$ is a quasi-geodesic in the Cayley graph (i.e. if we "join the dots" by geodesics) we have $a^+ \neq a^-$.

Theorem 3.57 *If X is a geodesic proper hyperbolic metric space then there exists a continuous surjection $\partial X \to \mathit{ends}(X)$ whose fibres are the connected components of ∂X.*

For example, for a free group the map is a bijection. Note that theorem 3.57 also has an analogue for CAT(0) spaces. There is a sketch proof of theorem 3.57 in [19]. For more details, see [44]. We can deduce from ends theory that the boundary of a hyperbolic group must have either have cardinality 0,1 or 2 or be uncountable. Hyperbolic groups are either finite (0-ended), properly virtually infinite cylic (2-ended) or non-elementary (either 1-ended or with an uncountable number of ends).

Exercise 3.58 Let G be δ-hyperbolic and let X be its Cayley graph. Let a and b be points of ∂X. Show that there exists a bi-infinite geodesic γ such that

$$\lim_{i \to \infty} (\gamma(i)) = a \text{ and } \lim_{i \to \infty} (\gamma(-i)) = b.$$

(This is also true in a general *proper* geodesic metric space. Here the proof depends on Ascoli's theorem.)

3.11 The Tits Alternative for Hyperbolic Groups.

The name "Tits alternative" comes from the following theorem due to Tits (see [65]).

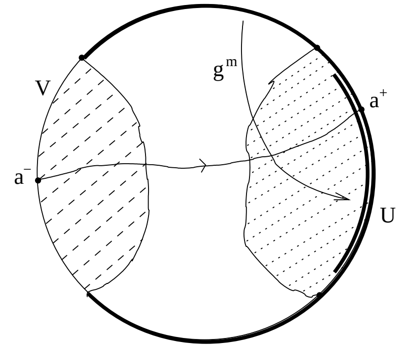
Theorem 3.59 *Let G be a finitely generated linear group. Then either G is virtually solvable or contains a free subgroup of rank* 2

More generally, any dichotomy about a class of groups of the form "either G contains F_2 as a subgroup or ..." is called a *Tits alternative* for that class of groups. We obtain such a dichotomy for the class of hyperbolic groups by considering the action of such a group on its boundary.

We say that an action of a group G by isometries on a hyperbolic metric space X is *parabolic* if there exists a point in ∂X which is fixed by every element of G.

Definition 3.60 *A hyperbolic group G is called* elementary *if it is finite or the action of G on ∂G is parabolic.*

Proposition 3.61 *Let U and V be neighbourhoods of two points a^+ and a^- on the boundary of a hyperbolic group G which are fixed by $g \in G$. Then there exists $N \in \mathbb{N}$ such that for all $m > N$ we have $g^m(\partial G - V) \subset U$ and $g^{-m}(\partial G - U) \subset V$.*



Proof. Let X be a Cayley graph of G and let x be a point of $\partial X - V$. Then there exists a sequence $\{x_n\}$ converging to x with $\liminf_n((x_n \cdot g^{-n})) \leqslant K$ for some $K > 0$. We want to find an integer m such that $g^m(x) \in U$. It's enough to find m such that $\liminf_n(g^m(x_n) \cdot g^n) \geqslant M$ where $N_M(a^+) \subset U$. By elementary property 9 of the boundary of a hyperbolic metric space, there exists $N > 0$ such that for all $m > 0$ and $n > 0$, if $(x_n \cdot g^{-m}) \leqslant K + \delta$, then

$$\frac{1}{2}\left(d(x_n, e) + d(g^{-m}, e) - d(x_n, g^{-m})\right) \leqslant K + \delta,$$

which implies (noting that $d(g^m(x_n), e) = d(x_n, g^{-m})$)

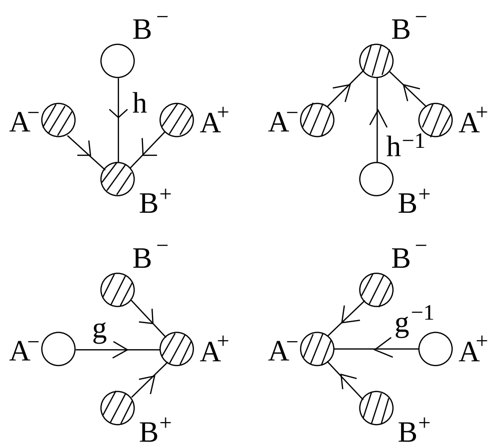
$$\frac{1}{2}\left(d(x_n, g^{-m})\right) \geqslant \frac{1}{2}\left(d(x_n, e) + d(g^{-m}, e)\right) - (K + \delta).$$

So we have

$$\begin{aligned}(g^m(x_n) \cdot g^n) &\geqslant \frac{1}{2}(d(x_n, e) + d(g^{-m}, e) + d(g^n, e) \\ &\quad - d(x_n, g^{n-m})) - (K + \delta) \\ &= \frac{1}{2}(d(g^n, e) + d(g^{-m}, e) - d(g^{n-m}, e)) \\ &\quad + (x_n \cdot g^{n-m}) - (K + \delta),\end{aligned}$$

and by taking m large enough, $(g^m(x_n) \cdot g^n) \geqslant M$. □

Lemma 3.62 (Tits' Ping-Pong Lemma) *Let G be a group generated by two elements r and s. Let G act on a set X and let A^+,A^-,B^+ and B^- be disjoint subsets of X such that $r(A^+ \cup B^- \cup B^+) \subset A^+$, $s(B^+ \cup A^- \cup A^+) \subset B^+$, $r^{-1}(A^- \cup B^- \cup B^+) \subset A^-$ and $s^{-1}(B^- \cup A^- \cup A^+) \subset B^-$. Then G is isomorphic to the free group on two generators.*



Proof. It is enough to show that only the trivial word on r and s is equal to the identity in G. Let $w = r^{n_1}s^{n_2}\cdots r^{n_{k-1}}s^{n_k}$ where $|n_1| \geqslant 0$,$|n_k| \geqslant 0$ and $|n_i| > 0$ for $1 < i < k$. We distinguish 4 cases, depending on whether or not we have an r or an s at either end of w.

1. $n_1 \neq 0$ and $n_2 \neq 0$: Assume $n_k > 0$. Then if $x \in B^+$, $wx \in B^+$ as $r^{n_1}(s^{n_2} \cdots s^{n_k}x) \in A^+ \cup A^-$. So $w \neq 1$. Similarly if $n_k < 0$ and $x \in B^-$, then $wx \neq B^-$.

2. $n_1 = 0$ and $n_k \neq 0$: Similarly we can pick a point $x \in A^+$ or A^- such that $wx \notin A^+ \cup A^-$ so $w \neq 1$.

3. $n_1 \neq 0$ and $n_k = 0$: Pick $x \in B^+ \cup B^-$ such that $wx \notin B^+ \cup B^-$.

4. $n_1 = 0$ and $n_k = 0$: Pick $x \in A^+ \cup A^-$ such that $wx \in A^+ \cup A^-$.

In each case, the result follows. □

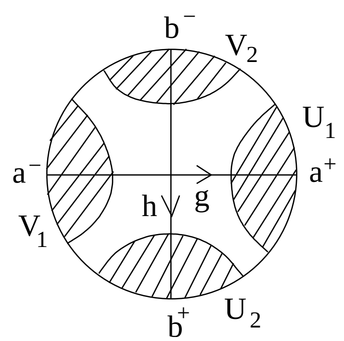
Proposition 3.63 *Let G be a hyperbolic group and $g \in G$ a non-torsion element. Let a be a point of ∂G fixed by G. Then under the action of G on ∂G, the stabilizer Stab(a) is a finite extension of $\langle g \rangle$, i.e. it is virtually cyclic.*

Proof. Let c be a shortest coset representative in $\mathrm{Stab}(a)/\langle g \rangle$. We show that $|c|$ is bounded. Assume $a = \lim(g^n)$. Then $\lim(cg^n) = a$. Thus $\lim((cg^n.g^n)) = \infty$. If c_1^n, c_2^n and c_3^n are the internal points of the triangle $[e, cg^n, g^n]$, we have that $\lim(d(e, c_1^n)) = \infty$ and $d(c_1^n, c_2^n) \leqslant \delta$ for all n. Fix n to be sufficiently large. As the image of $\{g^m | m \in \mathbb{Z}\}$ is a quasi-geodesic in the Cayley graph of G, there exists $M > 0$ such that $d(c_1^n, g^k) \leqslant M$ and $d(c_2^n, cg^m) \leqslant M$ for some integers k and m. So $d(g^k, g^m) \leqslant 2M + 2\delta$. By increasing n we can find $k_1 > k_2$, $m_1 > m_2$ and a word w such that $cg^{m_1} = g^{k_1}w$ and $cg^{m_2} = g^{k_2}w$. But this implies that for all m there exists n such that $d(g^n, cg^m) \leqslant 2M + 2\delta$. Thus $|c| \leqslant 2M + 2\delta$. □

Thus if G is elementary and $a \in \partial G$ is fixed by G then $\mathrm{Stab}(a) = G$ and G must be virtually cyclic. Therefore a hyperbolic group is elementary if and only if it is either finite or virtually cyclic.

Theorem 3.64 *Let G be a hyperbolic group and let g_1 and g_2 be non-torsion elements of G. Then either $\langle g_1, g_2 \rangle$ is virtually cyclic or there exists an integer $k > 0$ such that $\langle g_1^k, g_2^k \rangle$ is a free group of rank 2.*

Proof. Let Γ be the Cayley graph of G, let a^+ and a^- be the fixed points of g_1 and let b^+ and b^- be the fixed points of g_2. If $\{a^+, a^-\} \cap \{b^+, b^-\}$ is nonempty then $\langle g_1, g_2 \rangle$ is virtually cyclic by the previous proposition. If not there exist open neighbourhoods $a^+ \in U_1$, $a^- \in V_1$,$b^+ \in U_2$ and $b^- \in V_2$ which are mutually disjoint. So there exists an integer k with $g_1^k(U_1 \cup U_2 \cup V_2) \subset U_1$, $g_2^k(U_2 \cup U_1 \cup V_1) \subset U_2$, $g_1^{-k}(V_1 \cup U_2 \cup V_2) \subset V_1$ and $g_2^{-k}(V_2 \cup U_1 \cup V_1) \subset V_2$.



Hence, by Tits' lemma, $\langle h^k, g^k \rangle \cong F_2$. □

This generalises in the obvious way to subgroups generated by any finite number of elements. In fact, Delzant has shown the following in [17].

Theorem 3.65 *Let G be a torsion free hyperbolic group. Then G only contains a finite number of conjugacy classes of subgroups not freely generated by two elements.*

Furthermore, Rips and Sela have given a proof in [50] of the following result, first stated in [23].

Theorem 3.66 *Let G be a hyperbolic group and let H be a finitely presented torsion free freely indecomposable noncyclic subgroup. Then G contains finitely many conjugacy classes of subgroups isomorphic to H.*

Recall that a hyperbolic group only has finitely many conjugacy classes of torsion elements. In contrast, Gromov has proved the following in [23].

Theorem 3.67 *In a nonelementary hyperbolic group there are infinitely many conjugacy classes of primitive elements.*

It follows from theorem 3.67 that infinite torsion groups can't be hyperbolic.

4 Canonical Representatives and Equations In Hyperbolic Groups.

In this section we present some work of Rips and Sela [49, 57, 50, 58].

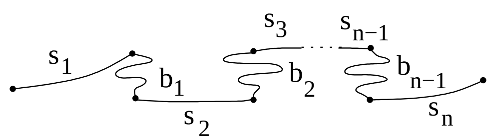
4.1 Coarse Geodesics

The following type of path, due to Rips, lies at the heart of the construction of canonical representatives. Throughout, assume that we are in a δ-hyperbolic graph X.

Proposition 4.1 *[23] Let γ be a geodesic between vertices v_1 and v_2 of X. Then there exists a constant $K_{\delta,\lambda}$ such that any λ-quasigeodesic q between v_1 and v_2 lies within $N_{K_{\delta,\lambda}}(\gamma)$.*

Let $\rho_{\delta,\lambda} = 8(2\delta + K_{\delta,\lambda})$ and let $\mu_{\delta,\lambda} = 500\delta\rho_{\delta,\lambda}$. If λ and δ are understood, we shall omit the subscripts. (The reason we have so defined these constants will become apparent in the proof of the short-circuiting lemma below.)

Definition 4.2 *Let $L \in \mathbb{Z}$ with $L \geqslant \mu_{\delta,\lambda}$. An L-*coarse geodesic *$(c, \{s_1, b_1, s_2, b_2, \ldots, b_{n-1}, s_n\})$ is a $\rho_{\delta,\lambda}$-quasigeodesic c together with a partition of subpaths of c where $s_1, \ldots, s_n$ are $\mu_{\delta,\lambda}$-local geodesics such that $l(s_i) \geqslant L$ for all integers i with $1 \leqslant i \leqslant n-1$ and $b_1, \ldots, b_{n-1}$ are paths such that $l(b_i) \leqslant 2K_{\delta,\lambda} + 2\delta$ for each i. The s_i are called* sublocal geodesics *and the b_i* bridges*. L is called the* criterion *of c.*



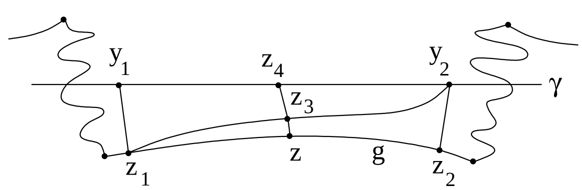
Note that if $n = 1$ then there is no condition on the length of the sublocal geodesic s_1. Thus, in particular, all geodesics are coarse geodesics. See the picture above. We think of the s_i as very long and not very flexible, whereas the b_i are very short and flexible. (Think of the s_i as planks of wood and the b_i as springs!) If Y is a path in a metric space X, then by $N_K(Y)$ we mean the set $\{x \in Y \mid d(x, Y) \leqslant K\}$.

Proposition 4.3 *[23] In a δ-hyperbolic metric space, for any $\lambda \in \mathbb{Z}$, every $(1000\delta, \lambda)$-local quasigeodesic is a 2λ-quasigeodesic.*

These above results will be used in the proofs of the following two propositions about coarse geodesics. In what follows, "sufficiently far from" will mean a distance greater than or equal to $\frac{1}{2}\mu_{\delta,\lambda}$ from.

Proposition 4.4 *Let γ be a geodesic between vertices v_1 and v_2 of X and let c be a coarse geodesic between v_1 and v_2 with g a sublocal geodesic of c. If $z \in g$ is sufficiently far from the endpoints of g then $d(z, \gamma) \leqslant 2\delta$.*

Proof. Let $\rho = \rho_{\delta,\lambda}$. By proposition 4.1, c stays within a distance $K_{\delta,\rho}$ of γ. Let $z_1, z_2 \in g$ satisfy $d(z, z_1) = d(z, z_2) = \frac{1}{2}\mu_{\delta,\lambda}$. Then $[z_1, z_2]_g$ is a geodesic by $\mu_{\delta,\lambda}$-locality of g. Let y_1 and y_2 be points on γ with $d(y_1, z_1) \leqslant K_{\delta,\rho}$ and $d(y_2, z_2) \leqslant K_{\delta,\rho}$. Then consider the rectangle $[z_1, y_1, y_2, z_2]$ which we can split into two triangles.



By thinness of the triangle $[z_1, z_2, y_2]$, $d(z, z_3) \leqslant \delta$, and by thinness of $[z_1, y_2, y_1]$, $d(z_3, z_4) \leqslant \delta$, so $d(z, z_4) \leqslant 2\delta$, i.e. $d(z, \gamma) \leqslant 2\delta$. □

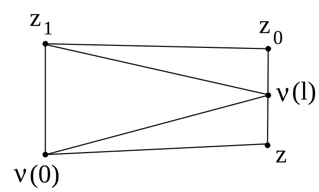
This result, which allows us to "short-circuit" coarse geodesics, is very important in the proof of the cylinder theorem later.

Proposition 4.5 (Short-circuiting Lemma) *Suppose c is an L-coarse quasigeodesic between vertices v_1 and v_2 of X and γ is a geodesic between v_1 and a third vertex v_3. Let g be a sublocal geodesic of c and suppose $z \in g$ is sufficiently far from $t(g)$ and that $l([i(g), z]_c) \geqslant L$. Let z_0 be a closest point to z on γ, and suppose that this is less than $2K_{\delta,\lambda} + 2\delta$ away. Then*

$$c' = [v_1, z]_c * [z, z_0] * [z_0, v_3]_\gamma$$

is an L-coarse geodesic.

Proof. We can consider a geodesic from z to z_0 as a bridge because it is no longer than $2K_{\delta,\lambda} + 2\delta$. As z is no less a distance than L from $i(g)$, the path $[i(g), z]$ is still a valid sublocal geodesic. Thus to show that c' is a coarse geodesic amounts to showing that it is a $\rho_{\delta,\lambda}$-quasigeodesic. By proposition 4.3, we only need to show that it is a $(1000\delta, 8\delta + 4K_{\delta,\lambda})$-local quasigeodesic. For this it is necessary to show that any subpath of c' of length no greater than $1000\delta(8\delta + 4K_{\delta,\lambda})$ is a $(8\delta + 4K_{\delta,\lambda})$-quasigeodesic, and by the symmetry of the following argument, it suffices to show this for any such subpath ν which starts on g and finishes on $[z, z_0]$. Assume that ν is not a path of length zero, because the result is trivial in this case. Let z_1 be a point on γ which is no further that $4\delta + 2K_{\delta,\lambda}$ from $i(\nu)$



Such a point exists because the previous lemma ensures that $i(\nu)$ is at most a distance 2δ from a geodesic segment between v_1 and v_2. A closest point to $i(\nu)$ on this segment is in turn at most a distance $2\delta + 2K_{\delta,\lambda}$ from γ by the geometry of geodesic triangles. By the fact that $L \geqslant \mu_{\delta,\lambda}$ and by $\mu_{\delta,\lambda} - \textit{locality}$ of g, $[\nu(0), z]_g$ is a geodesic. So we have

$$d(\nu(0), z) \leqslant 6\delta + 4K_{\delta,\lambda} + d(z_1, z_0).$$

Also, by the triangle inequality,

$$d(z_1, z_0) \leqslant d(z_1, \nu(1)) - d(\nu(1), z_0) \text{ and so}$$
$$d(\nu(0), z) \leqslant 6\delta + 4K_{\delta,\lambda} + d(\nu(0), \nu(1)) + d(\nu(1), z_0).$$

Hence we have

$$\begin{aligned} d(\nu(0), z) + d(z, \nu(1)) &\leqslant 6\delta + 4K_{\delta,\lambda} + d(\nu(0), \nu(1)) + d(\nu(1), z_0) + d(z, \nu(1)) \\ &= 8\delta + 4K_{\delta,\lambda} d(\nu(0), \nu(1)) \text{ i.e.} \\ l(\nu) &\leqslant (8\delta + 4K_{\delta,\lambda}) d(\nu(0), \nu(1)) \text{ (as } d(\nu(0), \nu(1) \geqslant 1). \end{aligned}$$

Thus c' is a $(1000\delta, 8\delta + 4K_{\delta,\lambda})$ -local quasigeodesic, and we are done. □

4.2 Cylinders

Definition 4.6 *Let X be the Cayley graph of a hyperbolic group Γ. A vertex $v \in X$ is called an* L-elector *of a vertex w of X if v lies on a path p between the identity e of Γ and w such that $[v, w]_p$ and $[v, e]_{p^-}$ are L-coarse geodesics and the concatenation of the opposite of the first $\mu_{\delta,\lambda}$-sublocal geodesic of $[v, e]_p$ and the first of $[v, w]_{p^-}$ is a $\mu_{\delta,\lambda}$-local geodesic. The set of all L-electors of w is called the* L-cylinder *of w, written $C_L(w)$.*

Proposition 4.7 *Every elector $v \in C_L(w)$ is 2δ-close to a geodesic segment $[e, w]$.*

Proof. Suppose that both of the coarse geodesics $[v, w]_p$ and $[v, e]_{p^-}$ in the definition of an elector have two or more sublocal geodesics (the first being of length greater than or equal to L). Then as v is sufficiently far from the endpoints of the $\mu_{\delta,\lambda}$-local geodesic it lies on (because $L \geqslant \mu_{\delta,\lambda}$), it is 2δ-close to $[e, w]$ by proposition 4.4.

On the other hand, one of these coarse geodesics may consist of only one sublocal geodesic. In this case, as mentioned by the note after the definition of a coarse geodesic, this sublocal geodesic isn't enforced to have length greater than or equal to L. If only one of them consists of one sublocal geodesic g, and $l(g) \geqslant \frac{1}{2}\mu_{\delta,\lambda}$, then the result follows as before. If $l(g) < \frac{1}{2}\mu_{\delta,\lambda}$, then as the length of the first sublocal geodesic h of the other coarse geodesic is greater than or equal to $\mu_{\delta,\lambda}$, there exists a point z on h with $d(z, [e, w]) \leqslant 2\delta$. By taking the triangle formed by e,z and z', where z' is a closest point of $[e, w]$ to z, noting that the triangle is geodesic by $\mu_{\delta,\lambda}$-locality of the concatenation of g and h^- and using the thinness of triangles in a hyperbolic metric space, we see that $d(v, [e, w]) \leqslant 2\delta$.

If both of the coarse geodesics have only one sublocal geodesic, where both of these sublocal geodesics have length less than $\frac{1}{2}\mu_{\delta,\lambda}$, then by $\mu_{\delta,\lambda}$-locality, of the combined path, it must be a geodesic. The result then follows from thinness of geodesic bigons (a special case of triangles). Otherwise the result follows as in the previous two cases. □

Note that as a group acts by isometries on its Cayley graph by left multiplication, we have that $v \in C_L(w) \iff w^{-1}v \in C_L(w^{-1})$. We shall call this the *invertibility* property of cylinders.

Denote the number of elements of a finite set X by $n(X)$.

Definition 4.8 *Let w be a vertex of the Cayley graph X and let v be an elector in $C_L(w)$. Then the* left *and* right neighbourhoods *of v in $C_L(w)$ are, respectively*

$$N_L^w(v) = \{x \in C_L(w) \mid d(e, x) \leqslant d(e, v) \text{ and } d(v, x) \geqslant 10\delta\} \text{ and}$$

$$N_R^w(v) = \{x \in C_L(w) \mid d(e, x) \geqslant d(e, v) \text{ and } d(v, x) \geqslant 10\delta\}$$

We define the difference *between two electors u and v of $C_L(w)$ to be*

$$\begin{aligned} \mathit{diff}_w(u, v) &= n(N_L^w(u) - N_L^w(v)) - n(N_L^w(v) - N_L^w(u)) + n(N_R^w(v) - N_R^w(u)) - \\ &\quad n(N_R^w(u) - N_R^w(v)). \end{aligned}$$

Lemma 4.9 *If A, B and C are any finite sets, then*

$$n(A - B) - n(B - A) + n(B - C) - n(C - B) = n(A - C) - n(C - A)$$

Proof. Put $n(A - B) = n((A - B) - C) + n((A - B) \cap C)$ etc. and expand. □

Lemma 4.10 *For all electors u, v and t in $C_L(w)$,*

1. $\mathit{diff}_w(u, v) + \mathit{diff}_w(v, t) = \mathit{diff}_w(u, t)$

2. $\mathit{diff}_w(u,v) = -\mathit{diff}_w(v,u)$

Proof.

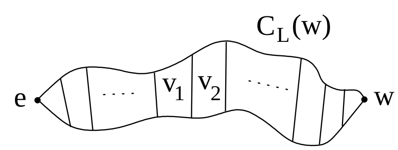
1. We have

$$\begin{aligned}
\mathrm{diff}_w(u,v) + \mathrm{diff}_w(v,t) &= n(N_L^w(u) - N_L^w(v)) - n(N_L^w(v) - N_L^w(u)) + n(N_R^w(v) - N_R^w(u)) - \\
&\quad n(N_R^w(u) - N_R^w(v)) + n(N_L^w(v) - N_L^w(t)) - n(N_L^w(t) - N_L^w(v)) + \\
&\quad n(N_R^w(t) - N_R^w(v)) - n(N_R^w(v) - N_R^w(t)) \\
&= n[(N_L^w(u) - N_L^w(v)) - n(N_L^w(v) - N_L^w(u)) + n(N_L^w(v) - N_L^w(t)) - \\
&\quad n(N_L^w(t) - N_L^w(v))] + n[(N_R^w(v) - N_R^w(u)) - n(N_R^w(u) - N_R^w(v)) + \\
&\quad n(N_R^w(t) - N_R^w(v)) - n(N_R^w(v) - N_R^w(t))] \\
&= n(N_L^w(u) - N_L^w(t)) - n(N_L^w(t) - N_L^w(u)) + n(N_R^w(t) - N_R^w(u)) - \\
&\quad n(N_R^w(u) - N_R^w(t)) \text{ (by lemma 4.9)} \\
&= \mathrm{diff}_w(u,t)
\end{aligned}$$

2. This follows from the last part if we put $t = u$.

□

Definition 4.11 *If w is a vertex of a Cayley graph and v is an elector of w, then the* slice *of v with respect to w, $\mathit{slice}_v^w = \{x \in C_L(w) \mid \mathit{diff}_w(x,v) = 0\}$. We say that $\mathit{slice}_{v_2}^w$ is* consecutive *to $\mathit{slice}_{v_1}^w$ if $\mathit{diff}_w(v_2,v_1) > 0$, and for all $u \in C_L(w)$, either $\mathit{diff}_w(v_1,u) \geqslant 0$ or $\mathit{diff}_w(u,v_2) \geqslant 0$.*



Proposition 4.12 *The notion of consecutivity between slices is well defined.*

Proof. Suppose that $u_1 \in \mathrm{slice}_{v_1}^w$, $u_2 \in \mathrm{slice}_{v_2}^w$, and $\mathrm{slice}_{v_1}^w$ is consecutive to $\mathrm{slice}_{v_2}^w$. Then $\mathrm{diff}_w(v_2,v_1) > 0$, and for every $u \in C_L(w)$, either $\mathrm{diff}_w(v_1,u) \geqslant 0$ or $\mathrm{diff}_w(u,v_2) \geqslant 0$. Now by definition of u_1 and u_2, $\mathrm{diff}_w(u_1,v_1) = 0$ and $\mathrm{diff}_w(u_2,v_2) = 0$. Hence, by the first part of the previous lemma, $\mathrm{diff}_w(u_1,u_2) = \mathrm{diff}_w(u_1,v_2)$. But we also have

$$\mathrm{diff}_w(v_1,v_2) + \mathrm{diff}_w(v_2,u_1) = \mathrm{diff}_w(v_1,u_1) = -\mathrm{diff}_w(u_1,v_1)$$

Hence

$$\mathrm{diff}_w(v_1,v_2) = -\mathrm{diff}_w(v_2,u_1) = \mathrm{diff}_w(u_1,v_2) = \mathrm{diff}_w(u_1,u_2).$$

Thus we can say that

$$\mathrm{diff}_w(v_2,v_1) > 0 \iff \mathrm{diff}_w(u_2,u_1) > 0.$$

Also suppose that for all electors u of w, we have either $\mathrm{diff}_w(v_2,u) \geqslant 0$ or $\mathrm{diff}_w(u,v_1) \geqslant 0$. Then we have to show that either $\mathrm{diff}_w(u_2,u) \geqslant 0$ or $\mathrm{diff}_w(u,u_1) \geqslant 0$. But this follows again simply from the last lemma, as applying it to u_2, v_2 and u gives $\mathrm{diff}_w(v_2,u) = \mathrm{diff}_w(u_2,u)$. Similarly, if $\mathrm{diff}_w(u,v_1) \geqslant 0$, then $\mathrm{diff}_w(u,u_1) \geqslant 0$. So the notion of consecutivity is independent of which representatives of the slices we pick, and is hence well defined. □

Slices have been introduced to partition a cylinder into easily manageable subsets, upon which consecutivity gives us a partial ordering. In particular, the slices can be thought of as "small":

Proposition 4.13 *1. For all electors $v \in C_L(w)$, $\mathit{diam}(\mathit{slice}_v(w)) \leqslant 10\delta$*

2. *If $slice^w_{v_2}$ is consecutive to $slice^w_{v_1}$ and $|v_1| \geqslant 10\delta$, then*

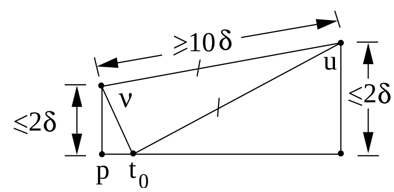
$$diam(slice^w_{v_1} \cup slice^w_{v_2}) \leqslant 20\delta + 1$$

Proof.

1. Suppose $u \in \text{slice}^w_v$ with $|u| \geqslant |v| \geqslant 10\delta$ and $d(u,v) > 10\delta$. If $\text{slice}^w_u = \text{slice}^w_v$, then $\text{diff}_w(u,v) = 0$. Thus

$$n(N_L(u) - N_L(v)) - n(N_L(v) - N_L(u)) = n(N_R(u) - N_R(v)) - n(N_R(v) - N_R(u))$$

As $|u| > |v|$ and $d(u,v) > 10\delta$, it follows from the definitions of left and right neighbourhoods that $N^w_L(v) \subset N^w_L(u)$ and $N^w_R(u) \subset N^w_R(v)$. Thus $\text{diff}_w(u,v) = n(N_L(u) - N_L(v)) + n(N_R(v) - N_R(u))$. As u and v are electors, they are 2δ-close to a geodesic segment $\gamma = [e,w]$. Let $t_0 \in \gamma$ satisfy $d(t_0,u) = d(u,v)$ and $|t_0| \leqslant |u|$.



By triangle $[t_0, v, p]$, we have $d(t_0, v) \leqslant 2\delta + d(t_0, p)$, i.e. $d(t_0, v) \leqslant d(t_0, p) + 2\delta$, giving

$$10\delta \leqslant d(t_0, u) \leqslant d(t_0, q) + 2\delta \Rightarrow d(t_0, q) \geqslant 8\delta.$$

So we have

$$d(t_0, p) = d(p, q) - d(t_0, q) \leqslant 10\delta - 8\delta = 2\delta.$$

This tells us that $d(t_0, v) \leqslant 2\delta + 2\delta = 4\delta$, i.e. $t_0 \notin N^w_L(v)$ but $t_0 \in N^w_L(u)$. Hence

$$\text{diff}_w(u,v) \geqslant 1 + n(N_R(v) - N_R(u)) \geqslant 1$$

which contradicts the fact that u and v are in the same slice. Thus we must have $d(u,v) \leqslant 10\delta$.

2. Suppose $\gamma = [e,w]$ is a geodesic segment. Then any vertex in γ is an elector in $C_T(w)$. Assume that $d(v_2, v_1) > 20\delta + 1$ and let $t \in \gamma$ satisfy $|v_1| < t$ and $d(t, v_1) = 10\delta + 1$. By the first part, we know that $t \notin \text{slice}^w_{v_1}$, so $\text{diff}_w(t, v_1) > 0$. But if $d(v_2, v_1) \geqslant 20\delta + 2$, then

$$d(v_2, t) \geqslant (20\delta + 2) - (10\delta + 1) = 10\delta + 1.$$

So $t \notin \text{slice}^w_{v_2}$ either, so it must lie strictly between $\text{slice}^w_{v_1}$ and $\text{slice}^w_{v_2}$ with respect to the partial ordering on slices. This contradicts the consecutivity of the two slices that we started with.

□

It is this last result about slices which allows us to define canonical representatives.

4.3 Canonical Representatives

Suppose Γ is a torsion-free hyperbolic group and let $F(\Gamma)$ denote the free group on the generators of Γ. Denote the natural quotient map from $F(\Gamma)$ to Γ by $\natural$. In this section we construct *combings* θ_L of Γ, i.e. maps from Γ to $F(\Gamma)$ such that $\natural \circ \theta_L$ is the identity map on Γ. This means that to every group element of Γ we assign a word equal to that element.

To every subset A of the 10δ-ball $N_{10\delta}(e)$ in the Cayley graph X of Γ, we assign the complete graph

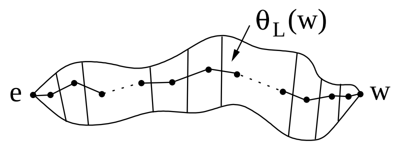
K_A on the vertices of A and colour the graph according to group elements joining the endpoints of an edge (these are words, not just generators). We say that K_A is equivalent to K_B if there exists a colour-preserving isomorphism between them. This is obviously an equivalence relation on the set of all such subsets of the 10δ-ball, and we call an equivalence class under this relation an *atom*. To each atom a, we arbitrarily assign a *centre*, which is an element of the vertex set. We have seen in the previous section that a slice is bounded in diameter by 10δ, so it is the image under left multiplication of some subset of the 10δ-ball, which corresponds uniquely to some atom. Hence we can assign a centre to a slice, $\mathrm{ce}(\mathrm{slice}_v^w)$ by taking the centre of this atom.

Next we want a canonical way to travel between centres of slices. To each vertex x in the $(20\delta+1)$-ball of X, we assign a geodesic segment $\mathrm{st}(x) = [e, x]$, called the *step* of x, such that $[\mathrm{st}(x)]^{-1} = st(x^{-1})$.

Definition 4.14 *Let w be a vertex in the Cayley graph of the torsion-free hyperbolic group Γ and let $\{slice_{v_1}^w, \ldots, slice_{v_\nu=w}^w\}$ be the sequence of consecutive slices of the L-cylinder $C_L(w)$. If $|w| > 10\delta$, then we define the L-*canonical representative *of w to be*

$$\theta_L(w) = st(ce(slice_{v_1}^w)) * (*_{i=2}^{\nu(w)} st(ce(slice_{v_{i-1}}^w)^{-1} ce(slice_{v_i}^w))) * st(ce(slice_w^w)^{-1} w)$$

If $|w| \leqslant 10\delta$, we define $\theta_L(w) = st(w)$.



Proposition 4.15 $[\theta_L(w)]^{-1} = \theta_L(w^{-1})$.

Proof. Because of the invertibility of cylinders, we have $C_L(w^{-1}) = w^{-1}C_L(w)$ and so the cylinder of w^{-1} is a left translate of that of w. Because the cylinder determines the slices, and we have chosen the steps such that $[\mathrm{st}(x)]^{-1} = st(x^{-1})$ for all x, the result follows. □

Let $C_L(v, w) = vC_L(v^{-1}w)$, i.e. the cylinder "between" v and w, and similarly let $\theta_L(v, w)$ denote the corresponding left-translated canonical representative. We shall also make use of the idea of the *canonical representative* of a finite subset Y of the Cayley graph X. This is defined to be

$$\theta_L(Y) = \bigcup_{y_i, y_j \in Y} \theta_L(y_i, y_j).$$

4.4 The Quasitree Property

This is the first main result and is quite technical to set up. It is this property which allowed Sela to prove the solvability of the isomorphism problem for torsion-free hyperbolic groups with no essential small action on a real tree, but with slight modification also implies the result on solutions of equations as in [49].

Let Y be a finite subset of a metric space X and let $n = |Y|$. An *approximating tree* $\mathrm{Tr}(Y)$ for Y is a union of at most $(n-1)$ geodesics such that every point of Y lies on this union.

Theorem 4.16 (Gromov) *There exists $K > 0$ such that if X is δ-hyperbolic then there exists an approximating tree for Y such that for all y_1 and y_2 in Y, the unique path along the tree between y_1 and y_2 is a $100\delta\log_2(n)$-quasi-geodesic.*

We shall assume from now on that the value λ, upon which our definition of a coarse geodesic depends, is equal to $100\delta\log_2(n)$. Throughout we shall call the number of points in a subset of the Cayley graph its *volume* and denote by $v_{2\delta}(X)$ the volume of the 2δ-ball around any vertex of the

Cayley graph (which is constant because the Cayley graph is a homogeneous metric space, i.e. a group acts *transitively* on its Cayley graph by isometries.

The following theorem justifies the definition of canonical representatives. Although slightly more general than Sela's version, its proof is almost the same.

Theorem 4.17 (Cylinder Theorem) *Let $Tr(Y)$ be as above and let $\{v_i\}_{i=1}^{s}$ be its set of nodes. Then there exists an integer R such that if we let $\{Y_a\}_{a=1}^{\lambda}$ be the set of connected components of*

$$Tr(Y) - \bigcup_{i=1}^{s} N_R(v_i)$$

and let $M_a = N_{100\delta}(Y_a)$ for each a, then there exists a criterion L such that for all $1 \leqslant a \leqslant \lambda$ and for all $1 \leqslant i_1, i_2, j_1, j_2 \leqslant n$,

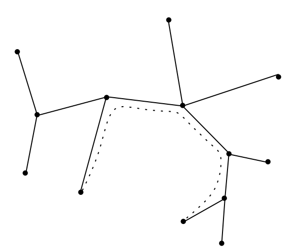
$$\begin{aligned} &C_L(y_{i_1}, y_{j_1}) \cap M_a \neq \emptyset \text{ and } C_L(y_{i_2}, y_{j_2}) \cap M_a \neq \emptyset \\ &\qquad \Rightarrow C_L(y_{i_1}, y_{j_1}) \cap M_a = C_L(y_{i_2}, y_{j_2}) \cap M_a \end{aligned}$$

We call such a criterion a *nice* criterion for the tree.

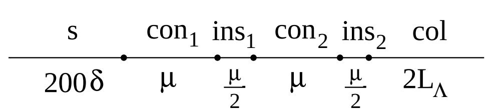
Proof. First, we show that given any fixed a, i_1, i_2, j_1 and j_2, there exists a criterion satisfying the above. Then a pigeon-hole argument is used to show that we can pick just one criterion so that the desired equality holds for all a, i_1, i_2, j_1 and j_2.

We shall also assume for the geometry of the proof that $i_1{=}i_2$ and simply call it i. We do not lose generality here, because to prove the case where the two cylinders start at distinct vertices, we apply what follows twice, the second time in the opposite direction, noting the invertibility of cylinders.

We define a finite set $\{L_i\}_{i=1}^{\Delta}$ of criteria, at least one of which will have the required property. Let $\Delta = 2\lambda n^4 \mathrm{Ca}(\mu) + 1$ and let $L_m = 2\mu(1+2m)$ for $1 \leqslant m \leqslant \Delta$. Then we define R as in the statement of the theorem to be $2L_\Delta + 3\mu + 100\delta$. Let η_1 and η_2 be geodesic segments between y_i and y_{j_1}, y_i and y_{j_2} respectively. Note that by proposition 4.1, these segments remain "close" to the part of the tree between these points.



This point is important when visualising the following proof. Let x and y be points 2δ-close to the endpoints of the geodesic segment Y_a, where x is closer to y_i than y_{j_1}. We can assume that η_1 and η_2 travel through Y_a in the same direction, because of the invertibility of cylinders. Let η' be the shortest path (a broken geodesic) along the tree from y_i to y_{j_1}, and consider the segment from y to the node y' after η' passes through Y_a. We partition this into segments as follows, where con stands for *conductor*, ins for *insulator* and col for *collector*.



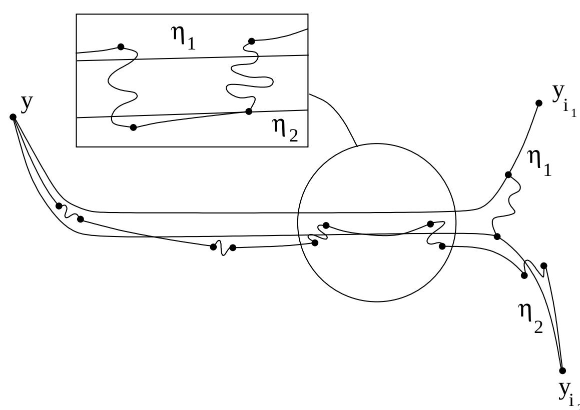
Let $\gamma = [e, x]$ be a geodesic in X. A geodesic not shorter than γ in $N_{2\delta}(\gamma)$ is called a *channel* of γ. The *μ-capacity* of X, $\mathrm{Ca}_X(\mu)$ is the maximum number of different channels of a geodesic of length μ. Note that $\mathrm{Ca}_X(\mu) \leqslant 2^{\mu v_{2\delta}(X)}$ as this bounds the number of subsets of a 2δ-neighbourhood of a geodesic of length μ.

If $C_L(y_i, y_{j_1}) \cap M_a \neq \emptyset$ and $C_L(y_i, y_{j_2}) \cap M_a \neq \emptyset$, but $C_L(y_i, y_{j_1}) \cap M_a \neq C_L(y_i, y_{j_2}) \cap M_a$, then we shall call L a *bad* criterion. In this case there must exist an elector $u_L \in C_L(y_i, y_{j_1})$ such that $u_L \notin C_L(y_i, y_{j_2})$, or vice versa. Our assumption during the proof will be that the former is true, and this is what we shall seek to contradict in the following claims. We then refer to u_L as a *bad* elector. As u_L is an elector we know that there exists a path β_L through u_L comprising of two coarse geodesics with the properties described in the definition of an elector. We call β_L a *witness* for the bad elector.

Claim 4.18 *If L_m is a bad criterion for which u_m is a bad elector with witness β_m, then β_m occupies a channel of either con$_1$ or con$_2$ which is neither con$_1$ nor con$_2$.*

Proof. β_m must occupy a channel of one of these conductors, because the enforced minimal length a sublocal geodesic ensures that if a bridge occurs while β_m passes through $N_{2\delta}(\text{con}_1)$ then no bridge can occur while it passes through $N_{2\delta}(\text{con}_2)$, and vice versa. Now the μ-locality of a sublocal geodesic guarantees that it occupies a geodesic of length greater than or equal to μ whilst passing through the relevant neighbourhood.

Assume w.l.o.g. that β_m occupies a channel of con$_2$ and suppose that this channel is con$_2$ itself. Then construct a path β'_m by travelling along β_m, through u_m and after passing through con$_2$, travel a distance L along η'_1 and form a bridge to η_2 which is a distance of at most $2\delta + 2K_{\delta,\lambda}$ away. This is because the coarse geodesic is at most 2δ from η_1, which is at most $2K_{\delta,\lambda}$ away from η_2 by twice applying proposition 4.1, where the quasigeodesics in this proposition are paths along the tree, firstly from y_i to y_{j_1} and secondly from y_i to y_{j_2}. Then continue along η_2 to y_{j_2}. This modified path is an L- coarse geodesic by the short-circuiting lemma, and thus forms the latter half of a path comprised of two coarse geodesics of the required type for u_m to be an elector in $C_L(y_i, y_{j_2})$. This contradicts the fact that u_m is a bad elector.

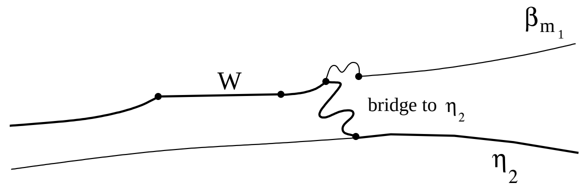


□

Next we aim to bound the number of bad criteria which may have bad electors with witnesses occupying the same channel. Suppose that L_{m_1} and L_{m_2} are two distinct bad criteria from our sequence (with $m_1 < m_2$) having bad electors u_{m_1} and u_{m_2} with witnesses β_{m_1} and β_{m_2} respectively, where β_{m_1} and β_{m_2} occupy the same channel W of con$_2$. Let ν_1 and ν_2 be the sublocal geodesics of β_{m_1} and β_{m_2} passing through W and denote by exc(ν_i) for $i = 1$ and 2 the *excess* of ν_i, i.e. the remaining length after passing through W.

Claim 4.19 *exc$(\nu_i) < L_{m_i}$ for $i = 1$ and 2.*

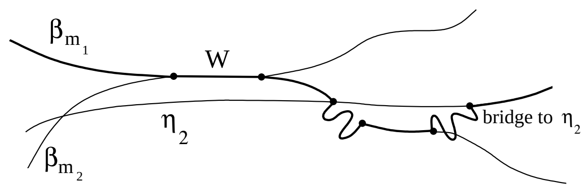
Proof. Suppose not, and say it is exc(ν_1) which is greater than or equal to L_{m_i}. Then we can modify β_{m_1} by passing through ν_1 a distance of $L_{m_1} - \mu$ and continuing to y_{j_2} along η_2 by adding a bridge as in the proof of the last claim.



This contradicts the fact that u_{m_1} is a bad elector. □

Claim 4.20 *$exc(\nu_2) < exc(\nu_1)$.*

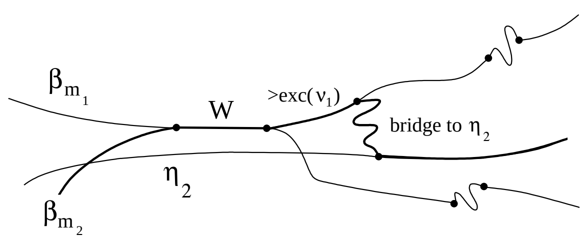
Proof. Suppose to the contrary that $\mathrm{exc}(\nu_2) \geqslant \mathrm{exc}(\nu_1)$. Then because the length of W is greater than or equal to μ, the path formed by going along ν_1 before W, through W, then out along ν_2 is still a μ- local geodesic of length greater than or equal to L_{m_1}, by our initial assumption. After this we continue along β_{m_2} through the next bridge, and we can travel a distance of $L_{m_1} < L_{m_2}$ along the *next* sublocal geodesic and form a bridge to η_2 as in fig. 2.11 to obtain a contradiction to u_{m_1} being a bad elector again.



□

Claim 4.21 *$exc(\nu_1) - exc(\nu_2) < \frac{1}{2}\mu$.*

Proof. If not, then we can modify β_{m_2} this time by continuing along β_{m_1} after emerging from W. After travelling a distance of $\mathrm{exc}(\nu_1)$ then we have certainly travelled a distance of L_{m_2} along a sublocal geodesic (we can do this by our assumption). As we are a greater distance than $\frac{1}{2}\mu$ from the other end of ν_1, we can form a bridge to η_2.



This contradicts the fact that u_{m_2} is a bad elector. □

Claim 4.22 *There are at most $2\mu Ca_X(\mu)$ bad criteria for fixed a, i_1, i_2, j_1, j_2.*

Proof. The 2 comes from the fact that we had to apply the arguments of the last four claims twice, as previously mentioned. Both of these times we fix a channel W of con_2 and take a corresponding $\mathrm{exc}(\nu_1)$. Then if L_{m_2} is distinct from L_{m_1}, $\mathrm{exc}(\nu_2)$ must be different. If we keep taking L_{m_i} which have not previously occurred in the sequence, then $\mathrm{exc}(\nu_i)$ must not have previously occurred in the sequence either. As the difference in these excesses, which must be integers, is bounded by $\frac{1}{2}\mu$, there are at most $\frac{1}{2}\mu$ different bad criteria for the given channel W. As there are at most $2\mathrm{Ca}_X(\mu)$ different such channels, (remember that we could be occupying a channel of either con_1 or con_2) there are at most $\mu\mathrm{Ca}_X(\mu)$ bad criteria altogether. □

To conclude the proof of the cylinder theorem, we now apply a pigeonhole argument. Apply the above claim $2\lambda s^4$ times, once for each choice of a, i_1, i_2, j_1 and j_2. As we have not enough bad criteria altogether to exhaust our sequence of criteria, there must be one which is simultaneously not bad for all of these choices. □

By a (B,m)-*quasitree* we mean a graph whose closed paths (i.e. paths with the same initial and terminal vertex) are all contained in at most m balls of radius B. We shall now see that for a finite subset Y of the Cayley graph, with repect to a nice criterion L for a tree $\mathrm{Tr}(Y)$, $\theta_L(Y)$ is a $(\beta, n(Y)(n(Y)-2))$-quasitree for some global bound B.

Theorem 4.23 (The Quasitree Property.) *There exists a global bound B, integers $t_{i,j}$ for $1 \leqslant i,j \leqslant n$ and sets of paths*

$$\{d_a\}_{a=1}^{\lambda} \text{ and } \{c_{i,j,b}\}_{1\leqslant i,j\leqslant n}^{0\leqslant b\leqslant t_{i,j}}$$

Such that with respect to the criterion of the cylinder theorem,

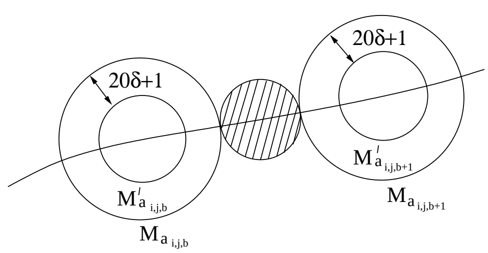
$$\theta_L(y_i,y_j) = c_{i,j,0} * d^{\pm 1}_{a_{i,j,1}} * c_{i,j,2} * \cdots * d^{\pm 1}_{a_{i,j,t_{i,j}}} * c_{i,j,t_{i,j}}$$

Where $l(c_{i,j,b}) \leqslant B$ for each i,j and b.

We call the d_a *distinguished paths* and the $c_{i,j,b}$ *auxiliary paths.*

Proof. Repeat the following argument for each i,j,a. Let M_a denote the same sets as in the cylinder theorem, and let M'_a be the set $M_a - N_{20\delta+1}(X - M_a)$. We know from the cylinder theorem that any two cylinders are the same in M_a between any pairs of points. As $20\delta+1$ bounds the distance between consecutive slices of a cylinder, then we know that the sequences of slices of these two cylinders are identical in M'_a, as they are determined by the distance from the starting point, and thus in here the canonical representatives are the same. So we can let $d_{a_{i,j,b}} = \theta_L(y_i,y_j) \cap M'_{a_{i,j,b}}$ for all those $a_{i,j,b}$ such that the relevant cylinder actually meets $M_{a_{i,j,b}}$, where $a_{i,j,b}$ ranges between 1 and λ and b indexes the $M'_a s$ our cylinder meets, ranging from 1 to $t_{i,j}$. Then by the cylinder theorem the set $\{d_a\}$ has the required properties (the ± 1 in the statement of the theorem takes into account that we could be travelling along the section of the canonical representative in M'_a in one of two directions.

We now must find the global bound on the size of the auxiliary paths. The diameter of a connected component of the union of the R- neighbourhoods of the nodes of $Tr(Y)$ is bounded by $2R[n+(n-2)] = 4(n-1)R$ (as the tree is a union of at most $(n-1)$ geodesics and so we have at most $(n-2)$ forks). To make the geometry a bit simpler, suppose w.l.o.g. that we are travelling neither to nor from a leaf of the tree, because the bound which we shall derive will suffice for this case also. Between $M'_{a_{i,j,b}}$ and $M_{a_{i,j,b+1}}$, our canonical representative must stay within 2δ of a geodesic between y_i and y_j

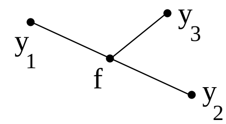


The maximum length of the segment γ of this geodesic between $M'_{a_{i,j,b}}$ and $M'_{a_{i,j,b+1}}$ is $4(n-1)R + 2(20\delta+1)$. As the canonical representative must always remain within 2δ of this geodesic and the number of points in $N_{2\delta}(\gamma)$ is loosely bounded by $[4(n-1)R + 2(20\delta+1)]v_{2\delta}(X)$, then the worst possible case for the length of an auxiliary path is where we take each of these vertices and join them by a step. This has length $[4(n-1)R+2(20\delta+1)](20\delta+1)v_{2\delta}(X)$. Thus we are done if we let

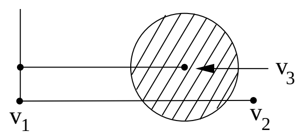
$$B = [4(n-1)R + 2(20\delta+1)](20\delta+1)v_{2\delta}(X)$$

□

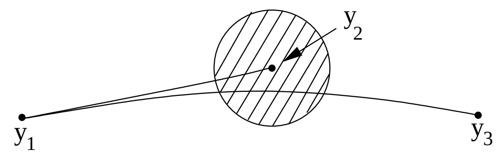
In the next section we shall want a corollary of this last result in the case where $n = 3$, i.e. when we are dealing with a triangle. Then the only possible tree we can have is like this:



Here, the only auxiliary paths we need are in the R-neighbourhood of the fork f. The reason we also included R-neigbourhoods in the cylinder theorem was in case the following situation occurred:

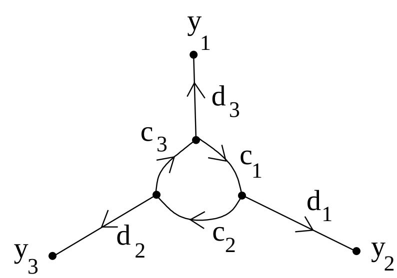


i.e. the canonical representative from y_1 to y_2 passes through an R-neighbourhood of y_3. This could only happen for a triangle if we had the combinatorial type as folows:



However, this is not a valid tree because y_1 is not a leaf. Hence the appropriate version of the quasitree property for a triangle is the following, which will imply the solvability of systems of equations.

Corollary 4.24 *Suppose y_1, y_2 and y_3 are three points in the Cayley graph of a torsion-free hyperbolic group. Then there exists a criterion L, a global bound B and paths $\{d_i\}_{i=1}^3$ and $\{c_j\}_{j=1}^3$ such that $\theta_L(y_1, y_2) = d_3^{-1} c_1 d_1$, $\theta_L(y_2, y_3) = d_1^{-1} c_2 d_2$ and $\theta_L(y_3, y_1) = d_2^{-1} c_3 d_3$.*



i.e. The canonical representative of Y is a $(B, 1)$-quasitree.

Note that in fact, what is needed in the next section is slightly more general, that we can pick one criterion which is simultaneously nice for q triangles. Clearly this is a trivial generalisation of the proof of theorem 4.17 as we only need to redefine Δ as $2\lambda q n^4 \mathrm{Ca}_X(\mu) + 1$.

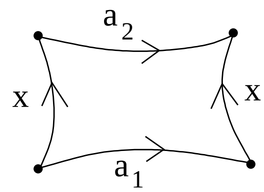
4.5 Systems of Equations in Hyperbolic Groups

Throughout this section, $\Gamma = \langle G \mid R \rangle$ will be a finitely presented torsion-free δ-hyperbolic group, with Cayley graph X. $x_1, \ldots, x_n$ will denote variables in Γ and $a_1, \ldots, a_n$ constants in Γ. A *system of equations* will be a finite set of equalities of the form

$$\Phi = \{\phi_j(x_1, \ldots, x_n, a_1, \ldots, a_k) = e\}_{j=1}^q$$

where each ϕ_j is a finite product of elements of $\{x_1^{\pm 1}, \ldots, x_n^{\pm 1}, a_1^{\pm 1}, \ldots, a_k^{\pm 1}\}$. If each ϕ_j is a product of exactly three such elements then the system is called *triangular*. For any system ϕ over a group with no 2-torsion we can construct an equivalent triangular system.

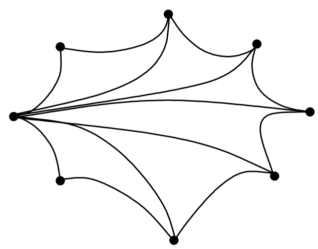
Example 4.25 (The Conjugacy Problem) Given $a_1, a_2 \in \Gamma$, is there an algorithm to decide if a_1 is conjugate to a_2? i.e. does there exist an algorithm to decide if there exists a solution to the equation $a_1 x a_2 x^{-1} = e$? Pictorially, in the Cayley graph X we have the outer quadrilateral



which we can triangulate by introducing the extra unknown z, giving the equivalent triangular system

$$
\begin{aligned}
a_1 x z &= e \\
x a_2 z &= e
\end{aligned}
$$

Similarly if Γ has no 2-torsion, then every circuit in the Cayley graph will be represented by a polygon of at least three sides, and we can split it up into triangles as follows:



We shall therefore assume from now on that our systems are triangular. Makanin and Razborov have independently shown [35, 46] that there exists an effective algorithm to decide whether a system of equations in a free group has a solution. We show the existence of such an algorithm for torsion-free hyperbolic groups by constructing a finite set of systems $\Psi(\Phi)$ in $F(\Gamma)$ such that

1. Every solution of a system in Ψ gives rise to a solution of Φ.

2. Every solution of Φ gives rise to a solution of at least one of the systems of Ψ.

The finiteness of the set of systems Ψ then enables us to decide if the system in the hyperbolic group has a solution by exhaustively applying the method for a free group to each system of Ψ. Then (1) above shows that if a system in Ψ has a solution, then Φ has a solution, and (2) shows that if Φ has a solution then we detect it somewhere in checking all of the systems in Ψ. Thus Φ has a solution if and only if there exists a system in Ψ with a solution. Hence there will exist an effective algorithm to decide whether a system of equations in a torsion-free hyperbolic group has a solution.

We shall denote a triangular system of equations Φ with q equations in it as $\Phi = \{z_{i(j,1)} z_{i(j,2)} z_{i(j,3)} = e\}_{j=1}^{q}$, where $i : \{1, \ldots, q\} \times \{1, 2, 3\} \to \{\pm 1, \ldots, \pm l\}$ is an indexing function, $l = n + k$ and $z_1 = x_1, \ldots, z_n = x_n$, $z_{n+1} = a_1, \ldots, z_l = a_k$, with the understanding that $z_{-i} = z_i^{-1}$. Thus z_i is a variable for $1 \leqslant i \leqslant n$ and a constant for $n+1 \leqslant i \leqslant l$. Suppose that $(w_1, \ldots, w_l)$ is a solution of Φ ($w_i = a_i$ for $n+1 \leqslant i \leqslant l$). Then for any j the points $w_{i(j,1)}$, $w_{i(j,1)} w_{i(j,2)}$ and $w_{i(j,1)} w_{i(j,2)} w_{i(j,3)} = e$ form a thin triangle if we join them by geodesics. This thinness, however, does not suffice to show that there exists an algorithm to solve equations. What we need to use is the quasitree property for a triangle (corollary 4.24).

If we consider joining the above mentioned points not by geodesics but by the canonical representatives with respect to a nice criterion, using the same notation as before for auxiliary and distinguished paths, we have the auxiliary paths forming a globally bounded loop $c_1c_2c_3$. We can then use the solution of the word problem in hyperbolic groups to effectively produce all possible sets of triples $\{c_1, c_2, c_3\}$ which could form this loop. Suppose that for a particular choice we have $\{c^j_{1,s}, c^j_{2,s}, c^j_{3,s}\}$.To construct the corresponding system in the free group, we equate first of all those terms in Φ with the same index, i.e. for our variables where $1 \leqslant i(j,a) \leqslant n$, we have

$$z_{i(j,a)} = z_{i(j',a')} \text{ if } i(j,a) = i(j', a').$$

Then we replace $z_{i(j,a)} = z_{i(j',a')}$ by the following. As $z_{i(j,a)}$ is a variable, we have:

$$z_{i(j,a)} = y^j_a c^j_{a,s} (y^j_{a+1,s})^{-1} \text{ and } z_{i(j',a')} = y^{j'}_{a'} c_{a',s} (y^{j'}_{a'+1})^{-1}$$

With the understanding that y^j_4 is a modulo 3 integer plus 1. Thus equating these we have

$$y^j_a c^j_{a,s} (y^j_{a+1})^{-1} = y^{j'}_{a'} c_{a',s} (y^{j'}_{a'+1})^{-1}$$

If $i(j,a) = i(j', a')$ and $1 \leqslant i(j,a) \leqslant n$.
Now if $z_{i(j,a)}$ is a constant this time, then in the system over the free group, we have

$$y^j_a c^j_{a,s} (y^j_{a+1})^{-1} = \theta_L(z_{i(j,a)})$$

for $n < i(j,a) \leqslant l$, where L is our nice criterion from the cylinder theorem. We don't, however, know *a priori* what $\theta_L(z_{i(j,a)})$ is, but we can certainly try constructing the canonical representatives constructed from all possible connected subsets of a 2δ neighbourhood of some geodesic $[e, z_{i(j,a)}]$.

So to construct our finite set of equations $\Psi(\Phi)$ with the desired properties mentioned at the beginning of this section, let r index all possible pairs

$$\{(\{c^j_a\}^{1 \leqslant j \leqslant q}_{1 \leqslant a \leqslant 3}, \{A_t\}^l_{t=n+1})_r \mid 1 \leqslant r \leqslant Q\}$$

where A_t is a connected subset of $N_{2\delta}([e, z_t])$. This is a finite set, and each pair gives rise to a system of equations Ψ_r. Thus we have constructed a set of systems over the free group $\Psi = \{\Psi_r\}^Q_{r=1}$ with the desired properties, and we have the following result,. This is interesting in its own right as it generalises the solvability of the word and conjugacy problems in hyperbolic groups. But it also plays a major part in the initial steps of Sela's proof of the solvability of the isomorphism problem. chapter.

Theorem 4.26 *Given a system of equations over a torsion-free hyperbolic group, it is decidable whether the system has a solution.*

Example 4.27 This should clarify how the system over the free group is constructed. Suppose that we have made a choice for $\{c^j_a\}^{1 \leqslant j \leqslant q}_{1 \leqslant a \leqslant 3}$ and the canonical representatives $\{\theta_L(z_t)\}^l_{t=n+1}$. Recall the conjugacy problem of the last example. In our new notation, x becomes z_1, z becomes z_2, a_1 becomes z_3 and a_2 becomes z_4. The indexing function i has values as in the following table.

$i(j,a)$	1	2 j
1	3	1
2	1	4
3	2	2
a		

For x we have $z_{i(1,2)} = z_{i(2,1)}$, giving

$$y^1_2 c^1_2 y^1_3 = y^2_1 c^2_1 y^2_2. \tag{3}$$

For z we have $z_{i(1,3)} = z_{i(2,3)}$, giving

$$y_3^1 c_3^1 y_1^1 = y_3^2 c_3^2 y_1^2. \tag{4}$$

And for a_1 and a_2 the equations which we obtain are

$$y_1^1 c_1^1 y_2^1 = \theta_L(a_1) \tag{5}$$

$$y_2^2 c_2^2 y_3^2 = \theta_L(a_2) \tag{6}$$

5 Formal Language Theory and Automatic Groups.

There is a large interaction between geometric group theory and the computer-scientific theory of formal languages. The definitive work on automatic groups is [18]. Here we give an introduction.

5.1 Regular Languages

Definition 5.1 *A* finite state automaton (F.S.A.) *is a quintuple* (S, A, μ, Y, s_0) *where* S *is a finite set called the* set of states*,* A *is a finite* alphabet*,* $\mu : S \times A \to S$ *is a function called the* transition function*,* Y *is a subset of* S *called the set of* accept states *and* $s_0 \in S$ *is called the* start (initial) state*.*

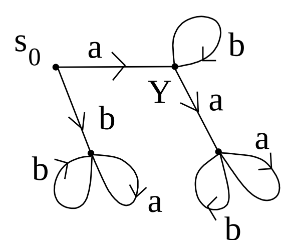
For $s \in S$ and $a \in A$ one usually writes sa instead of $\mu(s, a)$. We say that a word w is *accepted* by the finite state automaton if the sequence of transitions corresponding to w leads to an accept state, starting at s_0. If M is a finite state automaton then the *language* $L(M)$ *accepted by* M is the set of words in A which are accepted by M. A language L which is accepted by some finite state automaton is called a *regular language.*

Example 5.2 $L = \{(ab)^n \mid n \geqslant 0\}$ is a regular language. $L = \{a^n b^n \mid n \geqslant 0\}$ is not.

Exercise 5.3 Let G be a group with an ordered set of generators. Denote by $\mathcal{G}(G)$ the language of all geodesic words in G and denote by $\mathcal{SL}(G)$ the language (a combing) of all *short-lex* geodesics in G, that is, geodesics which are shortest under the lexicographic ordering on the set of words in G inherited from the ordering on the generators. Show that if the $\mathcal{G}(G)$ is regular, then so is $\mathcal{SL}(G)$ for any ordering of the generators.

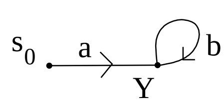
We associate to a F.S.A. M an oriented graph. The set of vertices are its set of states. There is an oriented edge from s_1 to s_2 labelled by $a \in A$ if and only if $\mu(s_1, a) = s_2$. The language of M is then the set of words corresponding to paths in this graph starting at s_0 and ending at an accept state.

Example 5.4 Let $A = (a, b)$ and let M be the finite state automaton given by the following graph.



The only designated accept state is Y. Then $L(M) = \{ab^n \mid n \geqslant 0\}$.

We call a state a *dead state* if there is no path from it to an accept state. When we draw the graph corresponding to a finite state automaton, we usually omit the dead states and just draw the following. For example, instead of the last graph we would just draw the following.



Note that we can have states which are neither dead nor accept states.

We now give an alternative characterisation of regular languages. By the *concatenation* KL of two languages K and L over an alphabet S, we just mean the set of all strings we can obtain by writing a string in K followed by a string in L. The *Kleene closure operation* $K*$ of a language K is defined by

$$K^* = \cup_{n \geqslant 0} K^n,$$

where $K^0 = \{\varepsilon\}$ and for all $n > 0$, $K^n = K^{n-1}K$.

Definition 5.5 *If S is an alphabet then a* rational language $L \subset S^*$ *is a language which can be obtained from the set of singleton subsets of X by a finite number of concatenations, unions and Kleene closure operations.*

Theorem 5.6 (Kleene) *A language is rational if and only if it is regular.*

See [30] for a proof.

5.2 The Chomsky Hierarchy.

More generally than regular languages, we have the *Chomsky hierarchy* of languages, i.e. each property in the list implies the following one: Regular, *context free*, *indexed*, *context sensitive* and *recursively enumerable*. These languages are accepted respectively by a finite state automaton, a *pushdown automaton*, a *nested stack automaton*, a *linear bounded automaton* and a *Turing machine*.

Example 5.7 $\{a^n b^m c^r | n, m, r \in \mathbb{N}\}$ is a regular language.
$\{a^n b^n c^m | n, m \in \mathbb{N}\}$ is a context-free language.
$\{a^n b^n c^n | n \in \mathbb{N}\}$ is an indexed language.

Definition 5.8 *A* pushdown automaton *(*PDA*) is a sextuple $P = (Q, \Sigma, \Omega, \delta, q_0, Z_0)$ where Q is a finite set called the* set of states*, Ω is a set called the* stack alphabet*, Σ is a set called the* alphabet*, $q_0 \in Q$ is the* initial state*, $Z_0 \in \Omega$ is the* start symbol *and δ is a mapping $Q \times \Omega \times (\Sigma \cup \{\varepsilon\}) \to \mathcal{P}_F(Q \times \Omega^*)$, the set of finite subsets of $Q \times \Omega^*$.*

If for all $(q, \omega, \sigma) \in Q \times \Omega \times (\Sigma \cup \{\varepsilon\})$ $\delta(q, \omega, \sigma)$ is a singleton then we say that P is a *deterministic* pushdown automaton. A *configuration* of a pushdown automaton is an element of $Q \times \Omega^*$. The *language accepted by* P is the set of inputs words which lead to an empty stack.

Example 5.9 Show that every language which can be accepted by a pushdown automaton can also be accepted by a pushdown automaton with only one state. (i.e. we can perform all calculations on the stack)

Just as rational languages are precisely those accepted by finite state automata, we can describe the set of languages accepted by pushdown automata.

Definition 5.10 *A* context free grammar *is the data $G = (V, T, P, S)$ where V is a set called the* set of variables*, T is a set disjoint from V called the* set of terminals*, P is a set called the* set of productions *and $S \in V$ is called the* start symbol*. A production is a rule of the form $A \to \alpha$, where A is a variable and α is a string of symbols from $(V \cup T)^*$. We denote variables by capital letters, and terminals by lower case. A grammar is specified by writing down its productions.*

Example 5.11 Arithmetic expressions, e.g. $x + (x^2 + (xy + 7yx) + 3y) + 5y$. The productions are $A \to A + B$, $A \to AB$, $A \to (A)$, $A \to x$ and $A \to y$. Obtain the above expression using these rules.

The *language accepted by* a context free grammar, $L(G)$ is the set of words on the set of terminals which can be derived from S by application of finitely many productions. A *context free language* is the accepted language of some context free grammar.

Theorem 5.12 *A language is context free if and only if it is accepted by a (nondeterministic) PDA.*

Chomsky normal form for context free grammars.

5.3 Word Problem Languages of Groups.

Given a group G with an inverse closed set of generators S, by the *word problem language* $E(G)$ we mean the set of words in S which are equal to the identity in G. The word problem for G asks precisely, when is $E(G)$ recursive. $E(G)$ is recursively enumerable if and only if G is recursively presentable. Suppose $E(G)$ belongs to a certain class of languages, e.g. regular languages. (In this case we say that G is *regular*, etc.) Then what can we say about G? (see also [48].)

Theorem 5.13 (Anisimov,1971) *A group is regular if and only if it is finite.*

Proof. Suppose that $G = \{g_1, \ldots, g_n\}$ is a finite group. The *multiplication table* presentation is the presentation

$$\langle g_1, \ldots, g_n | g_i g_j = g_k,\ 1 \leqslant i, j \leqslant n \rangle.$$

We construct a finite state automaton M with set of states $g_1, \ldots, g_n$ where g_1, (the identity in G), is the start state and the only accept state. The transition function is given by $(g_i, g_j) \mapsto g_i j$.

Conversely, suppose that $G = \langle S|R \rangle$ is an infinite finitely generated group. Then there exists arbitrarily long words w on $S^{\pm 1}$ such that no nonempty subword of w represents the identity in G. Let M be a finite state automaton with input alphabet $S^{\pm 1}$. Let w be a word whose length is greater than the number of states of M, such that no nonempty subword of w represents the identity of G. If M begins reading w, then there must exist distinct prefixes of w, say u and uv, such that M is in the same state after reading these. uu^{-1} is equal to the identity in G but uvu^{-1} is nonidentity in G since it is a conjugate of a nonidentity element. Now M must either accept both of uu^{-2} and uvu^{-1} or reject them both. Thus M cannot accept the word problem language of G. $\square$

The following example shows that we can accept the word problem language of more groups if we use pushdown automata.

Example 5.14 Consider F_2, the free group of rank 2. Let the pushdown automaton P have set of states S containing the single state s, stack alphabet $A = \{a, b, A, B\}$, input alphabet I equal to A and the following transition function (we consider the transition function as a map $A^* \times I \to A^*$, since $|S| = 1$).

$$\begin{array}{rr} (wb, a) \mapsto wba & (wb, b) \mapsto wbb \\ (wb, a^{-1}) \mapsto wba^{-1} & (wb, b^{-1}) \mapsto w, \end{array}$$

and similarly for pairs of the form $(wa, -)$,$(wb^{-1}, -)$ and $(wa^{-1}, -)$. Then P clearly accepts the word problem language of F_2. We can extend this example easily to F_n and using the construction of Anissimov's theorem to $F_n \times G$ where G is a finite group.

Example 5.15 $\mathbb{Z} \oplus \mathbb{Z}$ is not context free. A word of the form $x^m y^n x^i y^k$ is equal to the identity in A iff $j = m$ and $k = n$. Intuitively a PDA can't check if both equations hold. Formally to show this we use the pumping lemma for context free languages.

In fact given a context free group G, we can construct a Dehn presentation of G and so G is hyperbolic (see Gilman's notes). But we have the following stronger result. First of all note that context freedom is a group invariant.

Theorem 5.16 *If G is a finitely generated group and the word problem language for G is context free with respect to one finite set of generators then it is context free with respect to all finite sets of generators.*

We have the following characterisation of context free groups from [37].

Theorem 5.17 (Muller, Schupp) *A finitely generated group is context free if and only if it is virtually free.*

Proof. First we show that all virtually free groups G are context free. Let N be a free finite index subgroup of G which is again finitely generated and we can assume to be normal. Let N have free generators $y_1, \dots y_n$, suppose that $B = G/N = \{b_1, \dots b_t\}$ and let $\eta : G \to B$ be the natural projection. For $i = 1, \dots t$ choose elements d_i of G such that $\eta(d_i) = b_i$. Now $N = ker(\eta)$ so relations of the form $d_i y_d d_i^{-1} = u_{i,d}$ and $d_i d_j^{\varepsilon} = z_{i,j,\varepsilon} d_k$ hold, where the $u_{i,d}$ and $z_{i,j,\varepsilon}$ are elements of N. We claim that

$$\langle y_1, \dots y_n, d_1, \dots d_t | d_i y_d d_i^{-1} = u_{i,d}, d_i d_j^{\varepsilon} = z_{i,j,\varepsilon} d_k \rangle$$

is a presentation of G.

To see this, note that using these relations any word can be transformed to a unique word of the form wd_i, where w is a freely reduced word in the y_ds. The latter word represents the identity of G if and only if w is empty and d_i is the symbol d_1 with $\eta(d_1) = \mathrm{id}_B$. This proves the claim.

Exercise 5.18 Calculate such a presentation for the free product of two finite cyclic groups.

We now construct a PDA P accepting the word problem language of G with this presentation. Roughly speaking, P keeps the "free part" of a word in its stack and its image in B in its states. We have states $q_1, \dots q_t$ corresponding to $b_1, \dots b_t$ and other "working" states. If P is in state q_i and reads y_l^{ε} then since $d_i y_l^{\varepsilon} = u_{i,l}^{\varepsilon} d_i$ in G, P uses its working states to process the word $u_{i,l}^{\varepsilon}$ onto the stack, as in the automaton for the word problem of a free group, and returns to q_i. If p is in state q_i and reads d_j^{ε} then since $d_i d_j^{\varepsilon} = z_{i,j,\varepsilon} d_k$ in G, P uses its working states to process $z_{i,j,\varepsilon}$ onto the stack and then changes state to q_k. Thus, after reading an arbitrary word v, equal in G to wd_i, where w is a word on the $y_l^{\pm 1}$, P has w on its stack and is in state q_i. Thus v represents the identity of G if and only if P is in state q_1 with empty stack.

We now prove the converse, i.e. that if G is context free then G is virtually free. Since context freedom implies the existence of a Dehn algorithm, context free groups are hyperbolic and hence finitely presented. Now Dunwoody's theorem says that G must be accessible. We prove the main result by induction on the accessibility length s of G. If $s > 0$ then G has at most one end. But now we have a lemma.

Lemma 5.19 *An infinite context free group has more than one end.*

Proof. Let $G = \langle X | R \rangle$ be an infinite context free group. Let Γ be tha Cayley graph of G. As G is infinite there exist arbitrarily long geodesics in Γ, and by translating the midpoints to the origin, we get for any $i \geqslant 1$, elements u_i and v_i such that $d(u_i, 1) = d(v_i, 1) = 1$ and $d(u_i, e) = d(v_i, e) = i$ and $d(u_i, v_i) = 2i$. Since G is context free there exists a constant K such that every cycle in Γ can be K-triangulated. Pick $n \geqslant \frac{3}{2}K$ and let Γ^n be the ball of radius n around e in Γ. We show that if $i \geqslant n$ and u_i and v_i are as before, then u_i and v_i are in different components of $\Gamma - \Gamma^n$. This clearly implies the result. Suppose that u_i and v_i are in the same component of $\Gamma - \Gamma^n$. Let α be a geodesic $[e, u_i]$, γ be a geodesic $[v_i, e]$ and let β be a path from u_i to v_i in $\Gamma - \Gamma^n$. Let T be a K-triangulation of $\alpha\beta\gamma$. By the lemma, there exists a triangle t with vertices a,b and c on α, β and γ respectively. Each edge of t represents a path of length not exceeding K. Since $b \in \Gamma - \Gamma^n$ we must have $d(e, a) \geqslant n - K$, for otherwise one could go from e to a and then along a path of length not exceeding K to reach b, contradicting $b \in \Gamma - \Gamma^n$. Thus $d(a, u_i) \leqslant i - n + K$. Similarly, $d(c, v_i) \leqslant i - n + K$. But one can go from u_i to v_i by travelling along α^{-1} from u^{-1} to a, then along the path represented by the edge of t connecting a and c and then along α^{-1} from c to v_i. This is a distance not exceeding $2i + (3K - 2n)$, less than $2i$ since $2n > 3K$. But this contradicts the fact that $d(u_i, v_i) = 2i$. $\square$

We now continue with the main proof. We see that if $s = 0$ then G is finite by Stallings' ends theorem, and certainly virtually free. Now suppose that $s > 0$ and that $G = G_1 *_F G_2$ where F is finite (the case where G is an HNN extension is identical). Then G_1 and G_2 have accessibility length at most $(s-1)$. Now we invoke the following.

Lemma 5.20 *Finitely generated subgroups of finitely generated context free groups are context free.*

Thus G_1 and G_2 are context free and by induction are virtually free. Finally to complete the theorem we note that by the theorem of Gregorac, Karass et al, G is also virtually free. □

It is interesting to try and match the various subclasses of pushdown automata with subclasses of virtually free groups. In particular we have the following result of Thomas and Herbst. A *one-counter* automaton is a PDA whose stack alphabet consists of only one letter.

Theorem 5.21 *[28] A group G has a word problem accepted by a one-counter automaton if and only if G is virtually cyclic.*

More general languages than context free languages are the *indexed* languages. These are accepted by *nondeterministic one-way nested stack automata*. The nested stack automaton has the capability when the stack head is inside the stack in read only mode, to create a new stack. However this stack must be destroyed before the stack head can move up in its original stack.

Somewhere between pushdown automata and nested stack automata we have *stack automata*. These are similar to pushdown automata but have two additional features. The input is two way, read only with end markers, and the stack head can enter the stack in read only mode, travelling up and down the stack without rewriting any symbol

Example 5.22 $\{a^n b^b c^n | n \in \mathbb{N}\}$ is indexed but not context free. However we can accept this language by a stack automaton. The rough idea is to push all the a's onto the stack, move the stack head up the stack while reading b's and simultaneously move it back down again and pop a's while reading c's.

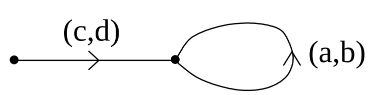
Linear bounded automata and context sensitive languages.

Some examples of groups with context sensitive word problem are $F_2 \times F_2$, linear groups over a countable field or $\mathbb{Z}$ and some finitely generated infinite torsion groups of intermediate growth. Turing machines, recursive sets and recursive enumerability.

5.4 Automatic Groups.

Let A_1 and A_2 be alphabets. By a *language over* (A_1, A_2) we mean a set of pairs (w_1, w_2) with $w_1 \in A_1^*$ and $w_2 \in A_2^*$. Such a language is called a *2-variable language*.

Example 5.23 A 2-variable language. $L = \{(ca^n, db^n) \mid n \in \mathbb{N}\}$



We would like to define regularity of a 2-variable language. Using our present definition, this cannot be done for languages which contain pairs (w_1, w_2) where w_1 and w_2 have different lengths. To solve this, we adjoin to the alphabet an *end-of-string* or *padding symbol* \$ and we define $B_i = A_i \cup \{\$\}$ for $i = 1$ and 2. The *padded alphabet* associated with (A_1, A_2) is $B = B_1 \times B_2 - \{(\$, \$)\}$. A *padded string* over B is a string $w = (w_1, w_2)$ such that once an end-of-string symbol occurs in one of w_1 or w_2 then all of the subsequent letters of that string are end-of-string symbols.

Definition 5.24 *Given a language L over (A_1, A_2) we define a 1-variable language $L^{\$}$ over the padded alphabet B associated with (A_1, A_2) by padding the string of (w_1, w_2) of shorter length until it has the same length as the longer one. To the resulting pair there corresponds a unique word in B. The set of all these words on B is called the* padded extension *of L.*

Example 5.25 (now,later) $\mapsto$ (now\$\$,later) $\mapsto$ (n,l)(o,a),(w,t)(\$,e)(\$,r)

Definition 5.26 *L is a* regular language *over (A_1, A_2) if $L^\$$ is a regular language over the padded alphabet B associated with (A_1, A_2). A finite state automaton accepting $L^\$$ is said to be a* 2-variable automaton over (A_1, A_2) accepting L*.*

Definition 5.27 *Let G be a finitely generated group. An* automatic structure *on G is a $(q+3)$-tuple $(S, M_0, \ldots, M_n)$ such that $S = \{s_0, \ldots, s_q\}$ is an inverse closed set of generators for G with s_0 equal to the identity, M is a finite state automaton with alphabet S such that $L(M)$ maps surjectively onto G under the natural map $F(S) \to G$, and for each $0 \leqslant i \leqslant q$, M_i are finite state automata such that $(w_1, w_2) \in L(M_i)$ if and only if $\overline{w_1 s_i} = \overline{w_2}$ and both w_1 and w_2 are in $L(M)$.*

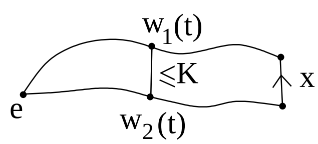
We call M the *word acceptor*, M_0 the *equality recogniser* and each M_i for $1 \leqslant i \leqslant q$ a *multiplier automator* for the automatic structure. Note that the diagonal subset of $L(M_0)$ is equal to $L(M)$.

Definition 5.28 *A group is* automatic *if it admits an automatic structure.*

In what follows if w is a word of length l and $0 \leqslant t \leqslant l$ then we denote by $w(t)$ the prefix of w of length t.

Proposition 5.29 (Fellow Traveller Property) *If a group G is automatic then there is a constant K (depending on the chosen automatic structure) such that if (w_1, w_2) is accepted by either the equality recogniser or one of the multiplier automators then for all t we have $d(\overline{w_1(t)}, \overline{w_2(t)}) \leqslant K$*

We call K the *Lipschitz constant* of the structure and we say that the words w_1 and w_2 *K-fellow travel.*



Proof. Let $A = \{a_0, \ldots, a_n\}$ be an inverse closed set of generators for G which forms the alphabet of an automatic structure for G. Let m be the maximum number of states in any of the multiplier automators or the equality recogniser and let M_i be the corresponding automaton. For any $0 \leqslant t \leqslant \max |w_1|, |w_2|$ assume that M_i has read $(w_1(t), w_2(t))$ and arrived at the state s and let (u_1, u_2) be the shortest path in $A^* \times A^*$ which leads to an accept state. Then $\overline{w_1(t) u_1 a_i} = \overline{w_2(t) u_2}$ with $|u_1|$ and $|u_2|$ both no greater than m. Then $d(\overline{w_1(t)}, \overline{w_2(t)}) \leqslant 2m + 1$, so a suitable Lipschitz constant is $2m + 1$. $\square$

There is a lot of data in the definition of an automatic structure, which we can reduce using the following construction. Suppose we are given a group G, a finite inverse closed set A of generators of G including its identity and a finite state automaton M whose language maps surjectively onto G. Suppose furthermore that there exists $K \geqslant 0$ such that whenever two strings w_1 and w_2 where $\overline{w_1 a} = \overline{w_2}$ for some a are accepted by M, then the corresponding paths $\overline{w_1}$ and $\overline{w_2}$ are a uniform distance less than K apart. Then we can construct an automatic structure for G with word acceptor M. The equality recogniser and multiplier automators obtained in this way are called the *standard automata* for G based on (M, K). Let B be the ball of radius K of the identity in the Cayley graph of G. Let N be a finite state automaton over A accepting the language $L(W)\$^*$ and let S be its set of states. Then we define the standard automata M_a for all $a \in A$ as follows. Its set of states is $S \times S \times B$ with initial state (s_0, s_0, e), where s_0 is the initial state of N and e is the identity element of G. The transition function of M_a is given by

$$((s_1, s_2, g), (y_1, y_2)) \mapsto (s_1 y_1, s_2 y_2, y_1^{-1} g y_2)$$

and the accept states of M_a are the states (s_1, s_2, a) where s_1 and s_2 are both accept states of N.

Proposition 5.30 *The standard automata equip G with an automatic structure.*

Proof. □

Example 5.31 1. F_n, the free group of rank n, is automatic. We shall see later that this generalises to hyperbolic groups.

2. Finitely generated abelian groups are automatic.

3. More generally than 1 and 2, let Γ be a finite graph and let G be the group generated by the vertices of Γ and with relators $[v_i, v_j]$ whenever v_i and v_j are adjacent. Then G is called the *graph group* on Γ and is known to be automatic (in fact biautomatic) [12],[67]. Graph groups are sometimes called *right-angled Artin groups*, *partially commutative groups* or *trace groups*. Examples of graph groups include direct products of finitely generated free groups and free products of finitely generated free abelian groups.

4. Direct products of automatic groups are automatic.

5. A finitely generated subgroup of an automatic group G is not necessarily finite or automatic. For example, the direct product $F(a,b) \times F(c,d)$ of two free groups of rank 2 is automatic, and we may construct a homomorphism $\theta : F_2 \times F_2 \to \mathbb{Z} = \langle g \rangle$ via $\theta(a) = \theta(b) = \theta(c) = \theta(d) = g$. Then $\ker(\theta)$ is finitely generated but not finitely presented. In fact we can consider the analogous map $\theta : F_2 \times F_2 \times F_2 \to \mathbb{Z}$. Then Mess REFERENCE has shown that $\ker \theta$ has infinitely generated third homology group, which means that it can't be automatic. It is unknown whether $\ker(\theta)$ satisfies a quadratic isoperimetric inequality.

5.5 The Quadratic Isoperimetric Inequality.

Lemma 5.32 *Let G be an automatic group and let L be the language of accepted words for some automatic structure of A. Then there exist constants P and Q such for all words w in G, the prefixes $w(t)$ of w all have representatives $u_t \in L$ of length at most $P|w| + Q$.*

i.e. automatic groups are "combable":

Definition 5.33 *A finitely generated group is said to be* **combable** *if it admits a quasigeodesic combing which satisfies the fellow traveller property.*

Theorem 5.34 *Automatic groups are finitely presented and satisfy a quadratic isoperimetric inequality.*

This shows that the word problem is solvable for automatic groups, since a quadratic function is certainly recursive. (See exercise 3.19.) Unlike hyperbolic groups, which are characterised by the linear isoperimetric inequality, automatic groups are not characterised by the quadratic isoperimetric inequality [18].

5.6 Strongly Geodesically Automatic Groups.

In what follows when we say that a language L is "part of an automatic structure" we mean that L is the language accepted by the word acceptor of the automatic structure.

Definition 5.35 *If the language of all geodesics of a group G is part of an automatic structure for G then we say that G is* strongly geodesically automatic. *If a language consisting only of geodesics is part of an automatic structure for G then we say that G is* weakly geodesically automatic.

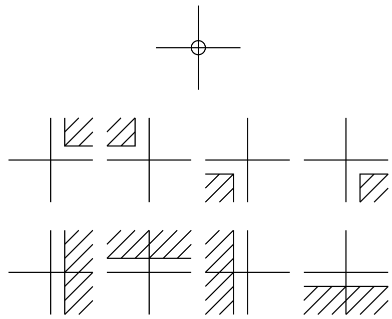
We now exhibit a large class of strongly geodesically automatic groups. The following construction is due to Jim Cannon REFERENCE.

Definition 5.36 *Let G be a group with an inverse closed set of generators S. Let $w \in S^*$. Then the* cone type *of w is*

$$C(w) = \{w' \in S^* \mid ww' \textit{ is a geodesic}\}.$$

Note that $C(w)$ is nonempty if and only if w is a geodesic. We now extend the definition to group elements. Let $g \in G$. We define $C(g)$ to be equal to $C(w)$ where w is a geodesic word for G. This is clearly well defined.

Example 5.37 In $\mathbb{Z} \oplus \mathbb{Z}$ with the standard generators there are 9 cone types, as in the following picture.



Definition 5.38 *Let G be a presented group and let $g \in G$. Then the* n-level *of g is*

$$\{h \in g||h| \leqslant n, d(gh, e) < d(g, e)\}.$$

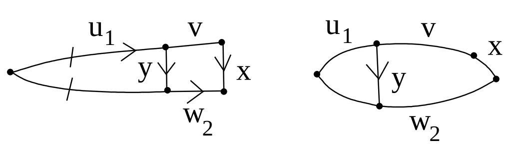
Let G be a hyperbolic group. Let L be the language of all geodesics in G. Then the language L is easily seen to K-fellow travel for some K, depending on the thinness δ of triangles.

Proposition 5.39 *If G is hyperbolic with K as above then the $(K+1)$-level of an element $g \in G$ determines its cone type.*

Proof. Suppose u_1 and u_2 are geodesic words with the same $(K+1)$-level. We show by induction on $|v|$ that u_1v is a geodesic if and only if u_2v is a geodesic. This is clearly true if $|v| = 0$. For the inductive step, let x be a generator of G and suppose that u_1v, u_2v and u_1vx are geodesics. We aim to show that u_2vx is a geodesic. Suppose not. Then there exist words w_1 and w_2 with $\overline{u_2vx} = \overline{w_1w_2}$, $|w_1w_2| < |u_2vx|$ and $|w_1| = |u_2| - 1$. Consider the word $y = u_2^{-1}w_1$ in the first picture. $|y| < K+1$ by considering the centre triangle, and

$$d(u_2y, e) = d(w_1, e) = |u_2| - 1 \leqslant d(u_2, e).$$

So y is in the $(K+1)$-level of $\overline{u_2}$. Thus, by assumption, y is in the $(K+1)$-level of $\overline{u_1}$ and we have $|\overline{u_1y}| < |\overline{u_1}|$.



Now $|\overline{u_1yw_2}| = |\overline{u_1vx}|$ and if w_1' is a geodesic word for $\overline{u_1y}$ then $|w_1'| < |u_1vx|$, which contradicts the fact that u_1vx is a geodesic. $\square$

Theorem 5.40 *Hyperbolic groups are strongly geodesically automatic.*

Proof. Let $G = \langle A|R\rangle$ be a hyperbolic group. We construct a finite state automaton M which accepts all geodesic words on A. The states of M are all possible cone types of elements of G plus the empty set $\emptyset$. The start state is $C(e)$ and the accept states are all but $\emptyset$. The transition function μ of M is as follows. If $s = C(w)$ is a state, then for $a \in A$ we have $\mu(s,a) = C(wa)$. Now the standard automaton corresponding to the ball of radius K (defined above) equips G with a strongly geodesically automatic structure. $\square$

Note that the above structure is also biautomatic (see section 5.7) and that the above automaton can be constructed for any group with finitely many cone types, but we need the fellow traveller property to get an automatic structure.

We now give an application of automatic groups to hyperbolic groups.

Definition 5.41 *Let G be generated by a finite set S and let a_n be the number of elements $g \in G$ with $|g| = n$. Then the* **growth series** *of G is*

$$J_S(t) = \sum_{n=0}^{\infty} a_n t^n.$$

Example 5.42 1. Consider $\mathbb{Z}$ with generating set $S = \{x, x^{-1}\}$. We then have $a_0 = 1$, $a_1 = 2$ and $a_n = 2$ for all $n \geqslant 2$. Then

$$J_S(t) = 1 + \sum_{n=1}^{\infty} 2t^n = \frac{1}{1+t} + \frac{t}{1-t} = \frac{1+t}{1-t}.$$

Note that this is a rational function of t, i.e. a quotient of two polynomials in t.

2. For F_2 with the usual generators we have

$$J(t) = 5 + \sum_{n=1}^{\infty} 4.3^n$$

which we can sum using geometric series.

Theorem 5.43 *The growth series of a hyperbolic group is rational.*

Proof. Let G be a hyperbolic group generated by an inverse closed finite set A. Let M_ε be the equality recogniser of an automatic structure which accepts all geodesic strings on A. Take an ordering of A and the shortlex automaton M corresponding to M_ε with this ordering. Let B be the transition matrix of the graph of M, i.e. B_{ij} is the number of edges going from the i^{th} vertex to the j^{th} vertex. Then $B^n_{i,j}$ is equal to the number of paths of length n from i to j. If we let the set of states of M be S, its start state be s_1 and its set of accept states be Y then we have

$$J_A(t) = \sum_{n=1}^{\infty} \left(\sum_{j \in Y} B^n_{1j} \right) t^n.$$

Let $p(t) = c_0 + c_1 t + \cdots + c_k t^k$ be the minimal polynomial of the matrix B, i.e. the smallest polynomial dividing its characteristic polynomial, and let $q(t) = c_0 t^k + c_1 t^{k-1} + \cdots + c_k$. Then $q(t)J_A(t)$ is a polynomial in t $\square$

Note that in general, the rationality of the growth series of a finitely presented group is not an invariant of the presentation (see [64]).

Papasoglu [40, 42] has shown that the converse to theorem 5.40 also holds:

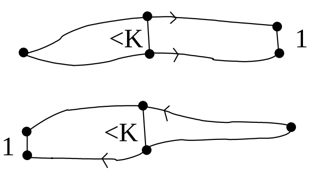
Theorem 5.44 *Strongly geodesically automatic groups are hyperbolic.*

5.7 Biautomatic Groups.

Let G be a finitely generated group and let S be a finite set of generators for G.

Definition 5.45 *A set of words $L \subset S$ is a* biautomatic structure *for G if L is surjective onto G, regular, and there exists K such that for all w_1 and w_1 in $F(S)$,*

1. *if $\overline{w_1 s} = \overline{w_2}$ for some $s \in S$ then w_1 and w_2 K-fellow travel.*
2. *if $\overline{s w_1} = \overline{w_2}$ for some $s \in S$ then w_1 and w_2 K-fellow travel.*



Definition 5.46 *We say that a subset H of a biautomatic group G with regular language L is* rational *if the inverse image $\{w \in L | \overline{w} \in H\}$ is regular. If, furthermore, H is a subgroup then we call it a* rational subgroup.

Example 5.47 A subgroup of a free group is finitely generated if and only if it is rational (see [AS]). Subgroups of abelian groups are rational. More examples are given in [25].

The following should be compared with theorem 3.44.

Theorem 5.48 *[25] Rational subgroups of biautomatic groups are biautomatic.*

Theorem 5.49 *[25] Let G be a biautomatic group.*

1. *For every finite subset X of G, the centraliser $C_G(X)$ is rational.*
2. *For every finitely generated subgroup H of G, the centraliser $C_G(H)$ is biautomatic.*
3. *For every rational subgroup H of G, the centre $Z(H)$ is biautomatic.*
4. *For every rational subset R of a biautomatic group, the centraliser $C_G(R)$ is biautomatic.*

5.8 Translation Numbers.

Definition 5.50 *Let G be a finitely presented group and let S be a finite set of generators for G. Then for $g \in G$ we define the* translation number *of g to be*

$$\tau_{G,A}(g) = \liminf_{n>0} \frac{|g^n|_A}{n}$$

In fact, by the triangle inequality we have $|g^{n+m}| \leqslant |g^n| + |g^m|$. Thus we can use the following theorem from analysis to show that in the above definition, the *limit* always exists, and we can dispense with the $\liminf$.

Theorem 5.51 (Polya-Szego) *If the sequence of nonnegative real numbers (a_n) satifies $a_{n+m} \leqslant a_n + a_m$ for all m and n, then*

$$\lim \left(\frac{a_n}{n}\right) \text{ exists.}$$

Proof. Fix m and write $n = rm + q$ with $r > 0$ and $1 \leqslant q \leqslant m$. Then

$$\frac{a_n}{n} = \frac{a_{rm+q}}{rm+q} \leqslant \frac{ra_m + a_q}{rm+q} \to \frac{a_m}{m} \text{ as } n \to \infty.$$

Then for all $m \geqslant 1$ we have

$$\limsup \left(\frac{a_n}{a}\right) \leqslant \frac{a_m}{m},$$

and since we now have

$$\limsup \left(\frac{a_n}{n}\right) \leqslant \inf_m \left(\frac{a_m}{m}\right) \leqslant \liminf \left(\frac{a_n}{n}\right),$$

the result follows. □

If g is a torsion element of the group G, then clearly as $\{|g| \mid n \in \mathbb{N}\}$ is a bounded sequence, $\tau_G(g) = 0$ (the converse is however, false).

Example 5.52 For the free group F_n the translation number of g is the length of a cyclically reduced word representing $g \in F$. For a free abelian group with basis A, then the translation number of an element is the same as its l_1-norm, i.e. the sum of absolute values of the coefficients.

Note that the translation number of a group element depends on the generating set. However, we have the following.

Proposition 5.53 *If S_1 and S_2 are generating sets for G, and for an element $g \in G$, we have $\tau_{G,S_1}(g) \neq 0$, then $\tau_{G,S_2}(g) \neq 0$.*

Proof. We can write elements in A as words in B and vice-versa (as in the proof that the quasi-isometry type of the Cayley graph of a group is a presentation invariant). In this way there exists $\lambda > 0$ such that for all $g \in G$ we have

$$\frac{1}{\lambda}|g|_{S_2} \leqslant |g|_{S_1} \leqslant \lambda |g|_{S_2}.$$

The result hence follows. □

So we may talk unambiguously of the translation number of a group element being nonzero. Note that this result tells us that the translation number of a non-identity element of a free abelian group of finite rank is nonzero for any generating set.

Translation numbers provide class functions on G (i.e. they are constant on conjugacy classes):

Proposition 5.54 *For all x and g in G, $\tau_{G,S}(x^{-1}gx) = \tau_{G,S}(g)$.*

Proof. We have

$$\frac{|x^{-1}g^n x|}{n} \leqslant 2\frac{|x|}{n} + \frac{|g^n|}{n}.$$

Taking limits, we see that $\tau(x^{-1}gx) \leqslant \tau(g)$. Replacing g by xgx^{-1} and x by x^{-1} gives the reverse inequality. □

A simple calculation also yields the next result.

Proposition 5.55 *For all $g \in G$, $\tau_{G,S}(g^n) = |n|\tau_{G,S}(g)$*

Example 5.56 A cocompact torsion free Fuchsian group is the fundamental group of some hyperbolic surface, so an element g of the group represents a homotopy class of loops in this surface. The translation number of an element is then determined by the length l of the shortest geodesic in this homotopy class, i.e. we have

$$\tau(g) = \cosh(\frac{l}{2}) = \frac{1}{2}\mathrm{tr}(C) = \cosh\left(\frac{d}{2}\right),$$

Where C is the matrix of $\mathrm{SL}_2(\mathbb{R})$ representing g and d is the distance g translates by in the hyperbolic plane (Given a hyperbolic translation we join its two fixed points to obtain its *axis*. The distance of translation is measured along the axis of such a transformation.)

Exercise 5.57 Let G be a group. The image of the map $G \to \mathbb{R}$ given by $g \to \tau_G(g)$ is called the *length spectrum* of G. Prove the following theorem of Delzant: Let G be a hyperbolic group. Choose a finite generating set for G. Then the translation number of every element of G is rational and, moreover, G has discrete length spectrum.

5.9 Nilpotent Subgroups of Biautomatic Groups.

The main tool in this section will be the translation number.

Proposition 5.58 *For every rational subgroup H of a biautomatic group G and for every element $h \in H$, $\tau_G(h) \neq 0$ if and only if $\tau_H(h) \neq 0$.*

Proof. Take A to be a finite set of generators for G and let $L \subset A^*$ be a regular combing of G with H represented by a regular sublanguage $L' \subset L$. Under these circumstances (see [18]) there exists λ and ε such that all words in L are (λ, ε)-quasigeodesics.

We claim that if $w \in A^*$ represents the element $g \in G$ then $\tau_{G,A} \neq 0$ if and only if there exist positive constants μ and C such that for all $n > 0$,

$$\frac{1}{\mu} l(w^n) - C \leqslant |g^n| \leqslant l(w^n).$$

To see this, suppose $t = \tau(g) > 0$. As t is a limit, there exists $N \in \mathbb{N}$ such that for all $n > N$, $\frac{|g^n|}{n} > \frac{t}{2}$, i.e. there exists $C > 0$ such that $\frac{tn}{2} - C \leqslant |g^n|$. Since $n = \frac{l(w^n)}{l(w)}$ we may let $\mu = \frac{t}{2l(w)}$ and μ is as required.

Now for any finite set of generators B for H, there exist positive μ and ε such that for all $h \in H$,

$$\frac{1}{\mu}|h|_{H,B} - \varepsilon \leqslant |h|_{G,A} \leqslant \mu|h|_{H,B} + \varepsilon.$$

Replacing h by h^n, dividing by n and taking the limit as $n \to \infty$, we have for all $h \in H$,

$$\frac{1}{\mu}\tau_{H,B}(h) \leqslant \tau_{G,A}(h) \leqslant \mu\tau_{H,B}(h).$$

Suppose $\tau_H(h) \neq 0$. Then $0 < \frac{1}{\mu}\tau_H(h)\tau_G(h)$. So $\tau_G(h) \neq 0$. Similarly, if $\tau_G(h) \neq 0$ then $\mu\tau_H(h) \geqslant \tau_G(h) \neq 0$ so $\tau_H(h) > 0$. □

Proposition 5.59 *In a biautomatic group, every element of infinite order has strictly positive translation number.*

Proof. Let x be an element of infinite order in the biautomatic group G. Consider the subgroups $Z(C_G(x)) \leqslant C_G(x) \leqslant G$. These are all rational and hence biautomatic by theorem 5.49. So the translation number of x in G is zero if and only if it is zero in all of these subgroups by proposition 5.58. Now $Z(C_G(x))$ is abelian, and finitely generated, since it is biautomatic. By general theory of finitely generated abelian groups, x lies in a torsion free subgroup K of finite index in $Z(C_G(x))$. Thus $\tau_K(x) \neq 0$. Since subgroups of finite index in biautomatic groups are rational, we have that $\tau_{Z(C_G(x))}(x) \neq 0$ and that $\tau_G(x) \neq 0$. □

Corollary 5.60 *The Baumslag-Solitar group $B_{k,l}$ for $|k| \neq |l|$ is not isomorphic to a subgroup of a biautomatic group.*

Proof. Recall that $B_{k,l}$ has presentation

$$B_{k,l} = \langle x, y | y x^k y^{-1} = x^l \rangle.$$

By known theory, x has infinite order. Suppose that G is biautomatic and there exist x and y in G such that $yx^ky^{-1} = x$ where $|k| \neq |l|$. Then we show that x is of finite order. For

$$\tau_G(yx^ky^{-1}) = \tau_G(x^k) = \tau_G(y^l),$$

so we have $(|k| - |l|)\tau_G(x) = 0$. But $|k| - |l| \neq 0$ and it follows that $\tau_G(x) = 0$. Hence x has finite order. □

Theorem 5.61 *Finitely generated nilpotent subgroups of biautomatic groups are virtually abelian.*

Proof. Let H be a finitely generated nilpotent subgroup of the biautomatic group G. After passing to a subgroup of finite index we may assume that H is torsion free. Suppose that H is not abelian. Then there exist elements x, y and z of H such that $[x, y] = z$, where $x \in Z_2(H)$, the second term of the upper central series of H and $z \neq 1$. We may assume without loss of generality that x, y and z are all in our generating set. Recall that for all elements $g \in Z_2(H)$ and all elements $h \in H$ we have $[g, h] \in Z(h)$. Thus

$$[x, y^n]xyx^{-1}y^{-1} = [x, y^{n+1}].$$

And similarly one has

$$[x^n, y][x, y] = [x^{n+1}, y].$$

It follows by induction that

$$[x, y]^{n^2} = [x, y^n]^n = [x^n, y^n].$$

Thus

$$\tau_G(z) = \lim \left(\frac{|z^{n^2}|}{n^2}\right) = \lim \left(\frac{4n}{n^2}\right) = 0.$$

But this contradicts the fact that z has infinite order (z is in the torsion free subgroup H). □

Example 5.62 By the last theorem, $\mathrm{SL}_n(\mathbb{Z})$ is not biautomatic for $n \geqslant 3$. Thurston has proved this using different methods.

Corollary 5.63 *Finitely generated nilpotent subgroups of hyperbolic groups are virtually cyclic.*

There are other theorems which have been proved for biautomatic groups, but remain open questions for automatic groups. For example:

Theorem 5.64 *[26] The conjugacy problem is solvable for biautomatic groups.*

Gersten and Short in [25] extend theorem 5.61 to the following

Theorem 5.65 *Polycyclic subgroups of biautomatic groups are virtually abelian.*

This has recently been extended to automatic groups by Harkins [27] using Lie group methods and higher dimensional isoperimetric inequalities.

6 Suggested Further Reading.

The theory of hyperbolic groups is enormous and contains many deep results which interact with other branches of mathematics. In this section we hopefully give a flavour of some of these. If an interested reader were one day to write a book giving a self-contained exposition of even just what appears here, it would probably run to several volumes.

A beautiful theorem due to Gromov is the following:

Theorem 6.1 *A finitely presented group has polynomial growth if and only if it is virtually nilpotent.*

This appears in [21], and a further account is given in [66]. The techniques used to prove theorem 6.1 were very innovative and have since been used in other situations. Gromov constructs what is now called the *asymptotic cone* of a nilpotent group G. Let (X, d) be the (vertex set of the) Cayley graph of a nilpotent group G equipped with the corresponding word metric. For all integers n we obtain a metric space $X_n = (X, d_n)$ by defining $d_n(x_1, x_2) = d(x_1, x_2)/n$. Gromov makes sense of what is meant by a limit metric space of this sequence of metric spaces. This is called the asymptotic cone of G. The action on the G on (X, d) extends to an action on the asymptotic cone, and by studying this action one can learn important information about G. Various people have also used this idea in the case where G is a hyperbolic group. In this case, if X is δ-hyperbolic then X_n is δ/n-hyperbolic and the limit becomes a 0-hyperbolic metric space. Such a space is called an $\mathbb{R}$-*tree*. The theory of actions of groups on $\mathbb{R}$-trees is both an important tool in geometric group theory and an interesting subject in its own right. See [60] for an introduction to $\mathbb{R}$-trees. For instance, Rips has proved the following (see [4] or [34]).

Theorem 6.2 (Rips) *If G is a finitely generated group which acts freely on an $\mathbb{R}$-tree then G is a free product of surface groups and free abelian groups.*

So not many finitely generated groups admit free actions on $\mathbb{R}$ trees. An action of a group on an $\mathbb{R}$-tree T is said to be *small* if the stabilizer of no point of T contains a free group of rank 2. For hyperbolic groups this means that the stabilizer of every point of T has to be virtually cyclic. Rips and Sela in [50] call a group *rigid* if it does not admit a small action on an $\mathbb{R}$-tree. Paulin uses asymptotic cone methods in [43] (see also [10]) to prove the following.

Theorem 6.3 (Paulin) *If G is a rigid hyperbolic group then $Out(G)$ is finite.*

By using related techniques, Sela proves in [57] that the isomorphism problem is solvable for rigid hyperbolic groups. This uses canonical representatives. Sela has also used canonical representatives in [56] to construct a uniform embedding of the Cayley graph of a torsion-free hyperbolic group in a Hilbert Space. This has relevance to work of Gromov relating to the Novikov conjecture.

A group is called *hopfian* if it has no proper quotient isomorphic to itself. For example, the Baumslag-Solitar group $B_{2,3}$ (see exercise 3.34) is non-hopfian (see [36]). Finitely generated linear groups are hopfian (see [68]) so it follows that $B_{2,3}$ is not linear. Sela proves the following in [59].

Theorem 6.4 *Torsion free hyperbolic groups are hopfian.*

There is also a huge body of literature on boundaries of hyperbolic groups. One highlight is the following theorem of Bowditch [7], which came from a deep understanding of the action of a hyperbolic group on its boundary:

Theorem 6.5 (Bowditch) *If G is a one-ended hyperbolic group such that ∂G has no global cut points then G splits over a virtually cyclic subgroup with infinite index in both factors if and only if ∂G has a local cut point.*

This was complemented by the following theorem of Swarup [55]:

Theorem 6.6 (Swarup) *If G is a one-ended hyperbolic group then ∂G has no global cut points.*

Thus it follows that

Theorem 6.7 *A one-ended hyperbolic group G splits over a virtually cyclic subgroup with infinite index in both factors if and only if ∂G has no global cut points.*

See [32] for a recent survey on boundaries.

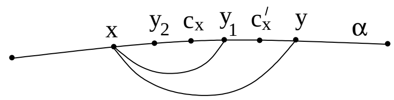
7 Hints and Answers To Exercises.

Exercise 2.27

1. Take a big square!

2. Fix x to be any point on p. Let y be any other point on p. We must show that $d_p(x,y) \leqslant 2d(x,y)$. In fact we are going to prove, by induction on $d_p(x,y)$ that the following equivalent statement is true: If $d_p(x,y) = d_p(x,y_1) + 4\delta$ then $d(x,y) \geqslant d(x,y_1) + 2\delta$. For the base case, assume that $d_p(x,y) \leqslant 8\delta$. Then $d_p(x,y) = d(x,y)$, which is equal to $d_p(x,y_1) + 4\delta$. So $d_p(x,y) = d_p(x,y_1) + 4\delta$ implies that $d(x,y) \geqslant d(x,y_1) + 4\delta \geqslant d(x,y_1) + 2\delta$.

   Now suppose that the statement is true for $d_p(x,y) \leqslant 4k\delta$ for some $K \geqslant 2$. We shall show that it is then true for $d_p(x,y) \leqslant 4(K+1)\delta$. Let y_1 and y_2 be points on p between x and y such that $d_p(y_1,y) = 4\delta$ and $d_p(y_2,y_1) = 4\delta$.

   Consider the triangles $[x,y_1,y_2]$ and $[x,y_1,y]$. By the inductive hypothesis, $d(x,y_1) \geqslant d(x,y_2) + 2\delta$. Therefore, if we let c_x be the internal point of $[x,y_1,y_2]$ on $[y_1,y_2]$, then $d(c_x,y_1) > 2\delta$. Let c'_x be the internal point of $[x,y_1,y]$ on $[y_1,y]$. Then $d(c'_x,y_1) \leqslant \delta$. Otherwise there are $c_1 \in [c_x,y_1]$ and $c_2 \in [y_1,c'_x]$ such that $d(c_1,y_1) = d(c_2,y_1) > \delta$ implies that $d(c_1,c_2) > 2\delta$. But by thinness of the two triangles, $d(c_1,c_2) < 2\delta$, which is a contradiction.

3. Let G be a finitely generated group with Cayley graph Γ and suppose that there exists $K > 1$ such that no cycle in Γ is a K-local geodesic. Then we claim that G is hyperbolic. We show that G has a linear isoperimetric inequality. Suppose c is a cycle of perimeter n and that every cycle in Γ is not a K-local geodesic. Then there exists a geodesic γ' such that $|\gamma'| < |\gamma|$, where γ is a subpath of c. By repeating the argument we have

   $$A(C) = A(B_1) + \cdots + A(B_m) \leqslant mA,$$

   where $B_1, \ldots, B_n$ is the sequence of bigons we obtain, and $A = \max\{A(B_i) | 1 \leqslant i \leqslant m\} \leqslant K$. Thus G is hyperbolic.

   It follows from the claim that if for some $K > 0$ we have that every K-local geodesic is a quasigeodesic then G is hyperbolic.

Exercise 3.4

If a group is not finitely presented then there are bigger and bigger "holes" in the Cayley graph. Use covering spaces and the Cayley complex.

Exercise 3.7

The quasi-isometry (one way) is given by the orbit map $g \mapsto gx_0$ for some basepoint x_0.

Exercise 3.18

G is trivial since firstly $a^2b^4 = a(ab^2)b^2 = a(b^3a)b^2 = ab^6a$. But this is equal to $b^9a^2 = b^8a^3b$ and hence $a^2b^3 = b^8a^3$. By symmetry $b^2a^3 = a^8b^3$. So $a^2b^3 = b^6a^8b^3$ as $a^2b^3 = b^6b^2a^3 = b^6a^8b^3$. We thus obtain $a^2 = b^6a^8$ and hence $a^{-6} = b^6$. But from the first line we now have $a^2b^4 = a^{-4}$ from which $b^4 = a^{-6}$ and $b^2 = 1$. So $a = b^3a$ which tells us that $b^3 = 1$. Now we know that $b^2 = b^3$ and thus $b = 1$. Also $a^2 = a^3$ which gives $a = 1$.

Exercise 5.57 Clearly we only need to consider infinite order elements of G. Let $g \in G$ and let g^+

and g^- denote the limit points on ∂G of the sequences $\{g^n | n \in \mathbb{N}\}$ and $\{g^{-n} | n \in \mathbb{N}\}$ respectively. Choose a geodesic combing θ of G (e.g. shortlex). Let K be the constant for G such that a geodesic from g^- to g^+ stays within a distance K of $\{g^n | n \in \mathbb{N}\}$. Let A_n be the set of geodesics in our combing from g^{-n} to g^{+n}. We then have $|A_n| \leqslant K^2$. Let B_n be the subset of A_n which we can extend to geodesics from g^- to g^+. Then if we define B to be the set of all infinite geodesics from g^- to g^+ which always restrict to combing lines, $|B| \leqslant K^2$, i.e. in particular this is a finite set. The left action of G on its Cayley graph preserves this set, so there exists $0 \leqslant N \leqslant K^2$ such that given a geodesic γ from g^- to g^+ through the identity and given $g \in G$, g^N lies on γ, and hence does g^{mN} for all $m \in \mathbb{N}$. Then if we let $d = |g^N|$, we have

$$\frac{|g^m N|}{mN} = \frac{md}{mN} = \frac{d}{N},$$

from which the result follows.

8 Open Problems.

Most of these are well known open problems and have stood for at least 10 years. See [70] for up-to-date information, and many other questions. Another good list of open problems in geometric group theory is available at [2].

1. A group G is *linear* if there exists a field F such that G has a faithful representation in $\mathrm{GL}_n(F)$ for some n. Is every hyperbolic group linear? (See [69].)

2. [70] A group G is *equationally noetherian* if, given an infinite system of equations over G, we can always replace it by a finite system with the same solution set. Is every hyperbolic group equationally noetherian? (Note that finitely generated linear groups are equationally noetherian REFERENCE, so a positive answer to question 1 would imply a positive answer to this one.)

3. The following two questions are shown to be equivalent in [33].

   (a) Is every hyperbolic group residually finite? (Note that finitely generated linear groups are residually finite [68], so a positive answer to question 1 would imply a positive answer to this one.)

   (b) Does every hyperbolic group have a proper subgroup of finite index? Note that if this were true then every hyperbolic group would have an infinite descending chain of proper subgroups of finite index. For an example of a finitely generated group with no proper subgroup of finite index, see [54, 29]

4. Is every hyperbolic group virtually torsion free?

5. Is it true that every group which admits a finite system of canonical representatives is hyperbolic?

6. Is it true that every finitely generated group with indexed word problem is virtually free?

7. [18] Is every automatic group biautomatic?

8. [18] Is the conjugacy problem solvable for automatic groups?

9. [18] Is the isomorphism problem solvable for automatic groups?

10. [25] Is the length spectrum of a biautomatic group always discrete? Note that it is for a hyperbolic group by exercise 5.57.

11. [25] Is every abelian subgroup of a biautomatic group finitely generated? If so, then it would follow from the flat torus theorem and a result of Mal'cev's that every soluble subgroup of a biautomatic group is virtually abelian.

12. Is Thompson's group biautomatic? Note that it has infinitely generated abelian subgroups so there is a strong connection with the previous question.

References

[1] J. M. Alonso, *Inégalités Isoperimetriques et Quasi-Isometries*, C.R. Acad. Sci. Paris t.311 Serie I (1990) 761-764.

[2] http://www.math.utah.edu/~bestvina/

[3] M. Bestvina, and N. Brady, *Morse Theory and Finiteness Properties of Groups*, Invent. Math. **129** (1997), 445-470.

[4] M. Bestvina and M. Feighn, *Stable Actions of Groups on Real Trees*, Invent. Math. 121 (1995), no. 2, 287-321.

[5] B. H. Bowditch, *Notes on Gromov's Hyperbolicity Criterion for Path-Metric Spaces*, in "Group Theory from the Geometrical Viewpoint" (Ed. E. Ghys, A. Haefliger and A. Verjovsky) World Scientific, Singapore (1991).

[6] B. H. Bowditch, *A Short Proof that a Subquadratic Isoperimetric Inequality Implies a Linear One*, Michigan Math J. **42** (1995), 103-107.

[7] B. H. Bowditch, *Cut Points and Canonical Splittings of Hyperbolic Groups*, Acta Math. **180** (1998), No. 2, 145-186.

[8] N. Brady, *Branched coverings of cubical complexes and subgroups of hyperbolic groups*, Jour. London Math. Soc., Vol 60 (1999), No. 2, 461-480.

[9] S. J. Brick, *Quasi-Isometries and Ends of Groups*, J. Pure Appl. Algebra **86** (1993), 23-33.

[10] M. R. Bridson and G. A. Swarup, *Equivariant Gromov Convergence and a Theorem of Paulin*, Enseignements Mathematiques (1995).

[11] R. G. Burns, *On the Intersection of Finitely Generated Subgroups of a Free Group*, Math. Z. **119** (1971) 121-130.

[12] R. Charney, *Artin Groups of Finite Type are Biautomatic*, Math. Ann. **292** (1992), no. 4, 671-683.

[13] M. Coornaert, T. Delzant and A. Papadopoulos, *Geometrie et Theorie des Groupes: Les Groupes Hyperboliques de Gromov*, Lecture Notes in Mathematics, 1441. Springer-Verlag, Berlin, 1990.

[14] M. Dehn, *Papers on Group Theory and Topology*, translated and introduced by J. Stillwell, Springer-Verlag Berlin (1987).

[15] W. Dicks and M. J. Dunwoody, *Groups Acting on Graphs*, Cambridge Studies in Advanced Mathematics 17, C.U.P. (1989).

[16] A. Dioubina, *Instability of Virtual Solvability and the Property of Being Torsion Free for Quasi-Isometric Groups*, preprint, 1999, arXiv:math.GR/9911099

[17] T. Delzant, *Sous-Groupes à deux générateurs des Groupes Hyperboliques*, in "Group Theory from the Geometrical Viewpoint" (ed. E. Ghys, A. Haefliger and A. Verjovsky) I.C.T.P. World Scientific, Singapore (1990) 177-189.

[18] D. B. A. Epstein, J. W. Cannon, D. F. Holt, S. V. F. Levy, M. S. Paterson and W. P. Thurston, *Word Processing in Groups*, Jones and Bartlett Publishers (1992).

[19] E. Ghys and P. de la Harpe, *Sur les Groupes Hyperboliques d'après Mikhael Gromov*, Birkhäuser Progress in Mathematics Series (1990).

[20] O. Goodman and M. Shapiro, *Dehn's Algorithm for non-hyperbolic groups*, preprint (2003).

[21] M. Gromov, *Groups of Polynomial Growth and Expanding Maps*, Publ. Math. I.H.E.S. Sci. 53 (1981) 53-78.

[22] M. Gromov, *Infinite Groups as Geometric Objects*, in Proceedings of the International Congress of Mathematicians (Warsaw, 1983), 385-392, PWN, Warsaw, 1984.

[23] M. Gromov, *Hyperbolic Groups*, in "Essays in Group Theory" (ed. S. Gersten) MSRI Publications No. 8. Springer (1988) 75-263.

[24] M. Gromov, *Asymptotic Invariants of Infinite Groups*, L.M.S. Lecture Note Series 182 (1993) C.U.P. Geometric Group Theory Volume 2, (Eds. G. A. Niblo and M. A. Roller).

[25] S. M. Gersten and H. Short, *Rational Subgroups of Biautomatic Groups*, Ann. Math. **134** (1991), 125-158.

[26] S. M. Gersten and H. Short, *Small Cancellation Theory and Automatic Groups II*, Invent. Math. 105 (1991) **3**, 641-662.

[27] A. Harkins, PhD Thesis, University of Newcastle upon Tyne, c. 2001.

[28] T. Herbst and R. M. Thomas, *Group Presentations, Formal Languages and Characterisations of One-Counter Groups*, Theoretical Computer Science **112** (1993) 187-213.

[29] G. Higman, *A Finitely Generated Infinite Simple Group*, J. LMS **26** (1951), 51-54.

[30] J. E. Hopcroft and J. D. Ullman, *Introduction to automata theory, languages, and computation*, Addison-Wesley Series in Computer Science. Addison-Wesley Publishing Co., Reading, Mass., 1979.

[31] A. G. Howson, *On the Intersection of Finitely Generated Free Groups*, J. LMS **29** (1954) 428-434.

[32] I. Kapovich and N. Benakli, *Boundaries of hyperbolic groups*, Combinatorial and geometric group theory (New York, 2000/Hoboken, NJ, 2001), 39-93, Contemp. Math., 296, Amer. Math. Soc., Providence, RI, 2002.

[33] I. Kapovich and D. Wise *The equivalence of some residual properties of word-hyperbolic groups*, J. Algebra 223 (2000), 562–583.

[34] M. Kapovich, *Hyperbolic manifolds and discrete groups*, Progress in Mathematics, 183. Birkhäuser Boston, Inc., Boston, MA, 2001.

[35] G. S. Makanin, *Equations in a Free Group* Math. U.S.S.R. Izvestiya 21(1983) pp 483-546

[36] Magnus, Karrass and Solitar, *Combinatorial Group Theory.*

[37] D. E. Müller and P. E. Schupp, *Groups, the Theory of Ends and Context-Free Languages*, J. Comp. and Systems Sciences **26** (1983) 295-310.

[38] H. Neumann, *On the Intersection of Finitely Generated Free Groups*, Publ. Math. Debrecen **4** (1955-56) 186-189; addendum, ibid. **5** (1957-58), 128.

[39] A. Yu. Ol'Shanskii, *Hyperbolicity of Groups with Subquadratic Isoperimetric Inequality*, Int. J. Alg. and Comp. **1** (1991) 281-289.

[40] P. Papasoglu, *Geometric Methods in Group Theory*, PhD Thesis, Columbia.

[41] P. Papasoglu, *On the Sub-Quadratic Isoperimetric Inequality*, To appear in the Proceedings of the Ohio-State Program in Geometric Group Theory (ed. R. Charney and M. Davis).

[42] P. Papasoglu, *Strongly Geodesically Automatic Groups are Hyperbolic*, Invent. Math. **121** (1995), no.2, 323-334.

[43] F. Paulin, *Outer Automorphisms of Hyperbolic Groups and Small Actions on* $\mathbb{R}$*-Trees*, article in “Arboreal Group Theory” (ed R. Alperin) Springer-Verlag MSRI series 19 (1991) 331-343.

[44] M. Pavone, *Boundaries of Discrete Groups, Toeplitz Operators and Extensions of the Reduced* C^**-Algebra*, Thesis, University of California, Berkeley, 1989.

[45] J. Pomroy, *A Characterisation of Hyperbolic Spaces*, Warwick M.Sc. dissertation (1994).

[46] A. A. Razborov, *On Systems of Equations in a Free Group* Math. U.S.S.R. Izvestiya 25(1985) pp 115-162

[47] E. Rips, *Subgroups of Small Cancellation Groups*, Bull. L.M.S. **14** (1982) 45-47.

[48] Redko and Lisovik, *Regular Events in Semigroups*, Problems of Cybernetics (1980).

[49] E. Rips and Z. Sela, *Canonical Representatives and Equations in Hyperbolic Groups*, Invent. Math. **120** (1995), no. 3, 489-512.

[50] E. Rips and Z. Sela, *Structure and Rigidity of Hyperbolic Groups I*, Geometric and Functional Analysis Vol. 4 No. 3 (1994) 337-371.

[51] E. Rips and Z. Sela, *Cyclic Splittings of Finitely Presented Groups and the Canonical JSJ Decomposition.*, Ann Math. **146** (1997) 53-109.

[52] J. J. Rotman, *An Introduction To The Theory Of Groups*, fourth edition. Graduate Texts in Mathematics, 148. *Springer-Verlag, New York*, 1995.

[53] P. Scott and C. T. C. Wall, *Topological Methods in Group Theory*, in “Homological Group Theory”, Proceedings of a symposium held in Sept. 1977 LMS **36** (1979) 137-203.

[54] J. P. Serre, *Trees*, Springer-Verlag (1980).

(Reviewer: Tatiana Smirnova-Nagnibeda) 20F18 (20E25 20F32)

[55] G. A. Swarup, *On The Cut Point Conjecture*, Electron. Res. Announc. Amer. Math. Soc. 2 (1996), no.2, 98-100.

[56] Z. Sela, *Uniform Embeddings of Hyperbolic Groups in Hilbert Spaces* Israel Jour. of Math. 80(1992) pp 171-181

[57] Z. Sela, *On the Isomorphism Problem for Hyperbolic Groups and the Classification of 3-manifolds*, Ann. Math. (2), **141**, (1995), no.2, 217-283.

[58] Z. Sela, *The Isomorphism Problem in Hyperbolic Groups II* (1993) In preparation.

[59] Z. Sela, *Endomorphisms of hyperbolic groups. I. The Hopf property*, Topology 38 (1999), no. 2, 301–321.

[60] P. Shalen, *Dendrology and its Applications*, in “Group Theory from the Geometrical Viewpoint” (ed. E. Ghys, A. Haefliger and A. Verjovsky) World Scientifiv, Singapore (1991) 543-616.

[61] J. M. Alonso, T. Brady, D. Cooper, V. Ferlini, M. Lustig, M. Mihalik, M. Shapiro and H. Short, *Notes on Word Hyperbolic Groups*, in *Group Theory from the Geometrical Viewpoint*, ed. E. Ghys, A. Haefliger and A. Verjovsky) World Scientific, Singapore, 1991, pp. 3-63.

[62] H. Short *Quasi-convexity and a Theorem of Howson's*, in "Group Theory from the Geometrical Viewpoint" (ed. E. Ghys, A. Haefliger and A. Verjovsky) World Scientific, Singapore (1991) 3-63.

[63] J. R. Stallings, *On Torsion-Free Groups with Infinitely Many Ends*, Ann. Math. (1968) **88** 312-334.

[64] M. Stoll, *Asymptotics of the Growth of 2-Step Nilpotent Groups*, J. London Math. Soc. (2) 58 (1998), no. 1, 38-48.

[65] J. Tits, *Free Subgroups in Linear Groups*, J. Alg. **20** (1972) 250-270.

[66] L. van den Dries and A. J. Wilkie, *Gromov's theorem on groups of polynomial growth and elementary logic*, J. Algebra 89 (1984), no. 2, 349-374.

[67] L. Van Wyk, *Graph Groups are Biautomatic*, J. Pure and Applied Algebra **94** (1994) 341-352.

[68] B. A. F. Wehrfritz, *Infinite linear groups. An account of the group-theoretic properties of infinite groups of matrices*, Ergebnisse der Matematik und ihrer Grenzgebiete, Band 76. Springer-Verlag, New York-Heidelberg, 1973.

[69] Mathematical Reviews 2003: Review of "The Residual Finiteness of Negatively Curved Polygons Of Finite Groups" by D. T. Wise. (Reviewer: I. Kapovich).

[70] http://zebra.sci.ccny.cuny.edu/web/

Michael Batty,
Department of Mathematics,
School Of Mathematics and Statistics,
Merz Court,
University of Newcastle upon Tyne,
Newcastle upon Tyne,
NE1 7RU,
United Kingdom.
e-mail: Michael.Batty@ncl.ac.uk

Panagiotis Papasoglu,
ADDRESS ?