Ends of Groups — a Computational Approach

Andrew David Menzies Clow

Thesis submitted for the degree of Ph.D.

Mathematics Institute

Warwick University

September 2000

Contents

List of Diagrams

Acknowledgements

I would like to first thank my research supervisor, David Epstein, for the many hours of enjoyable and helpful discussions we've had, for first teaching me geometric group theory, and getting me excited about the subject.

I would also like to thank my wife, Clare, who has given me so much support and practical help. With her help, this thesis is easier to read than it would have been, has better diagrams, and is finished earlier.

I'd like to thank the many mathematicians with whom I have had helpful discussions and who have taught me the mathematics I needed; Stephen Billington, Rob Foord and Jan-Mark Iniotakis, Martin Dunwoody, Derek Holt, Igor Mineyev, Peter Scott and Gadde Swarup.

Thanks also to my friends for making my Ph.D. studies fun, and particularly to Rob Reid, Andrew Stacey and Rob West, all three of whom have given me more than a reasonable amount of help when I've been overloaded.

Finally, I'd like to thank my family, for all the encouragement and care they've given me.

Declaration

The material in section 1.2.4, chapters 2, 3 (except section 3.1), 4 and 6 (except section 6.2) is, to the best my knowledge, original.

The remainder of chapter 1, the whole of chapter 5, and section 6.2 are expository.

Summary

We develop ways in which we can find the number of ends of automatic groups and groups with solvable word problem.

In chapter 1, we provide an introduction to ends, splittings, and computation in groups. We also remark that the 'JSJ problem' for finitely presented groups is not solvable.

In chapter 2, we prove some geometrical properties of Cayley graphs that underpin later computational results.

In chapter 3, we study coboundaries (sets of edges which disconnect the Cayley graph), and show how Stallings' theorem gives us finite objects from which we can calculate splittings.

In chapter 4, we draw the results of previous chapters together to prove that we can detect zero, two, or infinitely many ends in groups with 'good' automatic structures. We also prove that given an automatic group or a group with solvable word problem, if the group splits over a finite subgroup, we can detect this, and explicitly calculate a finite subgroup over which it splits.

In chapter 5 we give an exposition of Gerasimov's result that one-endedness can be detected in hyperbolic groups.

In chapter 6, we give an exposition of Epstein's boundary construction for graphs. We prove that a testable condition for automatic groups implies that this boundary is uniformly path-connected, and also prove that infinitely ended groups do not have uniformly path-connected boundary. As a result we are able to sometimes detect one endedness (and thus solve the problem of how many ends the group has).

Chapter 1

Introduction

1.1 Groups, graphs and ends

In this section, we introduce the terminology we use. The aim, then, is to provide a function from vocabulary to concepts, rather than to explain the concepts. For example, we sometimes make use of the concepts of CW complexes or the language of homology theory, without definition, because we feel that no ambiguity arises from doing so.

1.1.1 Preliminaries

1.1.1.1 Notation

We use the symbols $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$ to denote the natural numbers, the integers and the real numbers respectively. $\mathbb{N}$ does not contain 0. We denote the quotient group $\mathbb{Z}/n\mathbb{Z}$ by $\mathbb{Z}_n$. The closed interval from r to t in $\mathbb{R}$ is denoted by $[r, t]$.

We use the symbol $\subset$ to denote inclusion as a subset. Note that we consider a set to be a subset of itself, so for example, $\mathbb{N} \subset \mathbb{N}$. The set of elements of A which are not elements of B is denoted by $A \setminus B$. We use $\emptyset$ to denote the empty set.

Given $x \in \mathbb{R}$, we denote the smallest integer not less than x (the *ceiling* of x) by $\lceil x \rceil$, i.e. $\lceil x \rceil = \min \{z \in \mathbb{Z} \mid z \geqslant x\}$.[1]

Similarly, given $x \in \mathbb{R}$, we denote the largest integer not greater than x (the *floor* of x) by $\lfloor x \rfloor$, i.e. $\lfloor x \rfloor = \max \{z \in \mathbb{Z} \mid z \leqslant x\}$.

We often denote a sequence by $(a_n)_{n \in \mathbb{N}}$, and a finite sequence by $(a_i)_{i=1}^{n}$. (Sequences are infinite unless specified otherwise.)

Given a function $f : X \to Y$, we denote the image of f by $\operatorname{im} f$. If A is a subset of X, we may use f to define a function on this subset — the restriction of f to A — which we denote by $f|_A$.

1.1.1.2 Groups and words

Throughout, we denote the group we are considering by G. We denote the identity element of G by id_G. The group operation is denoted by juxtaposition.

When two groups G and H are isomorphic, we write $G \cong H$.

We use $\langle a_1, a_2, \ldots, a_n \mid r_1, r_2, \ldots, r_m \rangle$ to denote the group formed from the free group generated by $a_1, a_2, \ldots, a_n$ by quotienting out by the normal subgroup generated by the elements $r_1, r_2, \ldots, r_m$. Each r_i is called a *relator*. We also use the symbols $a_1, a_2, \ldots, a_n$ to denote the elements in this quotient

[1] The minimum, maximum, infimum and supremum of a set A are denoted by $\min A$, $\max A$, $\inf A$ and $\sup A$ respectively.

group. If we write $w = w'$ in place of a relator, we mean the relator $w^{-1}w'$.

A *word in a set* A is a finite sequence of elements of A, which we denote by writing its terms in order next to each other. The *length* of a word is the number of terms in the sequence, and a *prefix*, or *initial subword* of a word is a word consisting of the first n terms of the original for some n. Thus $ab^{-1}a$ is an example of a word of length 3 in the set $\{a, b, c, a^{-1}, b^{-1}, c^{-1}\}$.

For a group G with a fixed generating set, a *word of* G is a word in the set consisting of the generators and their inverses. The word is said to *represent* the element of G obtained by multiplying the terms of the sequence in the order they appear. Thus we use $ab^{-1}a$ to denote both the word and the element it represents. The *length* of an element of G is the length of the shortest word representing it.

If a group H acts on a set X and $A \subset X$, we denote the stabiliser of A by $\mathrm{stab}(A)$.

1.1.1.3 Metric spaces and paths

We rarely consider more than one metric space at a time; in fact, almost all of the time, the metric space in question is Γ, the Cayley graph of G. (See section 1.1.2.) If X is a metric space, we denote the distance in X between two points x and y of X, by $d(x, y)$. If A and B are subsets of X, we denote the Hausdorff distance between A and B by $d(A, B)$, i.e.

$$d(A, B) = \inf \{d(a, b) \mid a \in A, b \in B\}.$$

A *path* is a continuous function from a closed connected subset of $\mathbb{R}$. The domain of a path is usually $[0, 1]$. Paths may be composed, but we do not use any special notation for this. We always use the term *ray* to mean a continuous function (path) whose domain is $[0, \infty)$. By *infinite path*, we mean a ray. A path whose domain is the whole of $\mathbb{R}$ is called *bi-infinite*. A path from a to b can be considered a homotopy between the functions $\{\bullet \mapsto a\}$ and $\{\bullet \mapsto b\}$. Thus we sometimes abuse notation and write $\alpha : a \simeq b$ to mean that α is a path from a to b.

An *extension* of a path $\alpha : [0, t] \to X$ is a path $\beta : [0, t'] \to X$ with $t' > t$, or a path $\beta : [0, \infty) \to X$. In the latter case we call the extension an *infinite extension*. An *initial segment* of a path α is a path β, of which α is an extension.

A *geodesic* path is a path which is an isometry, i.e. a geodesic is a path α such that for all t and t' in the domain, $d(\alpha(t), \alpha(t')) = |t - t'|$. Of special interest in later chapters are *geodesic rays* — geodesics whose domain is $[0, \infty)$.

A set A is said to *meet* a set B, if $A \cap B \neq \emptyset$. A path α is said to *meet* a set A, if its image meets A. The number of *times* α meets A is the number of components in the inverse image of A under the function α. Similarly, if the image of a path α is a subset of a set A, we say α *lies inside* A. If α does not meet A, we say it *lies outside* A.

A subset A of a connected metric space X *disconnects* X if $X \setminus A$ is not connected. We say that A disconnects X into components $C_1, \ldots, C_n$ if $\{C_1, \ldots, C_n\}$ is the set of components of $X \setminus A$.

1.1.1.4 Graphs

A *graph* consists of a set of *vertices* together with a set of *edges*. An edge has an unordered pair of vertices with which we say it is *incident*. (These two vertices are allowed to be the same.) Sometimes we denote an edge which is incident with the vertices g and h by $[\mathsf{g}, \mathsf{h}]$.

A *directed* graph is one where the pair of vertices incident with an edge are ordered rather than unordered, so that each edge has a given direction.

A graph may be realised topologically as a 1-dimensional CW complex, with 0-cells which we call vertices, and 1-cells which we call edges. (Thus when we think of a graph as a CW complex, edges are the open 1-cells.)

We give all graphs the *path metric*, so that the distance between two points is the length of a shortest path between them, and edges have length one.

Thus we consider a graph to be a combinatorial object, a CW complex and a metric space, so that when we refer to an edge, we may think of it as an open 1-cell, or a metric subspace, or an element of the set of edges. A graph is *finite* if it has finitely many vertices and edges.

A *labelling* on a set A is a function from A to a set which we call the labels. A *labelled graph* is a graph together with a labelling on the set of edges (and sometimes also vertices). When thinking of a graph geometrically or topologically, the labelling is a labelling of the cells of the underlying CW complex, so it is the open edges (and possibly the vertices) that are labelled.

To *colour* a set is to give it a labelling, where the set of labels is finite and consists of elements that are named by colours. We then use the colours as adjectives.

A set may have more than one labelling, in which case we rely on the context to make it clear which we refer to. For example, colourings are labellings, but are never referred to as such, and to *mark* parts of a graph is to label them.

An *orientation* on an edge is a choice of *initial* and *terminal* vertices incident with it. We say that the edge is *oriented* from the initial vertex to the terminal vertex, or *points towards* the terminal vertex. Note that a directed edge can have an orientation inconsistent with its direction.

Orientation and direction are similar concepts; however, an orientation is equivalent to a labelling by an ordered pair, whereas a direction is part of the given structure of a graph. We never use the terms initial vertex or terminal vertex to refer to the direction on an edge.

An *edge path* from a to b is a finite sequence $(e_i)_{i=1}^{n}$ of oriented edges, where the initial vertex of e_1 is a, the terminal vertex of e_n is b, and for all $i \in \{1, \dots, n-1\}$, the terminal vertex of e_i is the initial vertex of e_{i+1}. Thus an edge path is not a path in the ordinary sense, but rather, determines a path.

A path α in a graph is said to *cross* an edge $e = [a, a']$ if there is a subset $[t, t']$ of the domain of α such that $\alpha(t) = a$, $\alpha(t') = a'$, and for all $x \in [t, t']$, $\alpha(x) \in e$.

A *spanning subgraph* is a subgraph which meets every vertex of the original.

A *tree* is a graph without loops — one where every vertex disconnects the graph.

1.1.2 The Cayley graph, Γ

Given a group G with a fixed finite generating set, we construct a directed, labelled graph Γ, called the *Cayley graph.* We suppress G from the notation, because we never consider more than one Cayley graph at a time.

The vertices of Γ are the elements of G, and the edges are the triples $\{(g, ga, a) \mid g \in G, a \in X\}$, where X is the set of generators of G. The edge (g, ga, a) is incident with g and ga, and is directed from g to ga. We say that an edge (g, ga, a) of Γ is *labelled with* a.

$$g \xrightarrow{a} ga$$

Diagram 1.1.1: An edge of Γ

Note that a (directed) edge is determined by its starting point and the generator with which it is labelled.

Since G is finitely generated, and for each $g \in G$, there is only one vertex ga, there are only finitely many edges incident with each vertex; Γ is a locally finite graph.

1.1.2.1 Geometry

As a graph and a metric space, Γ inherits all the definitions above. In particular, $d(x, y)$ usually refers to the distance between two points that are in the Cayley graph (and the Cayley graph has the usual path metric on graphs, as outlined above).

Let g be a vertex of Γ (i.e. an element of G). The open ball of radius n about g is written $B^{\circ}_{n}(g)$, and the closed ball is written $\overline{B}_{n}(g)$. The sphere at radius n about g is denoted by $S_n(g)$. This notation always refers to subsets of the Cayley graph. Of course, since the vertices of Γ are the elements of G, we may refer to elements of G which lie in $\overline{B}_n(id_G)$, for example.[2] For much of the time, we are concerned with balls centred at the identity element of G; we drop the centre, g, from the notation when $g = id_G$, so balls and spheres about the identity are called B°_n, $\overline{B}_n$, and S_n.

Note that if a geodesic in a graph has its endpoints on vertices, then if it meets the interior of an edge, it must cross the edge. Thus if such a geodesic does not meet $\overline{B}_{n-1}$, it cannot meet the interior of any edge incident with a vertex of S_{n-1}, so it cannot pass closer than distance n to the identity. Thus such a geodesic lies outside $\overline{B}_{n-1}$ if and only if it lies outside B°_n.

Recall that a ray is a path with domain $[0, \infty)$. Unless otherwise specified, rays in Γ start at the identity element.

Note that since the set of generators generates G, there is an edge path in Γ from the identity to any element of G, determined by a word in the generators representing it. This edge path determines a topological path with the same endpoints. Thus G is connected and path-connected. In fact, Γ is locally path-connected, so for subsets of Γ, path-components and components are the same.

1.1.2.2 The G-action on Γ

Given an element $g \in G$, and an edge $e = (h, ha, a)$ of Γ, we define

$$ge = (gh, gha, a).$$

[2]Note that the length of an element (defined as the length of a shortest word representing it), is the same as its distance from the identity in the Cayley graph.

Note that (gh, gha, a) really is an edge of Γ, because $(gh)a = gha$. Thus G acts by left multiplication on the vertices of Γ, and the action on edges is determined by the action on the vertices.

Note that the G-action on Γ preserves the direction of edges. Since G preserves the incidence relations between vertices and edges, it preserves the structure of the graph, so it preserves the length of paths, and so G acts by isometries on Γ.

G acts on itself by left multiplication with one orbit, and so the G action on the vertices Γ has only one orbit.

If two edges $e_1 = (h_1, h_1a, a)$ and $e_2 = (h_2, h_2b, b)$ are in the same orbit, then there exists $g \in G$ such that $(gh_1, gh_1a, a) = (h_2, h_2b, b)$, so $a = b$. Thus e_1 and e_2 are both labelled with the generator a.

Conversely, for any h_1 and h_2 in G, the edges (h_1, h_1a, a) and (h_2, h_2a, a) are in the same orbit, because $h_2h_1^{-1}(h_1, h_1a, a) = (h_2, h_2a, a)$.

Thus two edges are in the same orbit if and only if they are labelled with the same generator.

As usual, we extend the action of G on Γ to an action on the set of subsets of Γ; if A is a subset of Γ, a G-*translate* of A is the set gA for some $g \in G$.

1.1.2.3 Hyperbolic groups

A group G is said to be *hyperbolic* if there is some fixed constant δ', so that for any triple of geodesics (α, β, γ) that form a triangle in Γ,[3] every point in the image of α is at most distance δ' from the union of the images of the

[3]in the sense that there exist points a, b and c such that $\alpha : a \simeq b$, $\beta : b \simeq c$ and $\gamma : c \simeq a$

other two paths.

We touch only briefly on hyperbolic groups, mainly in chapter 5. The reader may want to refer to [GdlH90] or [GHV91], or to other literature for a fuller account of the theory of hyperbolic groups.

1.1.3 Ends and splittings

1.1.3.1 Ends

A locally compact, Hausdorff topological space X has at least e *ends* if there is a compact set K such that $X \setminus K$ has at least e components which have non-compact closure. Since the Cayley graph is a metric space, this is equivalent to saying that for some radius, the complement of the open ball has at least e infinite components.[4] We say X has e *ends* if it has at least e ends and for any $e' > e$, X does not have at least e' ends. If there is no bound on the number of ends, we say that X has infinitely many. (X has zero ends if it is compact.)

A group has e ends if its Cayley graph does. It is a theorem that the number of ends is independent of the generating set taken. It is also a theorem that a group has either zero, one, two or infinitely many ends. See, for example, [SW] for an introduction to theory of ends of groups.

Since G acts by isometries on Γ, for any $g \in G$, $\Gamma \setminus B_n^\circ$ is isometric to $\Gamma \setminus B_n^\circ(g)$; the number and nature of components of the complement of $B_n^\circ(g)$

[4]It may be the case that there are fewer components of the complement of an open ball than in the complement of the closed ball. In this case we may increase the radius and will find at least as many as we needed. This is an unimportant point.

and of B_n° are the same.

1.1.3.2 Splittings

Let A and B be groups with presentations $\langle a_1, \ldots, a_{n_A} \mid r_1, \ldots, r_{m_A} \rangle$ and $\langle b_1, \ldots, b_{n_B} \mid s_1, \ldots, s_{m_B} \rangle$ respectively. Let C be a group with generators $c_1, \ldots, c_{n_C}$, and let $f : C \to A$, $f' : C \to A$ and $g : C \to B$ be injective homomorphisms.

The *free product of* A *and* B *amalgamating* C is the group given by the presentation

$$\langle a_1, \ldots, a_{n_A}, b_1, \ldots, b_{n_B} \mid f(c_1) = g(c_1), \ldots, f(c_{n_C}) = g(c_{n_C}),$$
$$r_1, \ldots, r_{m_A}, s_1, \ldots, s_{m_B} \rangle.$$

When C is the trivial group, this is the *free product* of A and B.

The *HNN extension of* A *over* C is the group given by the presentation

$$\langle t, a_1, \ldots, a_{n_A} \mid f(c_1)t = tf'(c_1), \ldots, f(c_{n_C})t = tf'(c_{n_C}), r_1, \ldots, r_{m_A} \rangle.$$

If G is isomorphic to a free product with amalgamation over C or to an HNN extension over C, we say G *splits* over C. C injects into G, so we commonly think of C as a subgroup of G.

1.1.3.3 Stallings' theorem

Stallings' theorem states that a group has more than one end if and only if it splits over a finite subgroup. An outline of the proof appears in section 3.1 on page 65, and accounts may be found in [Sta68], [DD89], [SW], and others.

1.2 Computation in groups

1.2.1 Automatic groups

An automatic group is a group which has certain structures which make computation straightforward. [ECH+92] contains an introduction to the theory of automatic groups. In the terminology of [ECH+92], by *finite state automaton*, we mean a partial deterministic automaton.

1.2.1.1 Finite State Automata

A *finite state automaton* (*FSA*) is a directed graph A, whose edges are labelled by the elements of a set S called the *alphabet*. The vertices of A are called its *states*. There is a unique state called the *start state*, and some special states which are called *accept states*. Two edges incident with and directed away from the same vertex are not allowed to have the same label.

A *path in* A is an edge path in A that begins at the start state,[5] and whose edges have the orientation consistent with the direction of the edges of the directed graph A; edges of a finite state automaton are one-way only, and paths in an FSA are not allowed to go the wrong way. A *loop* is a path which ends at the same state it began.

A path in A (starting from the start state, as usual), determines a sequence of labels from S. Thus the path determines a word in the alphabet S. If the accept state at which the path finishes is an accept state of A, we say that A *accepts* the word.

[5]unless otherwise specified

The set of all paths accepted by A determines a set of accepted words which we call the *accepted language*.

1.2.1.2 Automatic structure of a group

Let G be a group with a fixed finite generating set.

A *word acceptor* for G is a finite state automaton WA, over the alphabet S equal to the set of generators and their inverses, such that every element of G can be written as a word accepted by WA.

We choose a symbol, \$, not amongst the generators of G, which we call a *padding symbol.*

In essence, a *general multiplier* for G is an automaton that accepts pairs of words (as words of pairs) that differ by a generator. More formally:

Given a word acceptor WA for G, a *general multiplier* for G is a finite state automaton GM over the alphabet $(S \cup \{\$\}) \times (S \cup \{\$\})$, with the following properties:

- The accept states of GM are labelled with $S \cup \{id_G\}$.
- GM accepts $(x_1, y_1)(x_2, y_2) \cdots (x_n, y_n)$ in an accept state labelled a if and only if the words $w = x_1 x_2 \cdots x_n$ and $w' = y_1 y_2 \cdots y_n$ satisfy the following properties:
  - After removing all padding symbols from w and w', they represent elements g and g' of G such that $ga = g'$.
  - Padding symbols occur only at the end of w or w'.

An *automatic structure* for G is a finite generating set, a word acceptor WA and a general multiplier GM. A group is said to be *automatic* if it has an automatic structure.

Sometimes an automatic structure is given by a word acceptor and a set of *multiplier automata*, $\{M_a \mid a \in S \cup \{id_G\}\}$, where M_a satisfies the same conditions as GM, except that it only accepts pairs of words which differ by a. It can be shown that the two are equivalent.

1.2.1.3 Languages

We usually use the term *accepted language* to refer to the language accepted by the word acceptor of an automatic group.

A language is said to be *prefix closed* if every prefix of every accepted word is also accepted.

The accepted language of an automatic group is said to have *unique representatives* if for all $g \in G$, there is one and only one word accepted by WA that represents g.

We say that a group is *short-lex*-automatic if it has unique, geodesic representatives, and there is a total ordering on the generators such that for all $g \in G$, the accepted word representing g comes first in the lexicographic (dictionary) order amongst all words of G that are equal to g.

1.2.1.4 Fellow-travelling

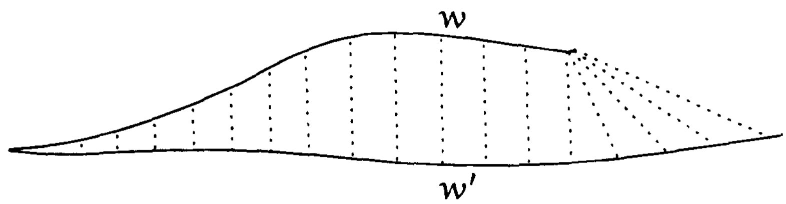

Diagram 1.2.1: Fellow-travelling

Let w and w', be words in G of length n and $n+m$ respectively. We say that w and w' f-fellow-travel if the distance in Γ between them is always bounded above by f; more precisely, they f-fellow-travel if

$$\forall i \in \{1, \ldots, n\}, \quad d(w(i), w'(i)) \leqslant f, \text{ and also}$$

$$\forall i \in \{1, \ldots, m\}, \quad d(w(n), w'(n+i)) \leqslant f.$$

Theorem 2.3.5 of [ECH+92] proves that for every automatic group G, there is a constant f such that every pair of words that represent elements of G that differ by a generator, the inverse of a generator or do not differ, f-fellow-travel. We call f the *fellow-travelling constant.*

1.2.1.5 Working with automatic groups

The package *KBMAG* is available via `ftp` from a link on Derek Holt's website `http://www.maths.warwick.ac.uk/~dfh/`. It consists of several programs for working with automatic groups. For example, the program *autgroup* can take a presentation for a group, and if the group is short-lex-automatic, calculate a short-lex automatic structure (given sufficient time and space).

1.2.2 General computation

In this section we explain our use of terminology for general programming.

1.2.2.1 Algorithmic decidability

The formal mathematical definition of whether a problem is solvable with an algorithm is given in terms of *Turing machines.* We do not adopt this approach, since we wish to show that certain algorithms exist in the sense that they could, in principle, be turned into a program. To turn an algorithm into a Turing machine would be time and space consuming, and probably would not help the reader to understand how the algorithm works. Instead, we describe the algorithm in less formal terms. Algorithm is to Turing machine as proof is to fully formal logical derivation.

Given a property P(O) of objects O, we say we can *detect* the property P if there is an algorithm which takes any valid object O, and terminates within a finite time with the output yes, if and only if O has property P. This corresponds to the standard terminology 'recursively enumerable'.

We say we can *decide* property P if we can detect both P and its negation. Since algorithms may be run concurrently, P is decidable if there is an algorithm which always terminates in a finite amount of time, with the output yes or no, depending on whether P is satisfied or not.

1.2.2.2 Depth first searches

Depth-first searches of graphs are well-known tools in computer science, and we use them in several forms in the algorithms and programs we describe.

Here the word 'depth' refers to the length of paths transversed whilst searching. 'Depth-first' refers to the fact that in this method of searching a graph, we follow a path as deep (i.e. long) as we can before retracing our steps and trying a different path.

There are two essential uses we make of depth-first searches. The first is to find all possible paths in a given graph, and the second is to calculate the components of a graph. The second type is more conventional in computation, and its purpose is to visit every vertex (performing some action, but not more often than necessary). It can be implemented as an altered form of the first.

The graph in question may be a subgraph of the Cayley graph Γ, or a finite state automaton. In the case of a subgraph of Γ, we do not construct a copy of the subgraph, but instead run the depth-first search in Γ itself, after marking certain edges *uncrossable* or vertices as *unreachable.*

There is an analogous process called a breadth-first search, but breadth-first searches can be both harder to implement, and less efficient.

Path depth-first search

Sometimes, we only wish to find paths of length at most D. If so, we perform a *path depth-first search to depth* D. Otherwise we perform a *path depth-first search.*

We find all paths and calculate the words that these paths define via the labels on the edges of the graph.

The algorithm is defined in terms of itself, in a procedure akin to reverse induction, but with an unfortunate clash of terminology, known to computer scientists as recursion. The algorithm 'calls' itself, and the called version 'returns' control and information to the caller.

Visiting a vertex We assume we are given a current depth d. (Initially we are given $d = 0$.

1. If we are performing a search to depth D, and $d = D$, return the set containing the empty word to the caller.

2. Visit each vertex incident with the current vertex (at depth $d + 1$). (But do not use uncrossable edges or visit unreachable vertices.) For each vertex visited, a set of words is returned. At the front of each returned word, write the label of the edge from the current vertex to the visited vertex.

3. Return the set of all these words to the caller.

To perform a path depth-first search starting at a vertex v, simply visit v. Once this is complete, a list of all words is obtained.

Vertex depth-first search

Sometimes, we only wish to find vertices at most D from our start vertex. If so, we perform a *vertex depth-first search to depth* D. Otherwise we perform a *vertex depth-first search*.

The algorithm uses recursion in the same way as the path depth-first search.

Visiting a vertex We assume we are given a current depth d. (Initially we are given $d = 0$.)

1. Mark the vertex as *visited.*
2. If we are performing a search to depth D, and $d = D$, return control to the caller.
3. Visit each vertex incident with the current vertex (at depth $d + 1$). (But do not use uncrossable edges or visit unreachable vertices, nor visit vertices already marked as visited.)
4. Return control to the caller.

To perform a vertex depth-first search starting at a vertex v, simply visit v. Once this is complete, all reachable vertices have been visited.

1.2.3 Solvable and unsolvable problems in group theory

We outline a number of the algorithmic problems in group theory, and prove that the JSJ problem is, in general, unsolvable.

1.2.3.1 The word problem

Given a group G with a fixed finite generating set, we say that G has *solvable word problem* if there is an algorithm that decides whether an arbitrary word of G represents id_G.

An automatic group has solvable word problem, but not every group with solvable word problem is automatic. Not every group has solvable word problem.

1.2.3.2 The triviality problem

A class of groups has *solvable triviality problem* if there is an algorithm which takes presentations of groups that are in the class, and decides whether they are trivial (have only one element).

The triviality problem is not solvable for the class of all groups, but is solvable for the class of all groups which have solvable word problem.

1.2.3.3 The isomorphism problem

A class of groups has *solvable isomorphism problem* if there is an algorithm which takes pairs of presentations of groups that are in the class, and decides whether they are isomorphic.

The isomorphism problem is not solvable for the class of all groups, because otherwise the triviality problem would be solvable. It is conjectured in [ECH+92] that the isomorphism problem for automatic groups is not solvable. Sela [Sel95] has published only half of his solution to the isomorphism problem for the class of hyperbolic groups. One step in the process of deciding whether two hyperbolic groups are isomorphic is determining whether a given group splits as a free product or splits over a finite subgroup. Gerasimov [Ger] proved an algorithm exists to do so.

1.2.3.4 The JSJ problem

We leave the concept of a JSJ decomposition of a group undefined. The interested reader may consult [DS99], [FP97], [Bow95] or [RS95] for definitions and existence proofs in different classes of groups.

In essence, the JSJ decomposition of a group G is a labelled graph (a graph of groups) which encapsulates the information about all the splittings of G over 2-ended subgroups.

Dunwoody and Sageev [DS99] generalise this to cover splittings over 'slender' subgroups, whilst Fujiwara and Papasoglu [FP97] use 'foldings' of graphs to express splittings over more than one class of subgroup at the same time.

The JSJ problem is solvable for a class of groups if there is an algorithm which takes a presentation for a group in the class that has one end and calculates its JSJ decomposition, including the nature of its hanging orbifold subgroups.

In a hyperbolic group, the non-orbifold vertex groups are rigid, and have solvable isomorphism problem. For this reason, Sela is able to solve the isomorphism problem for torsion-free hyperbolic groups by, in effect, solving the 'JSJ problem', the free product problem, and the isomorphism problem in the case of rigid groups. This work is to appear in *The Isomorphism Problem for Hyperbolic Groups II*, which is unfortunately not yet available as a preprint.

1.2.3.5 Related work: the free product problem

Clearly, the problem of whether a group given by finite presentation splits as a free product is not solvable for arbitrary finitely presented groups either — $G \times \mathbb{Z}$ splits as a free product if and only if G is trivial.

As mentioned above, Sela can solve the free product problem for torsion-free hyperbolic groups, but the proof is unavailable in written form. The algorithm he outlines seems infeasible as a real computer program, due to the complexity of the method he uses.

Gerasimov [Ger] proved that the free-product problem is solvable for arbitrary hyperbolic groups. In fact, he showed that the generalised problem of whether a hyperbolic group splits over a finite subgroup is decidable. We give an altered exposition of his result in chapter 5 on page 134.

1.2.3.6 More related work: ends of automatic graphs

Note that by Stallings' theorem, the problem of deciding whether an infinite group splits over a finite subgroup is equivalent to the problem of deciding whether the group has more than one end or not.

Olivier Ly [Ly00] defines the notion of an 'automatic graph' — one that may be generated using finite state automata, and proves that for arbitrary automatic graphs, the problem of deciding the number of ends is not solvable. Note that there are many automatic graphs which are not the Cayley graph of a group.

In chapter 4 on page 91 we show that for groups with 'good' automatic structures, the properties of having more than one end, and of having zero ends, are detectable. In chapter 6 on page 144, we show that there is an algorithm which *may* detect the property of having one end.

1.2.4 The JSJ problem for arbitrary finitely presented groups is unsolvable

It is not clear that anyone has noted the following, easily proven result; it is the only original comment in this otherwise expository chapter.

Theorem 1.2.1

The problem of determining the JSJ decomposition of a group from its presentation is not solvable for the class of all finitely presented groups.

Proof

We argue by contradiction. Suppose there is an algorithm which deduces the JSJ decomposition of a group from its finite presentation. Let G a group with a finite presentation. Then the direct product $G \times \mathbb{Z} \times \mathbb{Z}$ has a finite presentation which may be computed easily from that of G.

Since $\mathbb{Z} \times \mathbb{Z}$ is a surface, hence orbifold group, its JSJ decomposition is the graph of groups consisting of a single (hanging orbifold) vertex group, $\mathbb{Z} \times \mathbb{Z}$. If G is non-trivial, then $G \times \mathbb{Z} \times \mathbb{Z} \ncong \mathbb{Z} \times \mathbb{Z}$, so in this case, the JSJ decomposition is *not* a single vertex labelled $\mathbb{Z} \times \mathbb{Z}$.

Thus, calculating the JSJ decomposition for $G \times \mathbb{Z} \times \mathbb{Z}$ solves the triviality problem for G. But G was an arbitrary finitely presented group, and there can be no solution to the triviality problem for arbitrary finitely presented groups. Contradiction.

□

Note that in calculating the JSJ decomposition, we must find which $\mathbb{Z}$-splittings 'cross' each other and give rise to hanging orbifold vertex groups — the JSJ problem can only be solvable in a class of groups for which it is possible to decide whether a presentation is that of a surface (or orbifold) group.

Note also that one may be able to determine a JSJ decomposition without determining the isomorphism class of the group. As mentioned, it is necessary to determine the isomorphism class of the hanging orbifold subgroups, but one could conceivably determine a graph of groups decomposition where the non-orbifold vertex groups were given only by a presentation.

Chapter 2

Complements of balls in the Cayley graph

Recall that the Cayley graph, Γ, has at least e ends if the complement of the open ball of some finite radius has at least e infinite components.

In this chapter we deal with three issues relating to these complements of open balls. Firstly, how much of $\Gamma \setminus B_n^\circ$ do we need to examine, to be able to know what components it falls into? When computing, we can only calculate a finite part of the Cayley graph, and it is important to minimise the amount we need to calculate. In section 2.1 we show that there is a constant κ such that we only need to look at the ball of radius $n + \kappa$ if we want to find the components of $\Gamma \setminus B_n^\circ$.

In section 2.2, we prove a result which says that if B_n° disconnects the Cayley graph and there are at least two components of its complement that contain elements at distance $2n$ from the identity, then these components are infinite, so G has more than one end. This result is useful when we don't have an automatic structure for the group.

In section 2.3, we relate the question of whether components of $\Gamma \setminus B_n^\circ$ are infinite, to whether they contain rays. We use this twice: in an automatic group we can test very quickly which components of $\Gamma \setminus B_n^\circ$ are infinite, by checking in the word acceptor which elements can be infinitely extended. We also use these results in section 6.3 to show that 'patch shadows' are closed and open subsets of a boundary for the group, and that if the group has infinitely many ends, this boundary is not nicely connected.

2.1 Finding components: the connectivity constant, κ

Suppose Γ is a Cayley graph of a finitely presented group G. We wish to prove that there is a number κ depending only on the presentation of G such that the Γ-components[1] of $\overline{B}_{n+\kappa} \setminus B_n^\circ$ and the components of $\Gamma \setminus B_n^\circ$ are in one-to-one correspondence under inclusion; this will be the main theorem of the section (theorem 2.1.12 on page 51), and will allow us to represent infinite components of Γ as finite subsets of G.

Definition 2.1.1 (κ) *Fix a finite presentation of a group* G, *and denote its Cayley graph by* Γ. *Consider the length of each relator and choose the longest. Halve this length, round down to the nearest integer and subtract 1. This is the* connectivity constant *of* G, κ, *and for short we write*

$$\kappa = \left\lfloor \frac{1}{2}(\textit{max relator length}) \right\rfloor - 1.$$

[1] see 2.1.1.2 on page 35

Definition 2.1.2 *We call* $\overline{B}_{n+\kappa} \setminus B^{\circ}_{n}$ the band.

2.1.1 Coboundaries, Γ-connectedness, closeness of edges

We introduce some definitions enabling us to prove a generalised theorem, which will be used to prove the main theorem, but also a later result about the diameter of coboundaries.

2.1.1.1 Coboundaries

Notation 2.1.3 (complement) *Let* G *be a group, and let* S *be any subset of* G*. We denote the complement,* $G \setminus S$ *by the symbol* $S^{\star}$*.*

We use the term coboundary in a slightly different sense than in cohomology. Accordingly, we make the following definition:

Definition 2.1.4 (coboundary) *Let* G *be a finitely generated group, and let* Γ *be its Cayley graph. Let* S *be a subset of* G*. The* coboundary *of* S *is the set of edges of* Γ *which meet both* S *and its complement* $S^{\star}$*. We denote it by* δS*.*

Note S determines a characteristic function $\chi_S : G \to \mathbb{Z}_2$, with $g \in S \Leftrightarrow \chi_S(g) = 1$. We can think of χ_S as a 0-cochain with coefficients in $\mathbb{Z}_2$. As such it has coboundary (in the ordinary sense) equal to $\chi_{\delta S}$. This equality is why we use the term coboundary for δS.

We are only ever interested in subsets of G with finite coboundary, because we are interested in splittings over finite subgroups.

2.1.1.2 Γ-connectedness

Given a set of vertices S of Γ (i.e. elements of G), it is helpful to consider S as equivalent to the subset of Γ given by the union of the elements of S and all edges of Γ which have both vertices in S. Accordingly, we use the following terminology:

Definition 2.1.5 (Γ-within) *Let* G *be a finitely presented group with Cayley graph* Γ*. Given a set* S *of elements of* G *and an edge* e *of* Γ*, we say that* e *is* Γ-within S *if it has both its vertices lying inside* S.

Definition 2.1.6 (Γ-connectedness) *Let* G *be a finitely presented group with Cayley graph* Γ*. We say that a set* S *of elements of* G *is* Γ-connected *if for every pair of vertices of* S*, there is an edge path between them which has each edge Γ-within* S.

A Γ-component *of* S *is a maximal Γ-connected subset of* S.

2.1.1.3 Closeness of edges

Here we prove a lemma useful in the theorem that follows.

Definition 2.1.7 *Let* G *be a group with a fixed finite presentation, and let* Γ *be its Cayley graph. Let* e *and* e′ *be edges of* Γ.

We say that e *and* e′ *are* on the same relator *if there is an element* $g \in G$ *and a relator* r *such that starting from* g*, the word* r *determines an edge path in* Γ *which crosses both* e *and* e′.

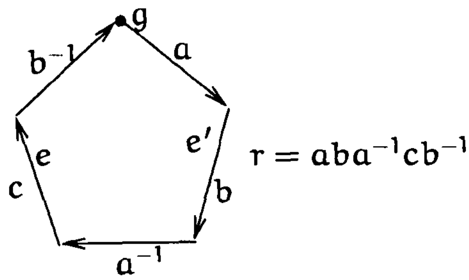

Diagram 2.1.1: Edges on the same relator

Lemma 2.1.8

Let G *be a group with a fixed finite presentation. Let* Γ *be its Cayley graph, and let* κ *be the connectivity constant of* G.

If $(s_j)_{j=1}^{m'+1}$ *is a sequence of edges of* Γ *such that for all* $j \in \{1, \ldots, n\}$, s_j *and* s_{j+1} *are edges on the same relator, then*

$$\forall j \in \{1, \ldots, n\}, \quad d(s_j, s_{j+1}) \leqslant \kappa.$$

Proof

Let $j \in \{1, \ldots, n\}$. s_j and s_{j+1} are on the same relator. Then there is an element $g \in G$, and a relator r such that starting from g, the word r determines an edge path in Γ which crosses both s_j and s_{j+1}.

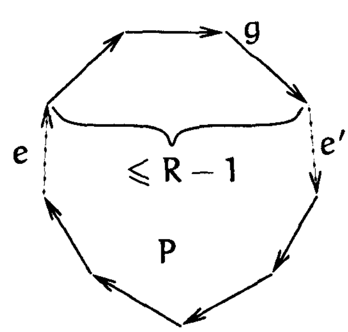

Diagram 2.1.2: P is the graph of the relator r.

Let P be the graph determined by the relator r; P is an n-gon, where n is the length of the word r. We label a vertex of P by g, and label one edge of P by s_j and the other by s_{j+1}, according to the position on the word r that s_j and s_{j+1} occur in Γ.

Let $R = \frac{1}{2}$(max relator length). P has the path metric, and has diameter at most R. If we remove the interiors of s_j and s_{j+1} from P, it falls into two path components.

If both of these path components had diameter greater than $R - 1$, then the word r would have length strictly greater than $2(R-1)+2 = 2R$. This cannot happen, because $R = \frac{1}{2}$(max relator length).

Thus one of the path components has diameter less than or equal to $R - 1$, thus less than or equal to $\lfloor R \rfloor - 1$, since it is of integer length; i.e. $d(s_j, s_{j+1}) \leqslant \lfloor R \rfloor - 1 = \kappa$.

□

2.1.2 Joining pairs of edges

We use the more generalised setting of coboundaries to prove the results we need about Γ-components of the band $\overline{B}_{n+\kappa} \setminus B^\circ_n$ and components of $\Gamma \setminus B^\circ_n$.

It is helpful in the theorem below to think of S as the vertices of B°_n (i.e. the vertices of $\overline{B}_{n-1}$) and $S^\star$ as the vertices of $\Gamma \setminus B^\circ_n$; indeed, this is how we will use the result. Being within distance $\kappa + 1$ of a vertex of B°_n is the same as being in $\overline{B}_{n+\kappa} \setminus B^\circ_n$.

Theorem 2.1.9

Let κ be the connectivity constant of G, and let S be a subset of G. Let

$e = [g_1, g_1 a]$ *and* $e' = [g_2, g_2 b]$ *be edges of the coboundary* δS, *with both* g_1 *and* g_2 *in* $S^\star$. *Suppose there is a path* α_1 *between* g_1 *and* g_2 *which does not meet* S, *and also that there is a path* α_2 *between* $g_2 b$ *and* $g_1 a$ *which does not meet* $S^\star$.

Then

1. *There is an edge path* $(e_i)_{i=1}^m$ *from* g_1 *to* g_2, *which is* Γ*-within* $S^\star$, *in the sense that the sequence* $(v_i)_{i=1}^{m+1}$ *of vertices it visits are all elements of* $S^\star$.

   The vertices of this edge path all lie within distance $\kappa + 1$ *of a vertex of* S.

2. *There is a sequence* $(s_j)_{j=1}^{m'+1}$ *of edges of* δS *such that* $s_1 = e$ *and* $s_{m'+1} = e'$, *and* $\forall j \in \{1, \dots, m'\}$, $\quad d(s_j, s_{j+1}) \leqslant \kappa$.

Proof

The proof is not very short, and rather than split it into a sequence of lemmas with similar hypotheses, we provide section headings.

We may remove unnecessary loops from the paths α_1 and α_2 so that they are both injective. The images of these two paths do not meet each other (because S and δS have empty intersection), and meet e and e' only at the endpoints g_1, g_2, $g_2 b$, $g_1 a$. Thus there is an embedded loop in Γ, passing across e and e'.

The Van Kampen diagram

This loop describes a word in the generators of G that represents a trivial element of the group. We may take a Van Kampen diagram for this word. There is a simplicial dimension-preserving map from the 1-skeleton of the Van Kampen diagram to Γ, which takes the boundary of the Van Kampen diagram to the loop in Γ we started with. We will use this map to label the vertices and edges of the diagram.

It must be stressed that we map the boundary of the Van Kampen diagram round the loop we started with, not starting at id_G as is the norm for Van Kampen diagrams. (Unless, of course, our loop meets id_G anyway.) In essence, we are translating the image of the diagram to where the loop occurs in Γ.

The diagram has a polygonal 2-complex structure.[2] The vertices are labelled by vertices of Γ (elements of G) and the edges are labelled by edges of Γ. The 2-cells (polygons) are labelled by relators of G.

Since the loop in Γ is embedded, neither the word it describes, nor any cyclic conjugate of it, has a proper subword representing a trivial element of G. An arbitrary Van Kampen diagram is planar, and is the union of disks and arcs.

[2]A ***polygonal 2-complex*** is a 2-dimensional CW complex whose 2-cells are open regular polygons of edge length one, whose 1-cells are isometric with the open interval $(0, 1)$, and where the attaching maps from the closed polygons to the 1-skeleton are determined by isometries of the interiors of the edges.

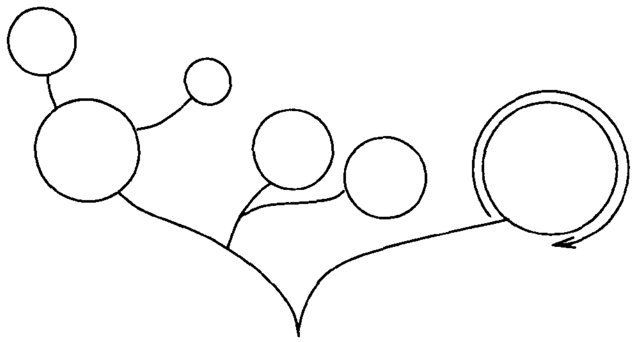

Diagram 2.1.3: An example of an arbitrary Van Kampen diagram

However, the boundary of each such disc region is labelled by a subword of the boundary word which is trivial in G (property (2) in §2 of [Str90]). Since there is no such subword for our word, the Van Kampen diagram must consist of only one disk, with no arcs.

Red and blue parts of the Van Kampen diagram

We colour all the cells of the Van Kampen diagram either red or blue. We colour vertices *red* if they are labelled by elements of S, and we also colour edges and 2-cells *red* if they are incident on red vertices. All other cells and vertices are *blue*. So in particular, vertices of $S^\star$, and edges which are Γ-within $S^\star$ are coloured blue.

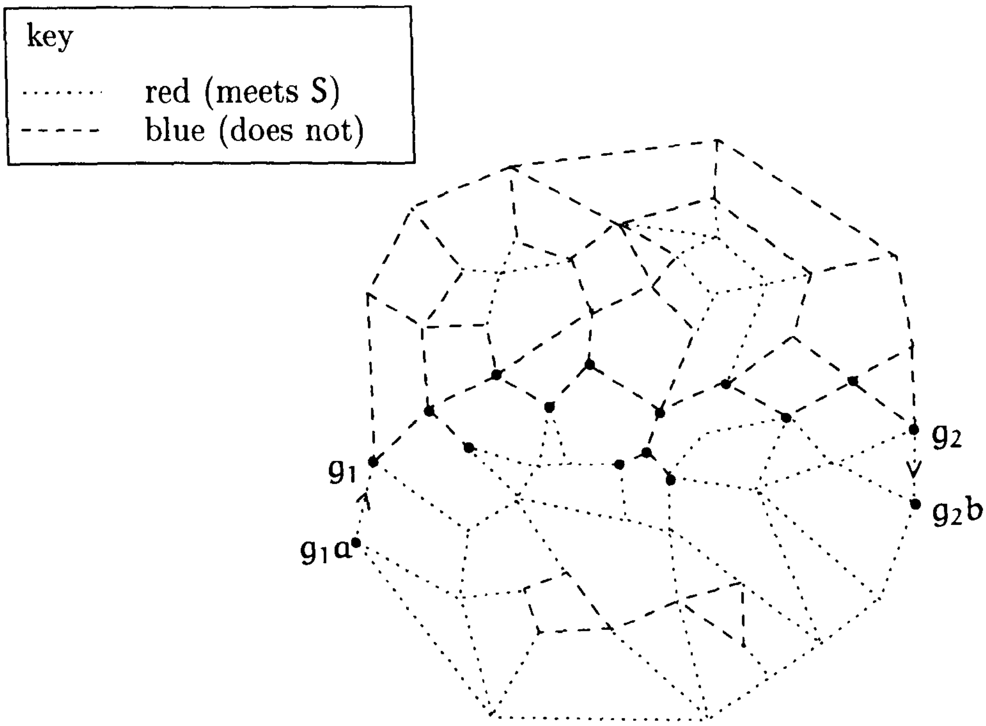

Diagram 2.1.4: Red and blue parts of the Van Kampen diagram

Firstly, note an edge of a red relator cell lies on the same relator as an element of S, and so each vertex on it is at most distance $\lfloor \frac{1}{2}(\text{max relator length}) \rfloor$ from a red vertex.[3] Now the distance between points in the Van Kampen diagram is greater than or equal to that between their images in Γ; to prove part 1 it is enough to construct the edge path so that its edges and vertices are blue (Γ-within $S^\star$) and are edges and vertices of red relator cells (so that the vertices are within distance $\kappa + 1$ of S).

Note that edges are labelled by elements of δS if and only if they have one vertex blue and the other red. (Such edges are themselves red.) To prove part 2 we need to find a sequence of red edges each with a single blue vertex, such that successive pairs lie on the same relator.

[3] Vertices are an integer distance apart.

From now on for shorthand, we may talk about vertices, edges and 2-cells of the Van Kampen diagram as if they were in Γ, rather than merely labelled by elements of Γ.

According to our colouring, the path α_1 between g_1 and g_2 that lies Γ-within $S^\star$ is represented by the blue part of the boundary circle of the Van Kampen diagram, and the path α_2 from g_1 to g_2 that is Γ-within S is the red remainder of the boundary circle. Diagrammatically, we put g_1 on the left, g_2 on the right, the blue boundary path on the top and the red boundary path on the bottom.

Orienting the 2-cells

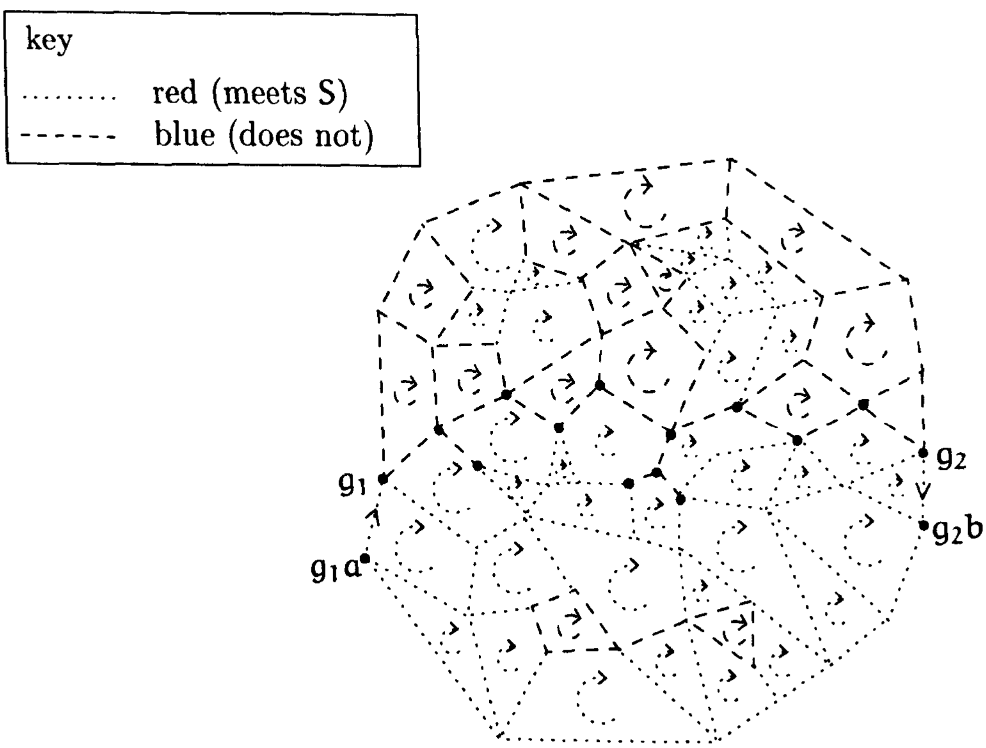

Diagram 2.1.5: Orienting the 2-cells

g_1a and g_2b are the second and penultimate vertices respectively on the red boundary path from g_1 to g_2. (The notation is meant to suggest that they differ from g_1 and g_2 by a generator.) Diagrammatically, we orient the relator cells clockwise, to correspond to the following definition; the pair (g_1a, g_1) is an ordering on the edge from g_1a to g_1. This determines an ordering on the (unique) relator cell incident with it. We orient all the other 2-cells coherently with this (so that considered as a homology 2-chain, the sum of the 2-cells has as boundary the (clockwise) oriented sum of the boundary edges). This is possible because the Van Kampen diagram is topologically a disk.

Inductively building up the edge path

We start by defining $v_0 = g_1a$, $v_1 = g_1$ and $s_1 = e_0 = [g_1a, g_1]$. g_1a is in S, so e_0 is red and on the boundary, so is incident on a unique red relator cell which we call r_1.

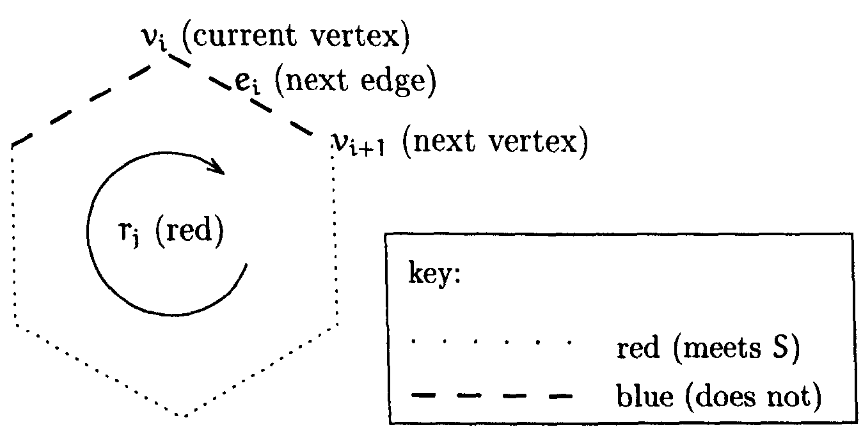

Diagram 2.1.6: Next edge

If we are given a vertex v_i, and a red relator cell r_j on which it is incident, then the orientation on r_j determines a *next edge* and a *next vertex* around r_j.

Case (1): If these are blue, we will call them e_i and v_{i+1} respectively, so that $e_i = [v_i, v_{i+1}]$. Note that e_i and v_{i+1} are also on the boundary of with r_j.

Case (2):

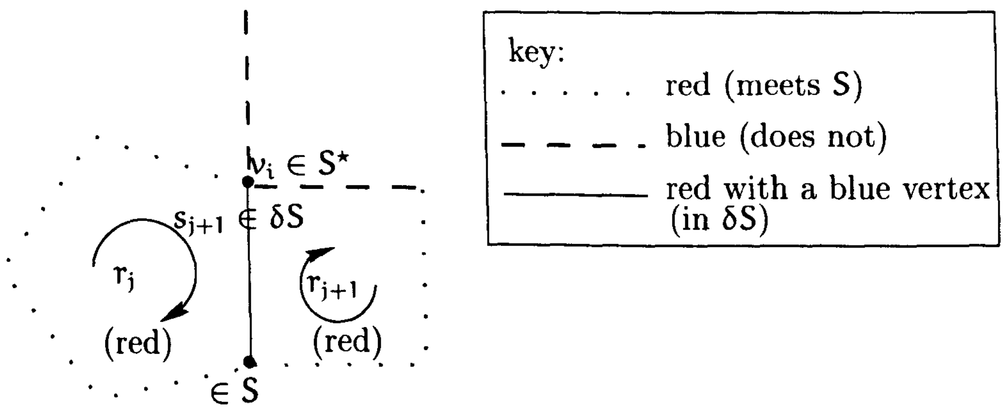

Diagram 2.1.7: The next edge is a coboundary edge.

If the next edge and next vertex are red, we define s_{j+1} to be this next edge. Note that s_{j+1} has one red vertex, and one blue vertex, i.e. it has one vertex in S, and the other in $S^\star$. Thus s_{j+1} is in δS.

Note also that both s_j and s_{j+1} are edges of the relator cell r_j, so are on the same relator.

If the next edge (s_{j+1}) is not on the boundary, then there is another relator cell $r_{j+1} \neq r_j$ incident both with it and with the next vertex. Since this cell is incident with a red vertex, it is red, so we keep the current vertex, but change relator cell to this new one.

Summary

At each stage we have a current vertex and relator cell, and may find that they determine a blue next vertex, in which case we change the current vertex, and add the intervening (blue) edge to our edge path. If not, we add the intervening (red, coboundary) edge to our other sequence of edges. In this case, if the current vertex and relator cell do not both touch the boundary, they determine another red relator cell, in which case we change the current relator cell, but leave the current vertex as it is.

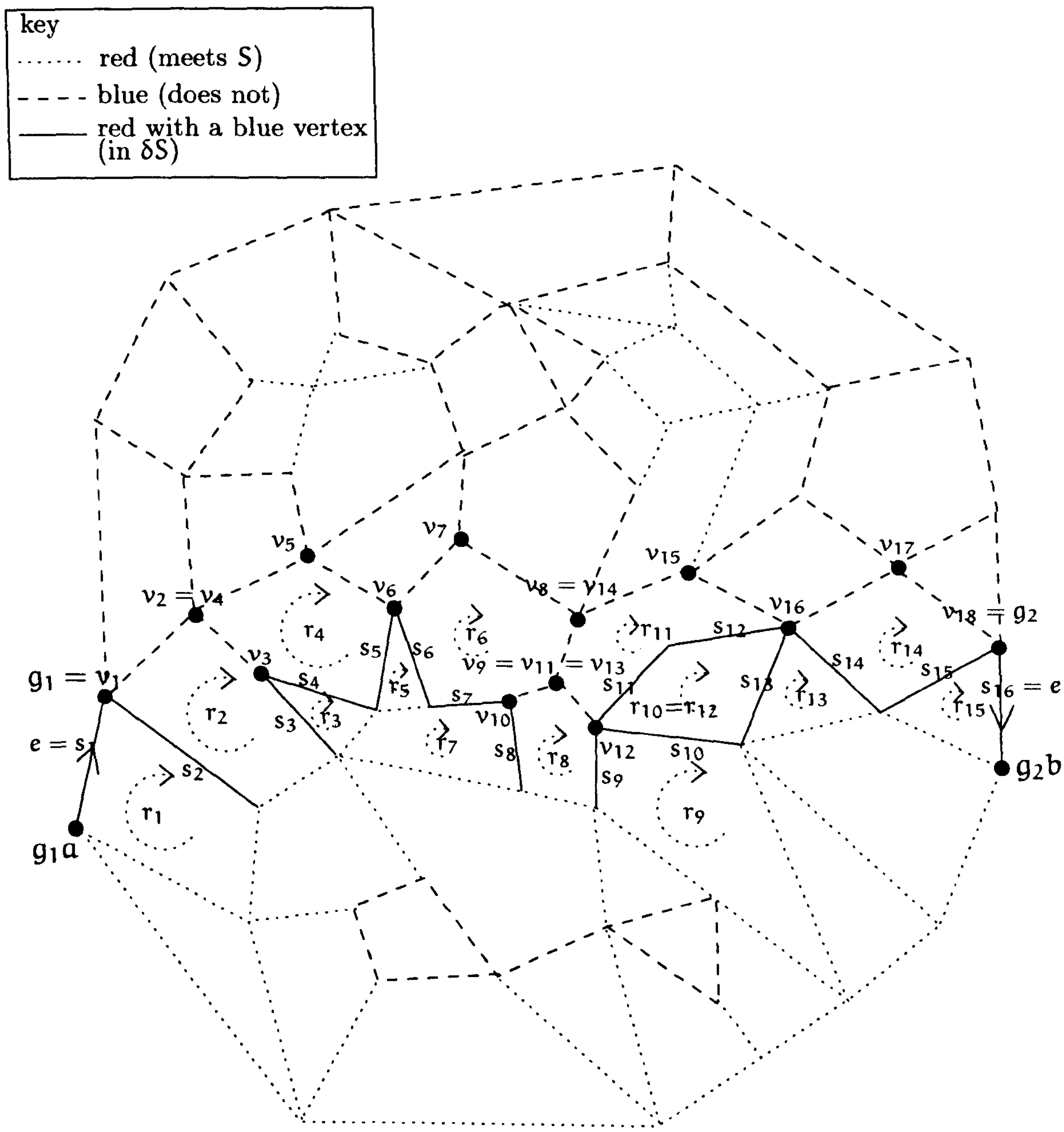

Diagram 2.1.8: Blue vertices on red relator cells, and coboundary edges.

The path can only end at g_2

The only other possibility that may arise during the induction is that the current vertex is on the boundary, with the next vertex and edge both red and on the boundary. The orientation this edge gets from the adjacent cell is from blue to red, i.e. it starts outside B_n° and ends inside it. There are only two edges on the boundary with both blue and red vertices, and we oriented the 2-cells using a red to blue orientation on one of them. Thus the next edge can only be the other one, $[g_2, g_2 b]$, and so the current edge must end at g_2.

Thus the last next edge must be the red edge $[g_2, g_2 b] = e'$, so $s_{m'+1} = e'$, where m' is the number of relator cells used in the construction of the edge path.

The path must end

We have proved that we may continue adding more edges to our blue edge path until we reach the point g_2, but what if we never reach it? In fact this cannot happen, since we can prove that the edge path transverses each edge at most once *in a given direction.*

Given $e_i = [v_i, v_{i+1}]$, there is only one edge that can be e_{i-1}:

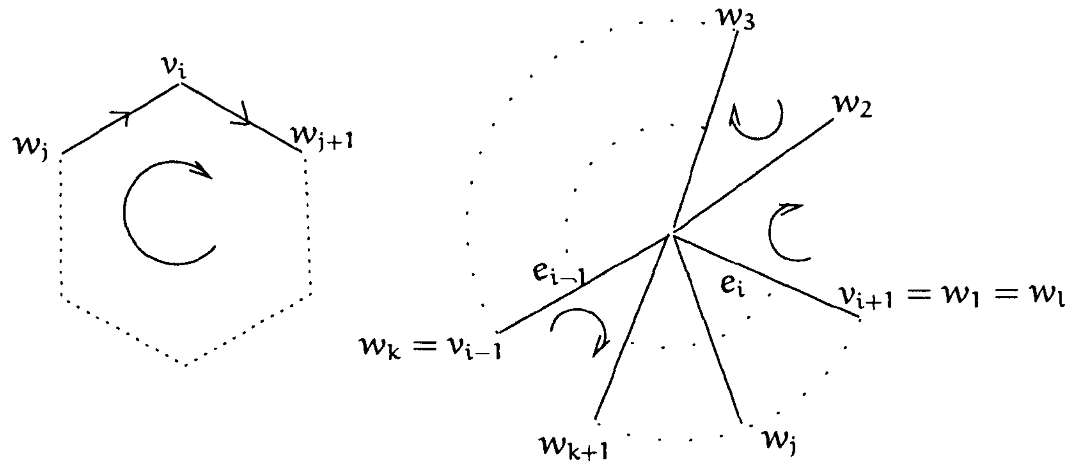

Diagram 2.1.9: Next edges around v_i

Fix v_i and consider the sequence $(w_j)_{j=1}^{l}$ of vertices which are arranged (anti-clockwise) around v_i, defined as follows: $w_1 = v_{i+1}$. Given w_j, consider the oriented edge $[w_j, v_i]$. This determines a unique relator cell having $[w_j, v_i]$ as an edge and oriented coherently with it. This in turn defines a next oriented edge, which we call $[v_i, w_{j+1}]$, thus defining w_{j+1}. We continue in this way defining a next w_j until we reach $w_l = w_1$ again.

Now for some k, $e_{i-1} = [w_k, v_i]$. Some of the edges around the vertex v_i are blue. If one of the edges $[v_i, w_j]$ for j strictly between k and l is blue, then by our definition, the next edge in the edge path after e_{i-1} cannot be $[v_i, v_{i+1}]$, but this is a contradiction. Thus there is no blue edge $[v_i, w_j]$ for j between k and l, so we have proved that $e_{i-1} = [w_k, v_i]$ where w_k is the last vertex in the sequence $(w_j)_{j=1}^{l}$ such that $[w_k, v_i]$ is blue. Thus an edge in our blue edge path determines its predecessor uniquely.

What we have shown implies that if $e_{i+1} = e_{j+1}$, then $e_i = e_j$. Note that e_1 is the first edge on the edge path, and does not have a blue predecessor, so cannot be reached a second time; by induction, the edge path does not meet the same edge with the same orientation more than once.

Conclusion

Thus since there are only finitely many edges in the Van Kampen diagram, and the edge path can meet each of them only twice, the inductive addition of edges to this edge path must terminate. As we found before, the only place this can happen is when the current edge has its terminus at g_2, so our edge path goes from g_1 to g_2. Since it is always on the boundary of a red relator cell, the vertices are always within distance $\kappa + 1$ of a vertex of S, proving part 1.[4]

We constructed a sequence $(s_j)_{j=1}^{m'+1}$ of edges of δS such that $s_1 = e$ and $s_{m'+1} = e'$. For each $j \in \{1, \ldots, m'\}$, s_j and s_{j+1} were on the same relator, so by lemma 2.1.8 on page 36, $d(s_j, s_{j+1}) \leqslant \kappa$, and we have proved part 2.

□

2.1.3 Finding components

Now we are ready to prove that the Γ-components of $\overline{B}_{n+\kappa} \setminus B^\circ_n$ and the components of $\Gamma \setminus B^\circ_n$ are in one-to-one correspondence under inclusion.

[4]Strictly speaking, we have defined the sequence of edges in the Van Kampen diagram, but these edges define a sequence of edges in the Cayley graph. The image points are at least as close to each other as the originals are in the Van Kampen diagram, so the sequence in the Cayley graph has the same property as that in the Van Kampen diagram.

Corollary 2.1.10

Let κ *be the connectivity constant of* G*, and let* g_1 *and* g_2 *be elements of* S_n*. If there is a path between them lying outside the open ball of radius* n*,* B_n°*, then there is necessarily a path between them which is* Γ*-within* $\overline{B}_{n+\kappa} \setminus B_n^\circ$*.*

Proof

Let $S = G \cap B_n^\circ = G \cap \overline{B}_{n-1}$. Then $S^\star = G \cap (\Gamma \setminus B_n^\circ)$.

Let g_1 and g_2 be two distinct elements of S_n, and suppose there is a path α_1 between them lying wholly in $\Gamma \setminus B_n^\circ$. We may construct a second path between them which lies inside B_n° except at its endpoints: take a geodesic from g_2 to id_G and from id_G to g_1 and compose them. Call the first edge of this path e', the last edge e, and the remainder of the path α_2.

α_1 does not meet S, α_2 does not meet $S^\star$, and both e and e' are in δS.

By theorem 2.1.9, there is an edge path $(e_i)_{i=1}^m$ from g_1 to g_2, which is Γ-within $G \cap (\Gamma \setminus B_n^\circ)$, such that the vertices of this edge path all lie within distance $\kappa + 1$ of a vertex of $\overline{B}_{n-1}$, so all lie in $\overline{B}_{n+\kappa} \setminus B_n^\circ$.

□

We now extend the result from elements of S_n to all elements of $\overline{B}_{n+\kappa} \setminus B_n^\circ$:

Corollary 2.1.11

Let κ *be the connectivity constant of* G*, and let* g_1 *and* g_2 *be elements of the band,* $\overline{B}_{n+\kappa} \setminus B_n^\circ$*.*

Take geodesics from each to vertices v_1 *and* v_2 *on* S_n*. Then these two paths lie entirely inside the band, and we may precompose or postcompose with them or their inverses without changing whether a given path lies in*

$\overline{B}_{n+\kappa} \setminus B^\circ_n$ *or* $\Gamma \setminus B^\circ_n$.

Thus if there is a path between g_1 *and* g_2 *lying outside* B°_n, *then there is necessarily a path between them which is* Γ*-within* $\overline{B}_{n+\kappa} \setminus B^\circ_n$, *using the previous result, corollary 2.1.10.*

□

Now we may prove the main result:

Theorem 2.1.12

Let κ *be the connectivity constant of* G. *There is a bijection, given by inclusion as a subset, between the* Γ*-components of* $\overline{B}_{n+\kappa} \setminus B^\circ_n$ *and the components of* $\Gamma \setminus B^\circ_n$:

Let C *be a component of* $\Gamma \setminus B^\circ_n$. *Since* C *is path-connected,*[5] *there is a path in* C *between any pair of elements of* $C \cap \left(\overline{B}_{n+\kappa} \setminus B^\circ_n\right)$. *By the previous result, there must also be a path which stays* Γ*-within* $C \cap \left(\overline{B}_{n+\kappa} \setminus B^\circ_n\right)$. *Thus* $C \cap \left(\overline{B}_{n+\kappa} \setminus B^\circ_n\right)$ *is a* Γ*-component of* $\overline{B}_{n+\kappa} \setminus B^\circ_n$.

Conversely, any Γ*-component of* $\overline{B}_{n+\kappa} \setminus B^\circ_n$ *lies wholly within one component of* $\Gamma \setminus B^\circ_n$.

We have proved that the Γ*-components of* $\overline{B}_{n+\kappa} \setminus B^\circ_n$ *are exactly the intersections of* $\overline{B}_{n+\kappa} \setminus B^\circ_n$ *with the components of* $\Gamma \setminus B^\circ_n$, *restricted to* G. *This inclusion is a bijection from the former to the latter.*

[5]In Cayley graphs, components and path-components are the same because Cayley graphs are locally path-connected.

2.2 A simple condition for having more than one end

Theorem 2.2.1

Suppose that there is an $n \in \mathbb{N}$ such that B_n° disconnects Γ. If two elements of length $2n$ lie in different components of $\Gamma \setminus B_n^\circ$, then Γ has more than one end.

Note The presence of an element of length $2n$ in a component of $\Gamma \setminus B_n^\circ$ does not on its own guarantee that the component is infinite. It is not inconceivable that the group might be finite of diameter at least $2n$, but yet still be disconnected by B_n°.

Proof

The proof spans the rest of this section. Throughout, fix n such that B_n° disconnects Γ, and let g_{-1} and g_1 be elements of length $2n$ lying in separate components of $\Gamma \setminus B_n^\circ$. We construct a path that is disconnected by B_n° into two infinite components.

2.2.1 Patches

Definition 2.2.2 *Let $g \in G$, and suppose B_n° disconnects Γ. Since G acts by isometries on Γ, $B_n^\circ(g)$ disconnects the Cayley graph into components $C_1, \ldots C_m$, say. The sphere of radius n about g, $S_n(g)$, is partitioned into subsets $S_n(g) \cap C_i$, which we call* patches *about g at distance n. (When the value of n is fixed or clear from the context, we do not mention it.)*

Fix i and let $c \in C_i$. Suppose a path from c to g meets $S_n(g)$ first at *some point of* C_j. *Then this initial segment of the path demonstrates that* C_i *and* C_j *are the same component of* $\Gamma \setminus B_n^\circ$. *Thus we define the* c*-patch about* g, $c \dashv g$, *to be the patch about* g *which any path from* c *to* g *must meet first, i.e.* $c \dashv g = C_i \cap S_n(g)$.

Lemma 2.2.3

Let $d \in \Gamma \setminus B_n^\circ(g)$. *The following are equivalent:*

(same patches) $d \dashv g = c \dashv g$

(all meet first) *All paths from* d *to* g *meet* $c \dashv g$ *first amongst the patches about* g.

(one meets first) *Some path from* d *to* g *meets* $c \dashv g$ *first amongst the patches about* g.

(all meet) *All paths from* d *to* g *meet* $c \dashv g$.

(one meets only) *Some path from* d *to* g *meets only the* c*-patch about* g.

(geods meet only) *All geodesics from* d *to* g *meet only the* c*-patch about* g.

(a geod meets) *A geodesic from* d *to* g *meets the* c*-patch about* g.

(path between) *There is a path from* d *to* c *that does not meet* $B_n^\circ(g)$.

(same component) $d \in C_i$

Proof

Note that if α is a path from b to g then since $\alpha^{-1}(S_n(g))$ is a closed subset of $[0, 1]$, it contains its infimum, t, say, so $\alpha|_{[0,t]}$ is a path from b to $b \dashv g$ which lies wholly in C_i, and so does not meet $B_n^\circ(g)$. We call this subpath the initial segment of α.

Note also that if α is a geodesic path, we can reverse it so that it becomes parameterised by distance from g, so it meets exactly one point of $S_n(g)$, and so meets exactly one patch about g.

(same patches) and (all meet first) are equivalent by definition.

(all meet first) $\Rightarrow$ (geods meet only) because any geodesic from d to g meets only one patch about g.

(geods meet only) $\Rightarrow$ (one meets first) because there is a geodesic from d to g.

(one meets first) $\Rightarrow$ (one meets only); take the initial segment of the path and compose it with a geodesic from its endpoint to g.

(one meets only) $\Rightarrow$ (path between); compose the initial segment of the path with a path outside $B^{\circ}_n(g)$ from its endpoint to c. Such a path exists because C_i is path connected.

(path between) $\Rightarrow$ (same component) trivially.

(same component) $\Rightarrow$ (all meet) because the initial segment of any path from d to g lies wholly in C_i and ends in $S_n(g) \cap C_i = c \mathbin{-\!\!\{} g$.

(all meet) $\Rightarrow$ (a geod meets) trivially.

(a geod meets) $\Rightarrow$ (same patches); the geodesic meets only one patch about g, so it meets $c \mathbin{-\!\!\{} g$ first, and so by the definition, $d \mathbin{-\!\!\{} g = c \mathbin{-\!\!\{} g$

□

Sometimes we write $g \mathbin{\}\!\!-} c$ instead of $c \mathbin{-\!\!\{} g$.

2.2.2 Constructing a bi-infinite path

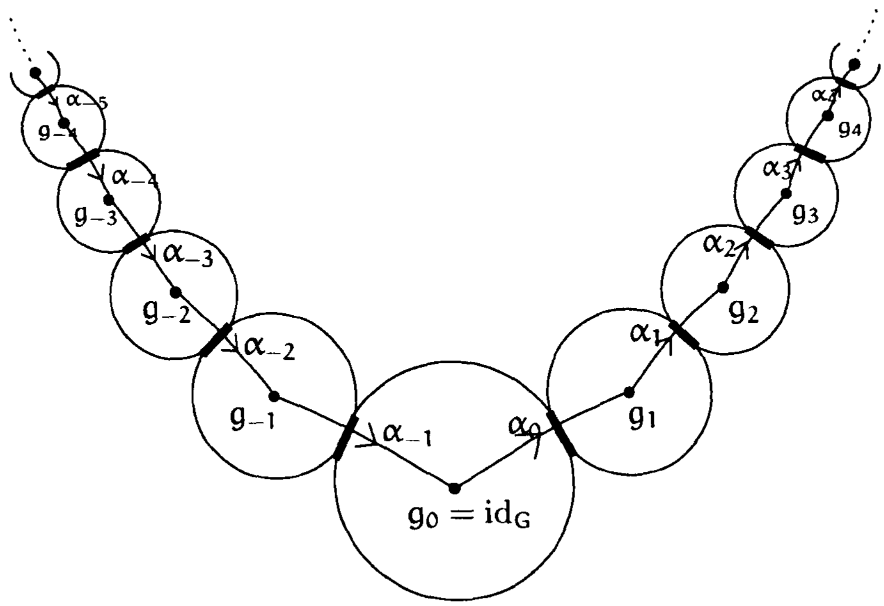

Diagram 2.2.1: Constructing a bi-infinite path

We define a bi-infinite path by inductively defining for all $i \in \mathbb{Z}$ points g_i and geodesics $\alpha_i : g_i \simeq g_{i+1}$ of length $2n$. We will prove that for $j > i$, $\overline{B}_n(g_{j+1})$ does not meet $\overline{B}_n(g_i)$, and from this, that the g_i are distinct.

The path α obtained by composing the $(\alpha_i)_{i \in \mathbb{Z}}$ is bi-infinite in the sense that for all $r \in \mathbb{R}$, both $\alpha((\infty, r])$ and $\alpha([r, \infty))$ contain infinitely many vertices of Γ.

Let $g_0 = id_G$. By hypothesis, there are points g_{-1} and g_1, with geodesics $\alpha_{-1} : g_{-1} \simeq g_0$ and $\alpha_0 : g_0 \simeq g_1$ both of length $2n$, with $g_{-1} \dashv (g_0 \neq g_0) \vdash g_1$.

Assume inductively that we have defined $g_{-i}, g_{-(i-1)}, \dots, g_{i-1}, g_i$ and paths $\alpha_{-i}, \alpha_{-(i-1)}, \dots, \alpha_{i-2}, \alpha_{i-1}$, with each α_i a geodesic of length $2n$ from g_i to g_{i+1}, and for all $j \in \{-(i-1), \dots, i-1\}$, $g_{j-1} \dashv (g_j \neq g_j) \vdash g_{j+1}$.

G acts by isometries on Γ, so for $h \in G$ there is more than one component of $\Gamma \setminus B^{\circ}_{n}(h)$ containing elements of distance $2n$ from h. Thus given g and h with $d(g,h) = 2n$ we can find k with $d(h,k) = 2n$ and $g\dashv(h \neq h)\vdash k$.

We make this construction for $g = g_{i-1}$ and $h = g_i$, defining $g_{i+1} = k$ and take α_i to be a geodesic between g_i and g_{i+1}. We repeat for $g = g_{-(i-1)}$ and $h = g_{-i}$, defining $g_{-(i+1)} = k$ and take $\alpha_{-(i+1)}$ to be a geodesic between $g_{-(i+1)}$ and g_{-i}. This completes the induction.

2.2.3 Only adjacent closed balls meet

Lemma 2.2.4

If $i \leqslant j$ *and* $\overline{B}_n(g_i)$ *meets* $\overline{B}_n(g_{j+1})$, *then* $i = j$.

Note that for $\varepsilon \in \{-1, 1\}$ and for $m \in \mathbb{Z}$, $\overline{B}_n(g_m) \cap B^{\circ}_{n}(g_{m+\varepsilon}) = \emptyset$ by the triangle inequality and the fact that g_m and $g_{m+\varepsilon}$ are distance $2n$ apart.

Proof

Suppose there are i and j in $\mathbb{Z}$ with $i < j$ and that $\overline{B}_n(g_i) \cap \overline{B}_n(g_{j+1}) \neq \emptyset$ We derive a contradiction.

Pick i and j so that $|j - i|$ is the smallest possible.

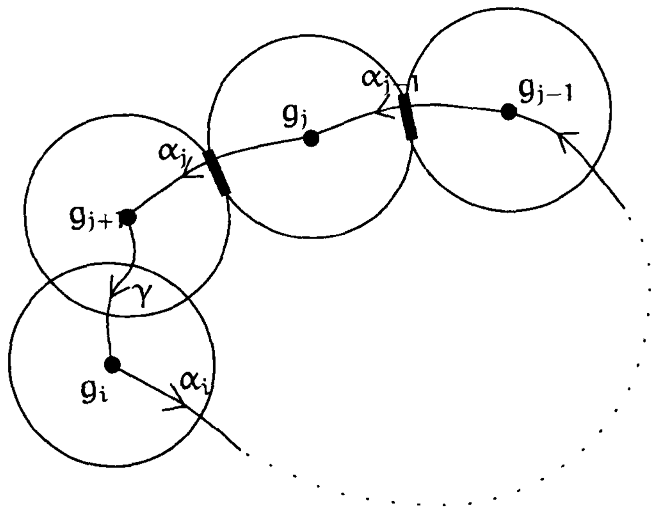

Diagram 2.2.2: Non-adjacent balls cannot meet like this.

Since $|j - i|$ is smallest, none of $\overline{B}_n(g_i), \overline{B}_n(g_{i+1}), \ldots, \overline{B}_n(g_{j-2})$ meet $\overline{B}_n(g_j)$, and thus none of them meet $B^\circ_n(g_j)$. Also, $\overline{B}_n(g_{j-1})$ does not meet $B^\circ_n(g_j)$ because g_{j-1} is adjacent to g_j. $(\overline{B}_n(g_i) \cup \overline{B}_n(g_{i+1}) \cup \cdots \cup \overline{B}_n(g_{j-1})) \cap B^\circ_n(g_j) = \emptyset$. Because of this, the composite path $\alpha_i\alpha_{i+1}\cdots\alpha_{j-2}$ from g_i to g_{j-1} lies wholly outside $B^\circ_n(g_j)$. (In the case that $j = i + 1$, this path is the trivial path.) Since α_{j-1} is a geodesic from g_{j-1} to g_j, the path $\alpha_i\alpha_{i+1}\cdots\alpha_{j-1}$ meets $S_n(g_j)$ only at $g_{j-1}\dashv(g_j$. Thus $g_i\dashv(g_j = g_{j-1}\dashv(g_j$.

Now since $\overline{B}_n(g_i)$ meets $\overline{B}_n(g_{j+1})$ there is a path $\gamma : g_{j+1} \simeq g_i$ lying wholly inside $\overline{B}_n(g_i) \cup \overline{B}_n(g_{j+1})$. Such a path cannot meet $B^\circ_n(g_j)$; $\overline{B}_n(g_{j+1}) \cap B^\circ_n(g_j) = \emptyset$ because g_j and g_{j+1} are adjacent, and $\overline{B}_n(g_i) \cap B^\circ_n(g_j) = \emptyset$ because we assumed that i and j were as close as possible. Thus by lemma 2.2.3, $g_i\dashv(g_j = g_{j+1}\dashv(g_j$. But now we have proved that $g_{j-1}\dashv(g_j = g_j)\vdash g_{j+1}$, which, by construction, does not happen.

□

2.2.4 Conclusion

We have constructed a bi-infinite path $\alpha = \cdots \alpha_{-2}\alpha_{-1}\alpha_0\alpha_1\alpha_2\cdots$. This path is bi-infinite both in the sense that its domain is $\mathbb{R}$ and that its image meets infinitely many vertices of Γ. For suppose $i \neq j$ and $g_i = g_j$. Then $\overline{B}_n(g_i)$ meets $\overline{B}_n(g_j)$, so by lemma 2.2.4, i and j are adjacent integers. But this cannot happen because in this case g_i and g_j are distance $2n$ apart.

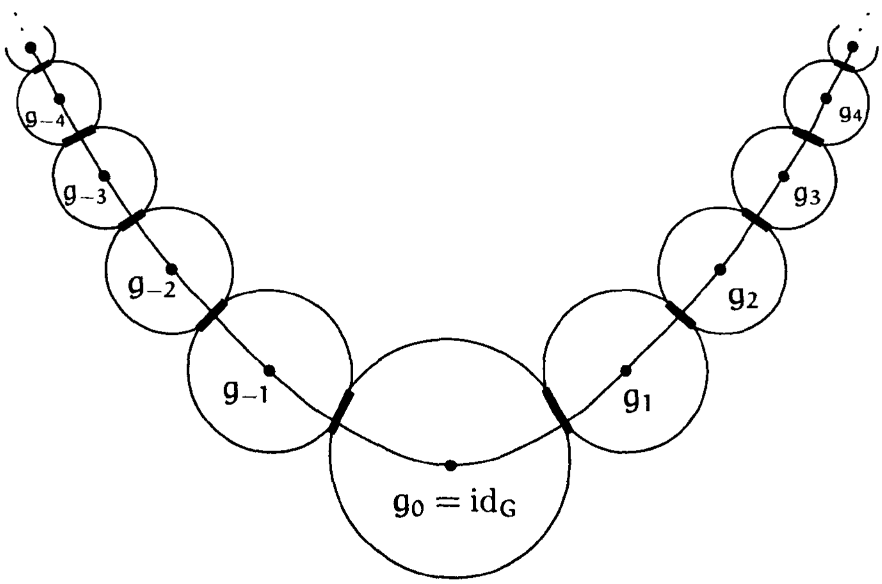

Diagram 2.2.3: α is bi-infinite.

The path $\alpha_1\alpha_2\alpha_3\cdots$ lies wholly outside $B^\circ_n(id_G)$, so the component of $\Gamma \setminus B^\circ_n$ in which g_1 lies is infinite. Similarly, the path $\cdots\alpha_{-4}\alpha_{-3}\alpha_{-2}$ demonstrates that the component of $\Gamma \setminus B^\circ_n$ in which g_{-1} lies is infinite. $B^\circ_n(id_G)$ separates the Cayley graph into more than one infinite component, and G has more than one end.

□

2.2.5 Alternative proof

The previous proof uses the G-action to construct a bi-infinite string of balls to demonstrate that G has more than one end. In this section we present an alternative proof that shows that two components of the complement of the ball can be moved strictly inside themselves using the G-action (which is of course by bijections) so are infinite.

Recall the statement of the theorem (theorem 2.2.1 on page 52):

Theorem

Suppose that there is an $n \in \mathbb{N}$ such that B°_n disconnects Γ. If two elements of length $2n$ lie in different components of $\Gamma \setminus B^\circ_n$, then Γ has more than one end.

The main work is done in the following lemma:

Lemma 2.2.5

Suppose that there is an $n \in \mathbb{N}$ such that B°_n disconnects Γ, and that S is one of the components of $\Gamma \setminus B^\circ_n$. If there exist elements x and y of G such that $B^\circ_n(x) \cap B^\circ_n(y) = \emptyset$, with $y \in xS$ and $x \notin yS$, then S contains infinitely many vertices.

Proof

Assume $B^\circ_n(x) \cap B^\circ_n(y) = \emptyset$, with $y \in xS$ and $x \notin yS$.

$B^\circ_n(x)$ is connected and does not meet $B^\circ_n(y)$, so lies entirely in one component of $\Gamma \setminus B^\circ_n(y)$. If that component were yS then x would be an element of yS, which is false by assumption, so $B^\circ_n(x) \cap yS = \emptyset$. Thus yS lies entirely within one component, C, of $\Gamma \setminus B^\circ_n(x)$.

We know that neither yS nor $B_n^\circ(y)$ meet $B_n^\circ(x)$, so neither does their union. However, $B_n^\circ(y) \cup yS$ is connected, so it lies wholy within one component of $\Gamma \setminus B_n^\circ(x)$, and since $yS \subset C$, we have that $y \in B_n^\circ(y) \cup yS \subset C$.

But $y \in xS$, so $xS = C$ and $yS \subset xS$. Now $y \in xS$ by assumption, but $y \notin yS$, so $yS \subsetneq xS$.

Thus the set of vertices of yS is a *proper* subset of the set of vertices of xS. However, the G-action provides a bijection between these sets and the set of vertices of S; there is a bijection between the set of vertices of S and a proper subset of them, so there are infinitely many vertices in S.

□

Now we give the alternative proof of the theorem.

Proof (of theorem 2.2.1)

Fix n such that B_n° disconnects Γ, and let g and h be elements of length $2n$ lying in separate components of $\Gamma \setminus B_n^\circ$. We prove that G has more than one end, by showing that the components of $\Gamma \setminus B_n^\circ$ containing g and h contain infinitely many vertices.

Note that it is enough to prove that the component of $\Gamma \setminus B_n^\circ$ containing g is infinite, since the fact that the component containing h is infinite follows by symmetry. Let S be the component of $\Gamma \setminus B_n^\circ$ containing g.

Note first that $B_n^\circ(g)$, $B_n^\circ(id_G)$ and $B_n^\circ(h)$ are pairwise disjoint: Neither $B_n^\circ(g)$ nor $B_n^\circ(h)$ meets B_n° by the triangle inequality. They do not meet each other because they are both path connected, so a non-empty intersection would imply that g and h lie in the same component of $\Gamma \setminus B_n^\circ$, which is false by assumption.

Thus, by the previous lemma, it is enough to find x and y amongst $\{g, id_G, h\}$ such that $y \in xS$ but $x \notin yS$.

Case (1): $id_G \in hS$. We know that $h \notin S = id_G S$, so if $id_G \in hS$ then let $x = h$ and $y = id_G$; the result follows from lemma 2.2.5 on page 59.

Case (2): $id_G \notin gS$. We know that $g \in S = id_G S$, so if $id_G \notin gS$ then let $x = id_G$ and $y = g$; the result follows from lemma 2.2.5.

Case (3): $id_G \notin hS$ and $id_G \in gS$. There is a geodesic path from id_G to g; it lies in $\overline{B}_n \cup B_n^\circ(g)$ by the triangle inequality, so does not meet $B_n^\circ(h)$. Thus id_G and g are in the same component of $\Gamma \setminus B_n^\circ(h)$; $id_G \notin hS \Rightarrow g \notin hS$.

Similarly, id_G and h are in the same component of $\Gamma \setminus B_n^\circ(g)$; $id_G \in gS \Rightarrow h \in gS$.

Thus $g \notin hS$ and $h \in gS$, so by lemma 2.2.5, S contains infinitely many vertices.

Thus in any case, S is infinite. $\Gamma \setminus B_n^\circ$ has at least two infinite components, and G has more than one end.

□

Note that this alternative proof, like the first, does not work when only one component of $\Gamma \setminus B_n^\circ$ has an element of length $2n$ in it.

2.3 Infinite components and infinite words

We prove a lemma relating whether a component of $\Gamma \setminus B_n^\circ$ is infinite to the existence of certain types of rays in it.

Lemma 2.3.1

Let C *be a component of* $\Gamma \setminus B_n^\circ$. *Then the following are equivalent:*

1. C *is infinite*
2. *there is a geodesic ray* r, *with* $r(m) \in C$ *for all* $m \geqslant n$
3. *there is an infinite injective path* γ *and an* $M \in \mathbb{N}$ *such that* $\gamma(m) \in C$ *for all* $m \geqslant M$.

Proof

Clearly, since geodesics are injective, (2) implies (3). Also, if γ is an infinite injective path then the points $\gamma(m) \in C$ for $m \geqslant M$ are all distinct elements of C, so C is infinite; (3) implies (1). All that remains is to prove that (1) implies (2).

We assume C is an infinite component of $\Gamma \setminus B_n^\circ$, and prove that there is a geodesic ray into it.

Since there are elements arbitrarily far from id_G in C (because there are only finitely many elements of each length), there is no bound on the length of geodesics into C from id_G. Thus, since there are finitely many entry points into C, at least one of the geodesics α to one of these entry points, a, must possess extensions of arbitrarily large lengths.

Consider the union of all the images of all the geodesic extensions of α. This forms an infinite connected graph based at a. (Think of the edges as directed away from a.)

Proceed along this directed graph, constructing a path. At each point we have finitely many choices (Γ is a locally finite graph). At least one of the choices must allow us to reach points arbitrarily far from a, because otherwise there would be a global bound on the length of geodesics through a. In this way we inductively build up an infinite geodesic extension of α.

By construction this ray α is a geodesic ray into C.

□

Chapter 3

Splittings and coboundaries

Stallings' theorem [Sta68] relates the number of ends of a group to whether it splits over a finite subgroup. Dicks and Dunwoody give a proof of Stallings' theorem in their book *Groups acting on Graphs*, [DD89]. The proof involves 'coboundaries', and Dunwoody more recently found a shorter proof using the 'Bergman norm' on coboundaries.

Dicks and Dunwoody use Bass-Serre theory to show that the group splits over a finite subgroup; they prove that the group acts in a certain way on an infinite tree, called the Bass-Serre tree for the splitting. For an introduction to Bass-Serre theory see [SW] or [Ser83].

In this chapter we define and study the coboundaries Dicks and Dunwoody use. Later, in section 4.4, we use some of these results in an algorithm which takes a group with certain computability properties, and if it has more than one end, calculates a finite subgroup over which it splits.

3.1 Stallings' theorem and coboundaries

We give a brief exposition of the proof of Stallings' theorem in order to introduce concepts we will need later.

Let G be a group, and let S be any subset of G. Recall from section 2.1.1.1 on page 34 that $S^\star$ denotes the complement of S, $G \setminus S$, and that the coboundary, δS, of S, is the set of edges of Γ which meet both S and $S^\star$.

3.1.1 Nestedness

The following tables are entrywise equivalent:

$$\begin{array}{c|c} S \subset T^\star & S \subset T \\ \hline S^\star \subset T^\star & S^\star \subset T \end{array} \Leftrightarrow \begin{array}{c|c} S \cap T = \emptyset & S \cap T^\star = \emptyset \\ \hline S^\star \cap T = \emptyset & S^\star \cap T^\star = \emptyset \end{array} \Leftrightarrow \begin{array}{c|c} T \subset S^\star & T^\star \subset S^\star \\ \hline T \subset S & T^\star \subset S \end{array}$$

This accounts for the use of the word nested below.

Definition 3.1.1 (nested) *Let* S *and* T *be sets of elements of* G*. We say that the pair* (S, T) *is a* nested *pair if one of the sets* $S \cap T$, $S \cap T^\star$, $S^\star \cap T$, *or* $S^\star \cap T^\star$, *is empty.*

A set of subsets of G *is said to be* nested *if it is pairwise nested.*

A subset S *of* G *is said to be* G-nested *if the set of all* G*-translates of it is a nested set of subsets of* G*, i.e. if* $\{gS \mid g \in G\}$ *is a nested set of subsets of* G.

A coboundary, δS*, is said to be* G-nested *if* S *is* G*-nested.*

Note that nestedness of the pair (S, T) is independent of the order and independent of whether we take S or $S^\star$, T or $T^\star$. As a corollary, $\{gT \mid T \in \{S, S^\star\},\ g \in G\}$ is a nested set of subsets of G if and only if $\{gS \mid g \in G\}$ is.

So S is G-nested if and only if $\{gT \mid T \in \{S, S^\star\},\ g \in G\}$ is a nested set of subsets of G.

3.1.2 The Bergman norm

Dicks and Dunwoody prove (in [DD89]) that if there is an infinite G-nested subset, S, of G, with infinite complement $S^\star$, and finite coboundary δS, then there is a tree on which G acts with edges $\{gT \mid T \in \{S, S^\star\},\ g \in G\}$, so that G splits over stab(S).

They prove also that if there exists any infinite subset T of G, with infinite complement $T^\star$, and finite coboundary, then there exists a G-nested set S with the same properties. In this way they prove Stallings' theorem, i.e. that a group with more than one end splits over a finite subgroup.

As mentioned earlier, Dunwoody gave a much shorter alternative proof of this latter point, using the 'Bergman norm' on coboundaries. The norm is due to George Bergman [Ber68].

Definition 3.1.2 (Bergman norm) *Let $E = \delta S$ be a finite coboundary. We define $\mu_i(E)$ to be the number of edge paths of length i which start in S and end in $S^\star$.*

We define the Bergman norm *of E to be the sequence $\mu(E) = (\mu_i(E))_{i \in \mathbb{N}}$. If S is a subset of G with coboundary δS, we define $\mu(S) = \mu(\delta S)$.*

We say $\mu(a) < \mu(b)$ if there exists $I \in \mathbb{N}$ such that for all $i < I$, $\mu_i(a) = \mu_i(b)$, and $\mu_I(a) < \mu_I(b)$, i.e. we use the lexicographic ordering with the Bergman norm.

Notice that $\mu(\delta S) = \mu(\delta(S^\star))$, so the asymmetry in the definition is only apparent. We could equally well have defined $\mu_i(E)$ to be the number of unoriented edge paths which cross edges of E an odd number of times.

3.1.3 Proof of Stallings' theorem

We concentrate only on those parts of the proof we will be using later; for the more interesting G-tree constructions, the reader is referred to [DD89].

Lemma 3.1.3 (Dunwoody)

Let S and T be subsets of G with finite coboundaries. Suppose that $\mu(S) = \mu(T)$ and that none of the sets $S \cap T$, $S \cap T^\star$, $S^\star \cap T$, nor $S^\star \cap T^\star$ is infinite with infinite complement and strictly smaller Bergman norm. Then the pair (S, T) is a nested pair. [Dun98]

⊠

Corollary 3.1.4 (Coboundary calculus)

Let S and T be subsets of G with finite coboundaries. If $\mu(S) = \mu(T)$ and (S, T) is not a nested pair, then one of the sets $S \cap T$, $S \cap T^\star$, $S^\star \cap T$, or $S^\star \cap T^\star$ has strictly smaller Bergman norm, and is infinite with infinite complement.

□

Theorem 3.1.5 (Dunwoody)

If S is a Bergman-minimal subset of G, then it is G-nested, i.e. the set $\{gT \mid T \in \{S, S^\star\},\ g \in G\}$ is a nested set of subsets of G.

Proof

$\delta S = \delta(S^\star)$, so $\mu(S) = \mu(\delta S) = \mu(\delta(S^\star)) = \mu(S^\star)$. Also, for any $g \in G$, gS has the same Bergman norm as S (because G acts by graph isometries on Γ).

Thus if S is Bergman-minimal, so are all of the sets $\{gT \mid T \in \{S, S^*\},\ g \in G\}$. By lemma 3.1.3, they must therefore be pairwise nested, i.e. S is G-nested.

□

The following theorem of Bergman in particular guarantees that a sequence of edge sets with decreasing Bergman norm is eventually constant.

Theorem 3.1.6 (Bergman)

The set of sets of edges of a graph is well-ordered by lexicographic ordering with the norm $\mu(\ )$. [Ber68]

⊠

Theorem 3.1.7 (Stallings)

If G is a finitely generated group with more than one end, then it splits over a finite subgroup.

The proof is Dunwoody's alteration of Stallings' original result, using Bergman's norm.

Proof

If G has more than one end, there exists a subset T of G which is infinite, has infinite complement and finite coboundary. (See [DD89].)

We define a sequence $(T_i)_{i \in \mathbb{N}}$ of infinite subsets of G with infinite complement. Let $T_1 = T$.

If T_i is G-nested, define $T_{i+1} = T_i$.

If T_i is not G-nested, there exists $g \in G$ with the pair (T_i, gT_i) non-nested. By corollary 3.1.3, one of $T_i \cap gT_i$, $T_i \cap gT_i^*$, $T_i^* \cap gT_i$, or $T_i^* \cap gT_i^*$ is infinite with infinite complement and has strictly smaller Bergman norm. Define T_{i+1} to be this set with smaller norm.

By theorem 3.1.6, this sequence has a minimal element with respect to the norm, so eventually must attain this minimum. This means that at some stage, $\mu(T_{i+1}) = \mu(T_i)$, but then T_i was G-nested. Define $S = T_i$.

Thus there exists an infinite G-nested set, S, with infinite complement. By theorem II.1.8 and I.4.1 of [DD89], G splits over stab(S). Now stab(S) = stab(S^*), so each element of stab(S) also stabilises δS, i.e. stab(S) $\subset$ stab(δS). Thus, since δS is finite, stab(S) is finite, so G splits over a finite subgroup.

□

3.2 Oriented coboundaries

In the proof of Stallings' theorem, we explained that G splits over a finite subgroup stab(S), for some infinite subset S of G which has a finite coboundary and infinite complement. If we wish to calculate this subgroup, it is impractical to find the stabiliser of an infinite set. We proved that stab(S) is finite by showing that stab(S) $\subset$ stab(δS). This inclusion may be proper, as is the case with the group $G = \langle a, b \mid a^2, b^2 \rangle \cong \mathbb{Z}_2 * \mathbb{Z}_2$, and the subset S consisting of elements whose shortest word in the generators starts with a.

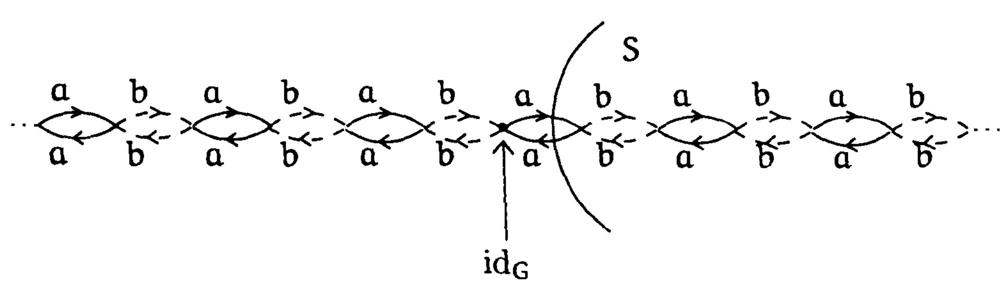

Diagram 3.2.1: $\mathbb{Z}_2 * \mathbb{Z}_2$

In this section, we define the *oriented* coboundary of a subset of G, and prove that the stabiliser of the oriented coboundary is equal to the stabiliser of the set itself. Since there are finitely many edges in the oriented coboundary, we can calculate its stabiliser.

Definition 3.2.1 (oriented coboundary) *Let* S *be a subset of* G *with coboundary* δS. *We define the* oriented coboundary *of* S, $\overrightarrow{\delta S}$, *to be the set* δS *together with the orientation on the edges in* δS *pointing towards* S.

Thus the coboundary δS *has two orientations,* $\overrightarrow{\delta S}$, *and* $\overrightarrow{\delta S^*}$.

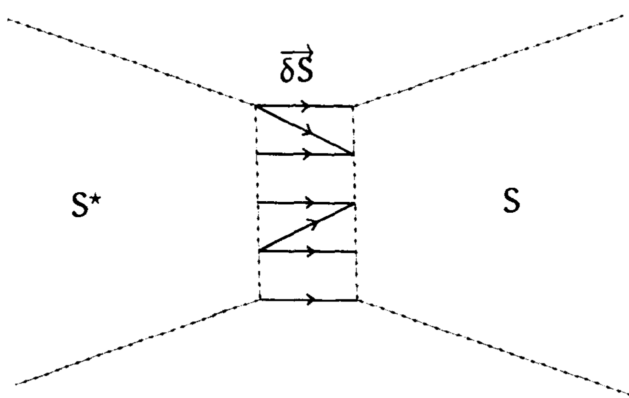

Diagram 3.2.2: An oriented coboundary

As with coboundaries, we are only ever interested in *finite* oriented coboundaries.

Lemma 3.2.2

Let $g \in G$, *and let* $\overrightarrow{\delta S}$ *be an oriented coboundary. Then* $g(\overrightarrow{\delta S}) = \overrightarrow{\delta(gS)}$.

Proof

Note that $g(S^*) = (gS)^*$.

Clearly, $g(\overrightarrow{\delta S}) \subset \overrightarrow{\delta(gS)}$: If $[x, y] \in \overrightarrow{\delta S}$, then $x \in S^*$ and $y \in S$, so

$gx \in g(S^\star) = (gS)^\star$ and $gy \in gS$, and so $g[x,y] = [gx, gy] \in \overrightarrow{\delta(gS)}$.

Conversely, if $[x,y] \in \overrightarrow{\delta(gS)}$, then $x \in gS^\star = g(S^\star)$ and $y \in gS$, so $g^{-1}x \in S^\star$ and $g^{-1}y \in S$. Thus $g^{-1}[x,y] = [g^{-1}x, g^{-1}y] \in \overrightarrow{\delta S}$, i.e. $[x,y] \in g(\overrightarrow{\delta S})$.

□

Corollary 3.2.3

Let $g \in G$, and let δS be a coboundary. Then $g(\delta S) = \delta(gS)$.

□

3.2.1 Stabilisers of oriented coboundaries

We make the following definition for clarification.

Definition 3.2.4 *Let S be a subset of G. We say that an element of G stabilises the oriented coboundary of S if it stabilises the coboundary of S and preserves the orientation on the edges.*

Lemma 3.2.5

Let S be a subset of G with finite coboundary δS and oriented coboundary $\overrightarrow{\delta S}$. Consider the action of G on its Cayley graph. Then $\mathrm{stab}(S) = \mathrm{stab}(\overrightarrow{\delta S})$.

Proof

First we prove the easier inclusion $\mathrm{stab}(S) \subset \mathrm{stab}(\overrightarrow{\delta S})$. Let $g \in \mathrm{stab}(S)$, so $g \in \mathrm{stab}(S^\star)$. Let e be an oriented edge in $\overrightarrow{\delta S}$. Then the initial vertex of e is in $S^\star$ and the terminal vertex is in S. Since $g \in \mathrm{stab}(S)$, g sends the terminal vertex of e to a vertex of S, and since $g \in \mathrm{stab}(S^\star)$, g sends the initial vertex of e to a vertex of $S^\star$. Thus e is sent to an edge which also starts in $S^\star$ and terminates in S, i.e. $ge \in \overrightarrow{\delta S}$. Thus every element of $\mathrm{stab}(S)$ stabilises $\overrightarrow{\delta S}$, i.e. $\text{stab}(S) \subset \text{stab}(\overrightarrow{\delta S})$.

Now we prove the reverse inclusion. Let $g \in \text{stab}(\overrightarrow{\delta S})$. We must prove that g stabilises S. Let $a \in S$.

Case (1): If a is incident with an edge e of $\overrightarrow{\delta S}$, it must be a terminal vertex (the oriented coboundary is defined to be oriented towards S). Since $g \in \text{stab}(\overrightarrow{\delta S})$, this edge is sent by g to another edge in $\overrightarrow{\delta S}$, so its terminal vertex is sent into S, and so $ga \in S$ as required.

Case (2): Assume now that a is not incident with any edge of $\overrightarrow{\delta S}$, and assume for contradiction that $ga \notin S$.

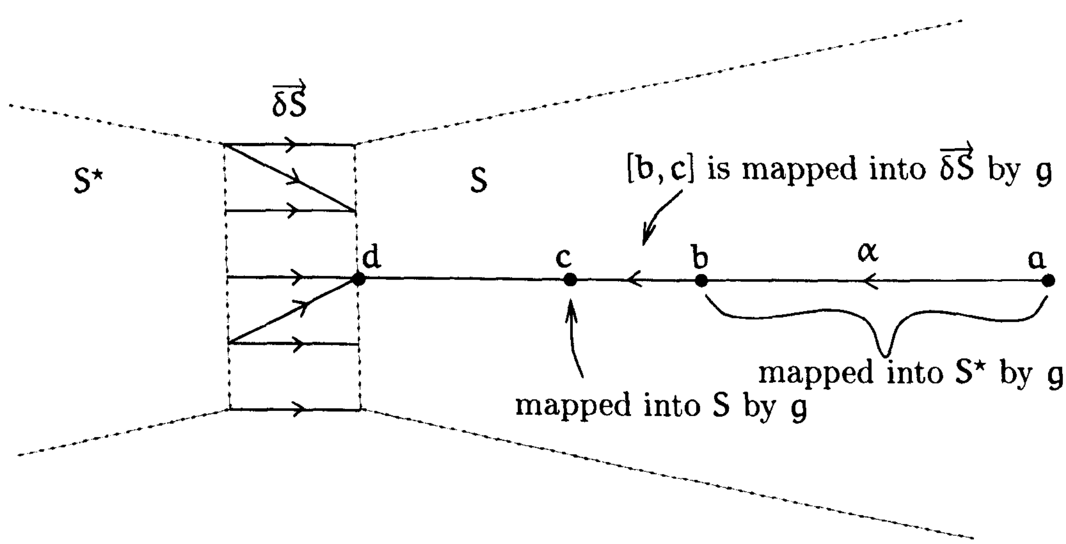

Diagram 3.2.3: If g stabilises the oriented coboundary of S, it must stabilise S.

Let $\alpha : a \simeq d$ be a geodesic from a to the set of vertices in S which are incident with edges of δS. α lies wholly in S because otherwise it would not be shortest. $gd \in S$ by the previous case. As we travel along the path α, there is a first vertex, c, for which $gc \in S$. The previous vertex on α, b, has $gb \notin S$.

Let e be the oriented edge $[b, c]$. We have $b, c \in S$ but $gb \in S^*$ and $gc \in S$, so $e \notin \overrightarrow{\delta S}$ but $ge \in \overrightarrow{\delta S}$. In other words, $(ge) \in \overrightarrow{\delta S}$ but $g^{-1}(ge) \notin \overrightarrow{\delta S}$, so there is an element $g^{-1} \in \mathrm{stab}(\overrightarrow{\delta S})$ which does not stabilise $\overrightarrow{\delta S}$. This is a contradiction, so our initial assumption that $ga \notin S$ was false, and $ga \in S$.

Thus in either case, $a \in S \Rightarrow ga \in S$, so g stabilises S. Thus we have proved $\mathrm{stab}(\overrightarrow{\delta S}) \subset \mathrm{stab}(S)$.

□

Corollary 3.2.6
$\mathrm{stab}(\overrightarrow{\delta S}) = \mathrm{stab}(S) = \mathrm{stab}(S^*) = \mathrm{stab}(\overrightarrow{\delta S^*})$.

□

3.2.2 Orbits in oriented coboundaries

An oriented coboundary, $\overrightarrow{\delta S}$, is partitioned into orbits of edges under the action of $\mathrm{stab}(\overrightarrow{\delta S}) = \mathrm{stab}(S)$. We make the following abbreviation:

Definition 3.2.7 (orbit) *Let* G *be a finitely presented group, and let* δS *be the coboundary of some set* S *of vertices of* G*. By an* orbit *in* δS *we mean an orbit under the action of* $\mathrm{stab}(\overrightarrow{\delta S}) = \mathrm{stab}(S)$.

Example 3.2.8

Consider again the group $G = \langle a, b \mid a^2, b^2 \rangle \cong \mathbb{Z}_2 * \mathbb{Z}_2$, and the subset S consisting of elements whose shortest word in the generators starts with a.

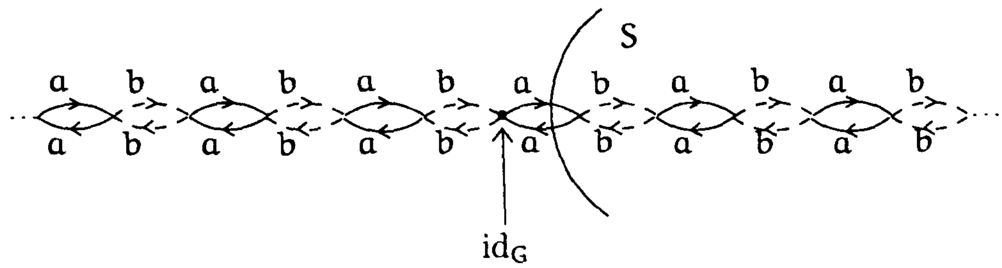

Diagram 3.2.4: $\mathbb{Z}_2 * \mathbb{Z}_2$

The coboundary of S consists of the two edges labelled a incident with id_G. The oriented coboundary consists of these two edges both with the orientation pointing towards S. The stabiliser of S is equal to the stabiliser of this oriented coboundary, and is the trivial subgroup of G. Thus this coboundary has two orbits. (This contrasts with the fact that under the G-action, these two edges are a (proper) subset of a single orbit, because they are both labelled with a.)

3.3 Some non-bounds on coboundaries

It would be convenient if there were some simple bound on the size of a (finite) Bergman-minimal coboundary, so that we would be able to conclude that if there were no coboundary within a certain radius the group has one end. We thought about several possible ways in which a computational approach might be successful, but unfortunately there is a counterexample to each of our ideas.

The groups are straightforward but their presentations are clearly unnecessarily complicated. However, pathological presentations of straightforward groups exist, and it is precisely for difficult-to-understand presentations that we require an algorithm to determine endedness.

3.3.1 Minimality does not imply single orbits

A first hope would be that there is only one orbit of each generator with a given orientation, as in example 3.2.8. Unfortunately, a minimal coboundary may have many orbits.

Example 3.3.1

Let $G = \langle a, b \mid b = a^3 \rangle \cong \mathbb{Z}$.

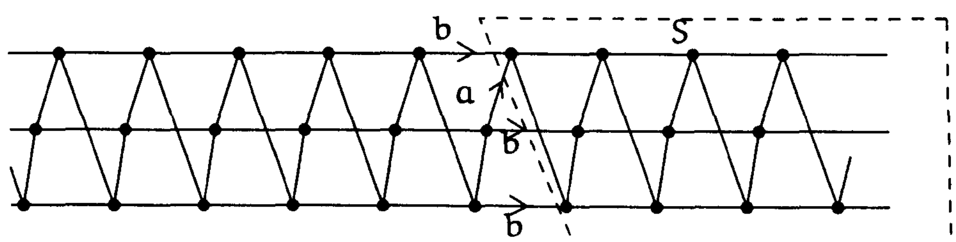

Diagram 3.3.1: $\mathbb{Z}$ generated by 1 and 3

A Bergman minimal coboundary is shown in the diagram; the reader may confirm that no set containing fewer edges disconnects the Cayley graph, and that any set with four edges which does so is a G-translate of the one indicated. Since G acts by graph isometries on its Cayley graph, all such sets of edges have the same Bergman norm.

Here the stabiliser of the oriented coboundary is again trivial. There are three orbits of edges labelled b and oriented into S.

3.3.2 Non linear growth in number of orbits with respect to presentation length

Consider the sequence of presentations

$$P_n = \langle a, b_1, b_2, \dots, b_n \mid b_1 = a^2, \{b_{i+1} = b_i^2 \mid i \leqslant n-1\}\rangle$$

of the group $\mathbb{Z}$.

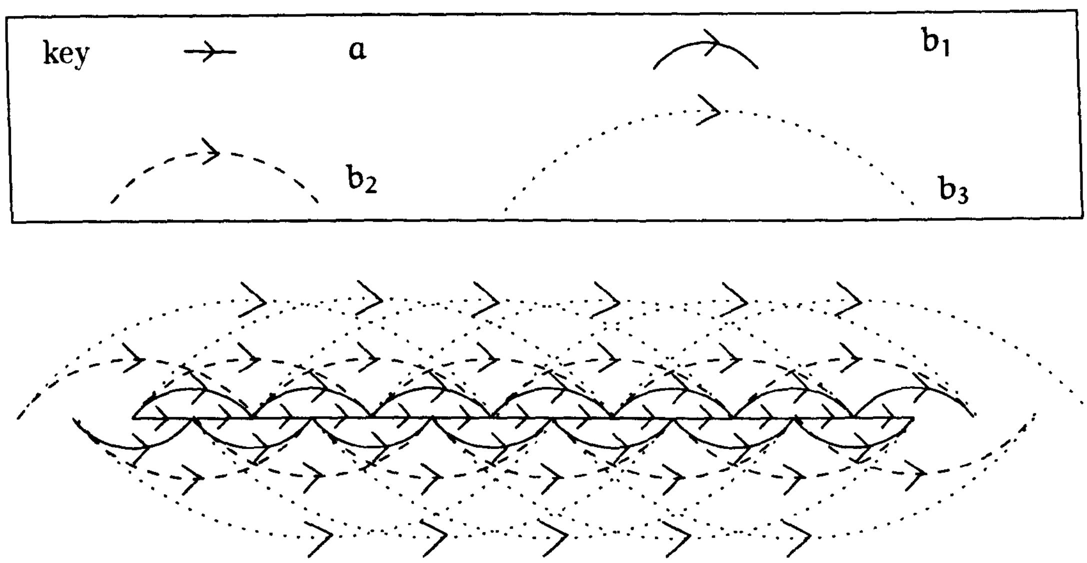

Diagram 3.3.2: $\mathbb{Z}$ generated by powers of 2

This presentation is of length $n+1+3n$, i.e. there are $n+1$ generators and words totalling length $3n$ in the presentation.

A generator b_i equal to a^{2^i} appears 2^i times in a minimal coboundary (without 2^i copies of such a generator, the Cayley graph cannot be disconnected by a finite set of edges). Thus there are $1+2+4+8+\cdots+2^n$ edges in a minimal coboundary. There are no finite subgroups of $\mathbb{Z}$ other than the trivial subgroup, so the stabiliser of a minimal coboundary *is* the trivial subgroup. Thus there are $2^{n+1} - 1$ orbits in a minimal coboundary with the presentation P_n.

Thus there can be no linear bound in terms of the length of the presentation on the number of orbits in a minimal coboundary.

Notice, however, that in this sequence of examples, the closed ball of radius $n + 1$ always disconnects the Cayley graph. (An element a^i can be written as a word of length $\lceil \log_2 i \rceil$ in the generators, and the minimal coboundary can be translated so that every edge of it starts at a vertex expressible as a^i for $i \leqslant 2^n$. So each edge of the minimal coboundary is within the closed ball of radius $\lceil \log_2 2^n \rceil + 1 = n + 1$.)

3.3.3 No bound for splitting ball diameter in terms only of κ

Recall from section 2.1 on page 33 that the connectivity constant, κ, of G is given by $\kappa = \lfloor \frac{1}{2}(\text{max relator length}) \rfloor - 1$. The presentations in the previous section each have the same value for κ, namely 0. Since the splitting ball diameter in the sequence is unbounded, there is no bound for the splitting ball diameter that depends only on κ.

3.4 kth coordinate of Bergman norm may be necessary

Recall from definition 3.1.2 on page 66 that the Bergman norm $\mu(\delta S)$ of a finite coboundary δS is the sequence $(\mu_i(\delta S))_{i \in N}$ where $\mu_i(\delta S)$ is the number of edge paths of length i which start in S and end in S^*.

In section 4.4 on page 114 we use the Bergman norm in an algorithm to find a finite subgroup over which a group with more than one end splits. We always use only finitely many elements of the norm, but can we bound in advance how many we need? In this section we answer the frequently asked question 'Do we need all the components of the Bergman norm?'.

Rephrased, the question comes in two forms: Firstly, is there some bound k such that for any group G, if a coboundary is minimal in the first k terms of its Bergman norm, then it is minimal with respect to all terms of the Bergman norm? (Can we determine Bergman-minimality with a bounded number of terms of the norm?) Secondly, is there some bound k such that for any group G, if a coboundary is minimal in the first k terms of its Bergman norm, then it is G-nested? (Can we determine G-nestedness with a bounded number of terms of the norm?)

Since Bergman-minimal coboundaries are nested (see Dunwoody's theorem 3.1.5 on page 67 from the research talk [Dun98]), a yes to the first is a yes to the second, and contrapositively, a no to the second is a no to the first.

The answer to both questions is no:

Example 3.4.1

For every $k \in \mathbb{N}$ there is a group G_k with coboundaries δS and δT which have $\mu_i(\delta S) = \mu_i(\delta T)$ for $i \leqslant k$, but for which S is G-nested, and T is not G-nested (hence not minimal).

The rest of this section is devoted to constructing such a sequence of groups.

Definition 3.4.2 (neighbourhood) *Let* G *be a finitely generated group with Cayley graph* Γ. *Let* X *be a subset of* Γ, *and let* $r \in \mathbb{R}$. *We denote the closed neighbourhood of radius* r *about* X *by* $\mathcal{N}_r(X)$, *i.e.*

$$\mathcal{N}_r(X) = \{x \in \Gamma \mid d(x, X) \leqslant r\}$$

3.4.1 A sufficient condition for having similar Bergman norm

Lemma 3.4.3

Let G *be a finitely generated group with Cayley graph* Γ. *Let* E_1, E_2 *and* E_3 *be sets of edges of* Γ, *such that* $E_1 \cup E_2 = \delta S$ *and* $E_1 \cup E_3 = \delta T$ *for some subsets* S *and* T *of* Γ. *Let* $k \in \mathbb{N}$.

Suppose that

$$\mathcal{N}_{k-1}(E_2) \cong \mathcal{N}_{k-1}(E_3)$$

in the sense that they are isomorphic as graphs, and that both $d(E_1, E_2)$ *and* $d(E_1, E_3)$ *are at least* $2k$. *Then for all* $i \leqslant k$, $\mu_i(\delta S) = \mu_i(\delta S)$.

Proof

Note first that a path of length i crossing an edge e of E must stay within distance $i - 1$ of e, so the number of such paths, $\mu_i(E)$ depends only on the path components of $\mathcal{N}_{i-1}(E)$. The distance between these path components does not affect the ith term of the norm.

Now $\mathcal{N}_{k-1}(E_1)$ and $\mathcal{N}_{k-1}(E_2)$ have empty intersection, because $d(E_1, E_2) \geqslant 2k > 2(k-1)$. Thus, for $i \leqslant k$, the path components of $\mathcal{N}_{i-1}(\delta S)$ are subsets of either $\mathcal{N}_{k-1}(E_1)$ or $\mathcal{N}_{k-1}(E_2)$.

Similarly, the path components of δT are subsets of either $\mathcal{N}_{k-1}(E_1)$ or $\mathcal{N}_{k-1}(E_3)$.

Since $\mathcal{N}_{k-1}(E_2)$ and $\mathcal{N}_{k-1}(E_3)$ are isomorphic as graphs, the path components of them are too, so their contribution to the ith entry in the Bergman norm is the same. Since there is no path of length less than $2k$ from either of them to $\mathcal{N}_{k-1}(E_1)$, we know that for $i \leqslant k$, $\mu_i(E_1 \cup E_2) = \mu_i(E_1 \cup E_3)$, i.e. $\mu_i(\delta S) = \mu_i(\delta T)$.

□

3.4.2 A sequence demonstrating need for the kth entry of the Bergman norm

Define

$$G_k = \langle a, b, c \mid a^2, b^{4k}, c^2, (b^{2k}a)^4, (bc)^2 \rangle.$$

Below is an incomplete fragment of the Cayley graph of G_k for $k = 2$. The edges are not all drawn at the same scale.

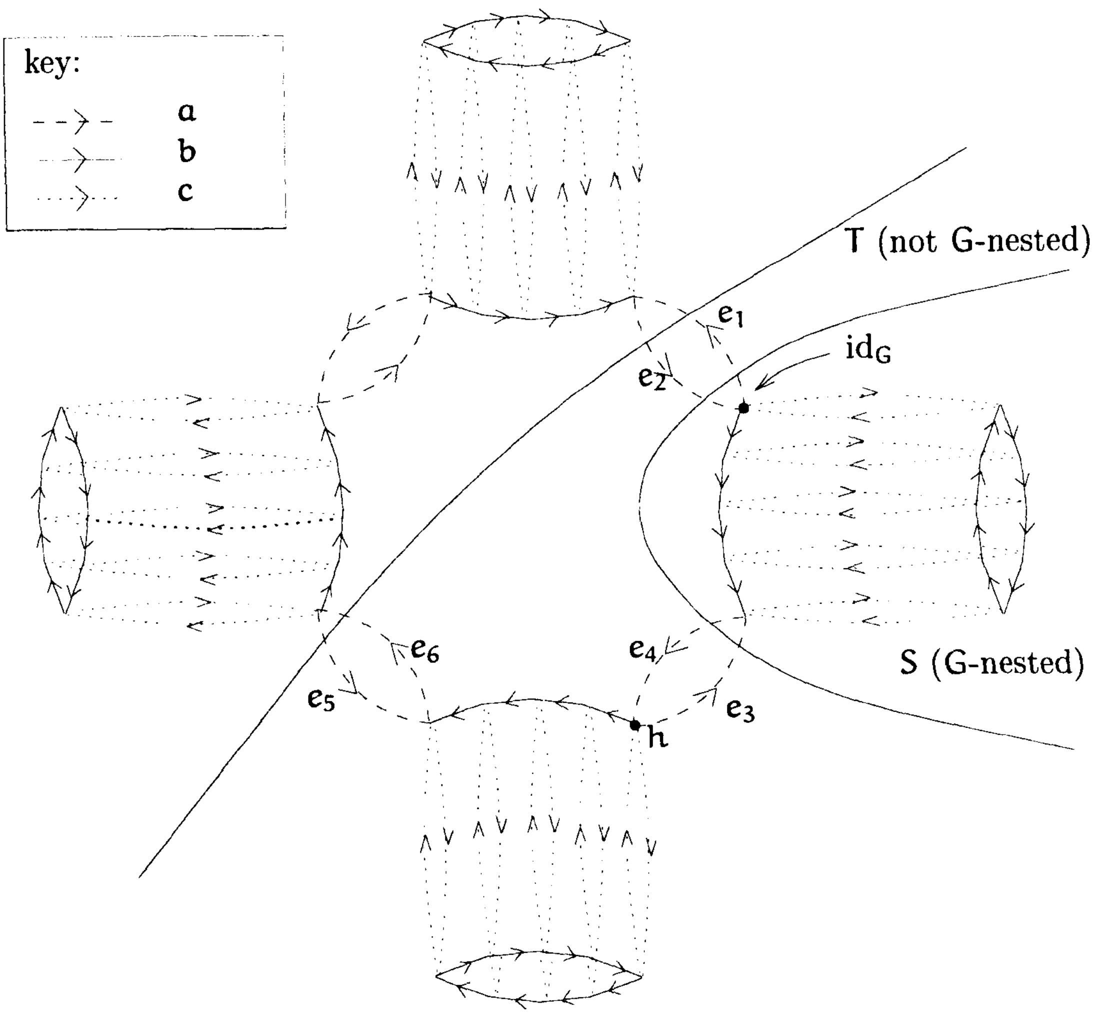

Diagram 3.4.1: T is a non-nested set with small norm.

Consider the edges e_1, e_2, e_3, e_4, e_5 and e_6, as labelled in the diagram. Define $E_1 = \{e_1, e_2\}$, $E_2 = \{e_3, e_4\}$ and $E_3 = \{e_5, e_6\}$.

Now $E_1 \cup E_2$ separates the Cayley graph into two infinite components, as does $E_1 \cup E_3$, so that $E_1 \cup E_2 = \delta S$ and $E_1 \cup E_3 = \delta T$ as indicated on the diagram.

Then E_2 and E_3 are both translates of E_1 under the G-action on Γ. Precisely, $E_2 = (b^{2k}a)E_1$, and $E_2 = (b^{2k}a)^2E_1$. Thus the sets $\mathcal{N}_{k-1}(E_j)$ are all isomorphic as graphs. Also, $d(E_1, E_2) = 2k$ and $d(E_1, E_2) = 4k + 1 \geqslant 2k$.

Thus the hypotheses of the lemma are satisfied, and $\mu_i(\delta S) = \mu_i(\delta T)$ for $i \leqslant k$.

The element $h = b^{2k}a$ has the property that $h \in T \cap hT$, $id_G \in T \cap hT^*$, $h^2 \in T^* \cap hT$, and $h^3 \in T^* \cap hT^*$. Thus T is not G-nested.

However, S *is* G-nested:

If we had $g \in G$ so that (S, gS) were not nested, g would need to move one edge of δS to to be Γ-within S and move another edge of δS to to be Γ-within S^*.[1] Whether an edge is Γ-within a set or not is determined only by whether its vertices are; e_1 is Γ-within some set if and only if e_2 is, and similarly for e_3 and e_4.

It is clear that $g(id_G) \in S$ if and only if $g(b^{2k}) \in S$, because the path between them described by the word b^{2k} cannot cross δS, which consists only of edges labelled by a. So if e_1 is Γ-within S, e_3 cannot be Γ-within S^*, and if e_1 is Γ-within S^*, e_3 cannot be Γ-within S. Thus there is no $g \in G$ for which (S, gS) is a non-nested pair.

Thus S is a G-nested subset of G, and T is not, but their Bergman norms do not differ in the first k places.

We do not know in advance how many components of the Bergman norm we will need to calculate to be able to find a nested set.

[1]This is proved in lemma 4.4.7 on page 120; S is Γ-connected, as is S^*. Lemma 4.4.7 gives us the required condition for the nestedness of the pair (S, gS). This example is not used later, so the is no circularity of reasoning here.

3.5 Quasi-path-connected coboundaries

In this section we prove that if S and $S^\star$ are Γ-connected, then the edges of δS are close to each other. As a corollary, we can bound the diameter of such a coboundary in terms of the number of edges it contains.

Definition 3.5.1 *Let* G *be a finitely generated group with Cayley graph* Γ*, and let* E *be a set of edges of* Γ*. We say* E *is* q-quasi-path-connected *if for every pair* $\{e, e'\}$ *of edges of* E*, there is a finite sequence* $(s_j)_{j=1}^{n+1}$ *of edges of* E *such that* $s_1 = e$*,* $s_{n+1} = e'$ *and*

$$\forall j \in \{1, \ldots, n\}, \quad d(s_j, s_{j+1}) \leqslant q.$$

The reader may recall a similar condition in the conclusion of theorem 2.1.9 on page 37.

Theorem 3.5.2

Let G *be a finitely presented group with Cayley graph* Γ*, and let* S *be a subset of* G*. Let* κ *be the connectivity constant of* G*. (See section 2.1.)*

If S *and* $S^\star$ *are* Γ*-connected subsets of* G*, then the coboundary of* S*,* δS *is* κ*-quasi-path-connected.*

Proof

Suppose that S and $S^\star$ are Γ-connected.

Let $e = [g_1, g_1 a]$ and $e' = [g_2, g_2 b]$ be edges of δS. Since $S^\star$ is Γ-connected, there is a path α_1 between g_1 and g_2 which does not meet S. Similarly, since S is Γ-connected, there is a path α_2 between $g_2 b$ and $g_1 a$ which does not meet $S^\star$.

By theorem 2.1.9 on page 37, part 2, there is a sequence $(s_j)_{j=1}^{m'+1}$ of edges in δS, such that

$$\forall j \in \{1, \dots, m'\}, \quad d(s_j, s_{j+1}) \leqslant \kappa.$$

Thus the sequence $(s_j)_{j=1}^{m'+1}$ between the arbitrarily chosen edges e and e' of δS shows that δS is κ-quasi-path-connected, as required.

□

Corollary 3.5.3

Let G be a finitely generated group with Cayley graph Γ, and let S be a subset of G. Let S and $S^\star$ be Γ-connected subsets of G.

If we know there are at most n edges in δS, then the diameter of δS is at most $n(\kappa + 1) + 1$.

□

3.6 A total ordering on the orbits

Recall that by *orbit*, we mean an orbit of coboundary edges under the action of the stabiliser of the oriented coboundary. (Definition 3.2.7.)

In this section, we further analyse the orbit structure of G-nested coboundaries, by defining 'innermost' and 'outermost' orbits with respect to the orientation of an oriented coboundary. This definition will give rise to a total ordering on the set of orbits in the coboundary of edges that are labelled by a given generator.

Lemma 3.6.1

Let $\overrightarrow{\delta S}$ be an oriented coboundary, and let $H = \mathrm{stab}(\overrightarrow{\delta S})$.[2] *Let* O_1 *and* O_2 *be orbits in* $\overrightarrow{\delta S}$ *(i.e. orbits of edges under the action of* H*), and suppose* g *is an element of* G *which translates an edge of* O_1 *to an edge of* O_2.

Then the set of all elements that translate some edge of O_1 *to some edge of* O_2 *is exactly the double coset* HgH. *Formally,*

$$HgH = \{k \in G \mid \exists e_1 \in O_1, e_2 \in O_2 \text{ such that } ke_1 = e_2\}$$

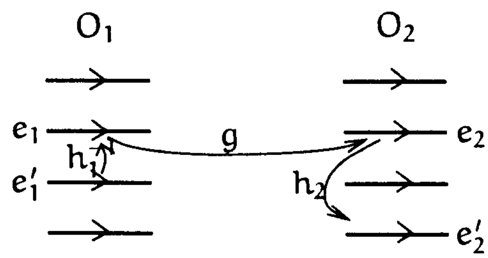

Diagram 3.6.1: Orbits and double cosets

Proof

Suppose $ge_1 = e_2$ with $e_1 \in O_1$ and $e_2 \in O_2$. Let $e_1' \in O_1$ and $e_2' \in O_2$. Then there exists $h_1 \in H$ such that $h_1e_1 = e_1'$, and there exists $h_2 \in H$ such that $h_2e_2 = e_2'$. Let $k = h_2gh_1^{-1}$. Then $ke_1' = h_2gh_1^{-1}e_1' = h_2ge_1 = h_2e_2 = e_2'$

Conversely, any element $k = h_2gh_1^{-1} \in HgH$ takes some edge h_1e_1 of O_1 to an edge h_2e_2 of O_2.

□

Definition 3.6.2 *Let* O_1 *and* O_2 *be orbits in a* G*-nested, oriented coboundary* $\overrightarrow{\delta S}$. *We say that* O_1 *is* further out *than* O_2 *if* $O_1 \neq O_2$ *and there exist* $g \in G$, $e_1 \in O_1$, *and* $e_2 \in O_2$ *such that* $ge_1 = e_2$ *and* $gS \subset S$.

[2]See section 3.2 on page 69 for definitions.

We write $O_1 \prec O_2$, *or* $O_2 \succ O_1$, *or say that* O_2 *is* further in *than* O_1.

Lemma 3.6.3

$\prec$ *is well-defined.*

Proof

Let $H = \text{stab}(\overrightarrow{\delta S})$. Suppose $O_1 \prec O_2$. Then there exists g in G and edges e_1 of O_1 and e_2 of O_2 such that $ge_1 = e_2$ and $gS \subset S$. Suppose also $k \in G$ with $ke_1' = e_2'$ for some e_1' of O_1 and e_2' of O_2. Then $k \in HgH$ by the previous lemma, so $k = h_1 g h_2$, for some h_1 and h_2 in H. Then $kS = h_1 g h_2 S = h_1 g S \subset S$, because $gS \subset S$ and $h_1 \in H = \text{stab}(\overrightarrow{\delta S})$.

□

Recall that there is an element in G taking e_1 to e_2, if and only if e_1 and e_2 are in the same G-orbit. This is true if and only if e_1 and e_2 are labelled by the same generator. Clearly then an orbit in a coboundary consists of edges that are all labelled by the same generator.

Lemma 3.6.4

Let $\overrightarrow{\delta S}$ *be a* G*-nested, oriented coboundary, and let* a *be a generator of* G. *Then the set of orbits in* $\overrightarrow{\delta S}$ *that have edges labelled* a *are totally ordered under the relation* $\prec$.

Proof

We prove that $\prec$ is transitive and satisfies the trichotomy condition:

Trichotomy

Let O_1 and O_2 be orbits of $\overrightarrow{\delta S}$ whose edges are labelled by the same generator, and suppose that $O_1 \neq O_2$. We prove that either $O_1 \prec O_2$ or $O_1 \succ O_2$.

Let $e_1 \in O_1$ and $e_2 \in O_2$. Then since e_1 and e_2 are labelled by the same generator of G, there is some element $g \in G$ such that $ge_1 = e_2$.

Since e_1 and e_2 are in $\overrightarrow{\delta S}$, their initial vertices are in $S^\star$ and their terminal vertices are in S, by the definition of oriented coboundary.

Consider the initial vertex of $e_2 = ge_1$. It is in $S^\star$, but also in $g(S^\star) = (gS)^\star$, so $S^\star \cap (gS)^\star \neq \emptyset$. Similarly, the terminal vertex of $e_2 = ge_1$ demonstrates that $S \cap gS \neq \emptyset$. Since S is G-nested, one of the sets $S \cap gS$, $S \cap gS^\star$, $S^\star \cap gS$, or $S^\star \cap gS^\star$ must be empty, so $S \cap (gS)^\star$ is empty or $S^\star \cap gS$ is empty.

Thus we have that $ge_1 = e_2$, and $S \subset gS$ or $gS \subset S$. Recall from 3.2.5 that $\mathrm{stab}(S) = \mathrm{stab}(\overrightarrow{\delta S})$.

Case (1): Both $S \subset gS$ and $gS \subset S$.

In this case, $S = gS$, so $g \in \mathrm{stab}(S)$, but then g cannot move e_1 outside its orbit, i.e. $e_2 = ge_1 \in O_1$, and so $O_1 = O_2$.

Note that conversely, if $O_1 = O_2$, then $g \in \mathrm{stab}(S)$, so both $S \subset gS$ and $gS \subset S$.

Case (2): $S \subset gS$ but $gS \not\subset S$.

Then $O_1 \neq O_2$ by the above remark, so by definition, $O_1 \prec O_2$.

Case (3): $gS \subset S$ but $S \not\subset gS$.

Then $O_1 \neq O_2$ by the remark above. Since $gS \subset S$, $S \subset g^{-1}S$. Now $g^{-1}e_2 = e_1$, so by definition, $O_2 \prec O_1$.

Transitivity

Suppose O_1, O_2 and O_3 are orbits in $\overrightarrow{\delta S}$, and assume that $O_1 \prec O_2$ and $O_2 \prec O_3$. We prove that $O_1 \prec O_3$.

Since $O_1 \prec O_2$, there exist $g_1 \in G$, $e_1 \in O_1$, and $e_2 \in O_2$ such that $ge_1 = e_2$ and $g_1S \subset S$. Since $O_2 \prec O_3$, there exist $g_2 \in G$, $e_2' \in O_2$, and $e_3 \in O_3$ such that $ge_2' = e_3$ and $g_2S \subset S$.

Let $H = \text{stab}(\overrightarrow{\delta S}) = \text{stab}(S)$. Since both e_2 and e_2' are in O_2, there is an element $h \in H$ such that $he_2 = e_2'$. Define $g = g_2hg_1$. Then $ge_1 = g_2hg_1e_1 = g_2he_2 = g_2e_2' = e_3$. Also, $g_1S \subset S$, so since $h \in \text{stab}(S)$, $h(g_1S) \subset S$. Thus since $g_2S \subset S$, $g_2(h(g_1S)) \subset g_2(S) \subset S$, so $gS = g_2hg_1S \subset S$. Thus either $O_1 \prec O_3$ or $O_1 = O_3$.

If $O_1 = O_3$ then $O_1 \prec O_2$ and $O_2 \prec O_1$, which cannot happen by the trichotomy condition above.

Thus $\prec$ is a total ordering, as required.

□

3.7 Could we bound the size of a minimal coboundary?

In this section, we use the term *minimal coboundary* to refer to a coboundary which is Bergman-minimal amongst the coboundaries of infinite subsets of G which have infinite complement.

We are interested in bounding the diameter of a minimal coboundary in terms of a computable function of the presentation (or an automatic structure, if we have one). If we were to find such a bound, we would be able to detect one-endedness in groups computationally. Briefly, this bound would give us a diameter within which, if the group has more than one end, there must be a coboundary δS with S and S^* infinite. If the group also has a solution to the word problem, or, even better, an automatic structure, we would then be able to completely solve the algorithmic problem of determining the number of ends, using the techniques outlined in chapter 4.

In section 3.3, we showed that some simple ideas for bounds on the size of coboundaries fail. It is the case, however, that in all our examples, the complexity of the presentation increases as we increase the number of orbits or the diameter of a minimal coboundary.

A direct bound on the diameter of a minimal coboundary would solve our problem, but it seems unlikely that this can be deduced directly from the presentation. A presentation for a group can be thought of as a specification of the geometry of a small part of the Cayley graph. This geometry may induce some unpredictable large scale geometry. The triangle groups are an example of this sort of behaviour — in some cases the geometry gives rise to positive curvature, and the group is finite, and in others, negative curvature, and the group is infinite. Of particular concern to us is that a presentation could conceivably be quite simple, but yet define a group which splits over a very large finite subgroup.

Perhaps we should assume that we must prove for ourselves a bound on the number of elements in finite subgroups of our group before we can calculate a bound on the size of a minimal coboundary. In a hyperbolic group, we can calculate such a bound from the constant of hyperbolicity, as shown in [BG95].

It would then be enough to bound the number of orbits in a minimal coboundary, because then we may obtain a bound on the number of edges in a minimal coboundary. Lemma 4.4.11 on page 126 implies that if δS is a minimal coboundary, then S and S^* must be Γ-connected, and in section 3.5, we proved that such coboundaries are κ-quasi-path-connected. Then a bound on the number of edges multiplied by $(\kappa + 1)$, plus 1, would be a bound on the diameter. (Corollary 3.5.3 on page 84.)

Chapter 4

Computing a splitting over a finite subgroup

4.1 Overview and related work

Sela was the first to prove that there is an algorithm which can decide whether a torsion-free hyperbolic group splits as a free product. However, the algorithm which detects one-endedness is a mammoth check which would be infeasible even on fairly simple examples. It is unlikely to ever be used in a computer program. Sela's algorithm is part of the content of the as yet unwritten paper *The isomorphism problem for hyperbolic groups part II.*

Sela's algorithm uses the method of canonical representatives to reduce the solution of equations with alternating quantifiers in a torsion-free hyperbolic group into large sets of such equations in a free group, which are then solvable. In this way he can calculate possible homomorphisms from the group to itself, and whether a given presentation of a splitting holds true in the group. If the group does split over a finite subgroup, the algorithm finds a presentation for the splitting, and if the group is one ended the algorithm finishes calculating its JSJ decomposition. It seems clear that, at many points, the approach is not feasible.

Recently, Gerasimov [Ger] generalised Sela's result, proving that there is an algorithm which decides whether an arbitrary hyperbolic group splits as a free product.

Gerasimov's algorithm involves constructing a sequence of simplicial complexes whose vertices are geodesic words of a given length n. If this complex is connected and satisfies a certain condition for large enough n, the boundary is connected and locally connected, so the group is one-ended. We discuss and re-present Gerasimov's algorithm in chapter 5.

Also, Dunwoody and his student Barker have been working towards finding an algorithmic solution to the endedness problem, assuming only that the group has solvable word problem. Their algorithms involve finding tracks in a simplicial presentation 2-complex for the group, and can detect one-endedness under certain conditions.

Our algorithm assumes that the group is automatic, and we have a separate approach to detecting each of the possible number of ends. The algorithms for detecting zero, two and uncountably many ends exist in the form of a program, kindly coded in C by David Hind as his fourth-year undergraduate project. The program runs quickly on our test presentations, and there is scope for making it much more efficient by avoiding some needless repetition. The algorithm for detecting zero ends is a simple check that there are no loops in the word acceptor, and we use a program written by Billington to detect two endedness. Our program for detecting more than two ends currently relies on the accepted words being geodesics, but theoretically this condition can be relaxed, as we shall see in sections 4.2.2.2 and 4.2.3.

In fact, theoretically, we can detect having more than one end assuming only that the group has a solution to the word problem (section 4.3). Because our algorithm which may detect one-endedness depends on an automatic structure, it cannot be used in this context. But the algorithms to detect zero or two ends *can* be used, because such groups are always hyperbolic and therefore automatic. (See [SW] for a proof that 2-ended groups have a subgroup of finite index which is isomorphic to $\mathbb{Z}$, [GdlH90] for a proof that such groups are hyperbolic, and [Pap95] or [ECH+92] for a proof that hyperbolic groups are automatic.)

4.1.1 Outline of the algorithm

Where we give two references to later sections, the first deals with the case when there is an automatic structure for the group, and the second deals with the case when there is only a solution to the word problem.

Let κ be the connectivity constant of G. (See section 2.1 on page 33.)

1. If the group is zero- or two-ended, terminate, saying so. See section 4.2.4 on page 106.

2. Calculate the vertices of $\overline{B}_{n+\kappa}$ for some $n \in \mathbb{N}$. See section 4.2.2, or 4.3.1.

3. Calculate the Γ-components of $\overline{B}_{n+\kappa} \setminus B^{\circ}_{n}$, using several vertex depth-first searches. (See section 1.2.2.2 on page 25.) This is the same as

calculating the components of $\Gamma \setminus B_n^\circ$, by theorem 2.1.12.

4. Decide by some means which components are infinite. See section 4.2.3 or 4.3.2.

5. If the number of infinite components of $\Gamma \setminus B_n^\circ$ is three or more, terminate, saying the group is infinitely ended. If not, increase n and start again.

4.2 Practical: automatic groups

We assume that G is automatic, and let WA be its word acceptor, and GM its general multiplier.

In this section we describe the algorithms used in the program **ends**. The program **ends** assumes that the words accepted by WA are geodesics, but we shall see that this assumption is unnecessarily restrictive — we used it only to simplify the program.

4.2.1 Representation of balls and their complement

Let $G = \langle X \mid R \rangle$. Finite subsets of the Cayley graph are represented by a collection of arrays of vertices.

Each vertex v has an array of adjacent vertices $\{w_x \mid x \in X \text{ or } x^{-1} \in X\}$. We think of an oriented edge as the pair (v, w_x) labelled by the generator x. Thus each edge of the Cayley graph appears twice, once for each vertex it is incident with. We think of this as using two oriented edges to represent each unoriented one.

When calculating the ball of a certain radius, we construct vertices and a spanning tree first and then the remaining edges. We calculate exactly those edges that begin and end at vertices which are already constructed.

In essence, the Cayley graph is stored as a set of vertices, with information about which vertices are adjacent via which generators or generator inverses. In this sense the computer representation of Γ contains no edges; when we use the word 'edge' in this context we really mean an adjacency relationship from v to w_x, labelled x.

In fact when we construct parts of Γ, an 'edge' exists if and only if both its vertices do; with the computer we can only test Γ-connectedness.

When we say that we have computed B_n° or $\overline{B}_{n+\kappa}$, we mean that we have constructed the vertices of these sets and the adjacency relationships between them. However, since an edge is present in a computed subset of Γ if and only if both its vertices are, we can test Γ-connectedness reliably within the ball we have constructed.

In fact, in the program, rather than calculate the Γ-components of $\overline{B}_{n+\kappa} \setminus B_n^\circ$, we calculate the Γ-components of $\overline{B}_{n+1+\kappa} \setminus \overline{B}_n$, which is the same as calculating the Γ-components of $\overline{B}_{n+1+\kappa} \setminus B_{n+1}^\circ$. (Using $n+1$ in place of n is a quirk of the original computer program that has persisted; it makes no difference mathematically.)

4.2.2 Constructing balls quickly ...

In this section we describe the computation of an arbitrary ball of radius n.

4.2.2.1 ... assuming that all accepted words are geodesic

Every element of G is represented by a word accepted by WA. We may, without loss of generality, assume that WA accepts a unique word for each element. (See section 2.5 of [ECH+92].) Thus each element is represented by a path in WA from the start state to an accept state. We wish to find all such paths of length at most n.

When we just want to find all the vertices of a graph, we do a vertex depth-first search, in effect, calculating a spanning tree. In this case we want to find all paths to all vertices, so we do a path depth-first search. (See section 1.2.2.2 on page 24.) In essence we are calculating a finite portion of a covering tree of the finite state automaton WA.

Whilst performing this path depth-first search, we make a list of the generators labelling the edges we travelled along. Thus we may calculate a list of all accepted words of length less at most n. This gives us a spanning graph for $\overline{B}_n$.[1]

To calculate the 'edges', i.e. adjacency relationships, it is enough to calculate all pairs of accepted words of length at most n which terminate at distance 1 from each other. Fortunately, the general multiplier, GM, accepts all such pairs, and as before, we may perform a path depth-first search to calculate them.

[1] Or a spanning tree, if the accepted language is prefix closed and has unique representatives for each element.

4.2.2.2 ... without assuming that all accepted words are geodesic

If not all the words accepted by WA are geodesic, then we may need to use words of length greater than n to construct $\overline{B}_n$.

However, since there is an accepted word for each element of G, and there are finitely many vertices in $\overline{B}_n$, there is an $m \in \mathbb{N}$ such that each element of $\overline{B}_n$ has an accepted path to it of length at most m. Thus, to construct $\overline{B}_n$ reliably, all we need is a way of checking that we have calculated all vertices of $\overline{B}_n$.

An algorithm to calculate $\overline{B}_n$ is as follows: Calculate and store the vertices and adjacency relationships[2] represented by words of length up to m for any $m \geqslant n$. Then begin at id_G in this stored graph, and perform a vertex depth-first search to depth $n-1$. At each vertex we check that all adjacent vertices of Γ have been calculated, by checking that all outgoing 'edges' exist. If so, the whole of $\overline{B}_n$ has been calculated. If not, we increase m and repeat the adjacency test.

Of course, once the ball $\overline{B}_n$ has been constructed, its manner of construction is unimportant.

An upper bound We can calculate an upper bound on the length of an accepted word ending in S_n. If we choose, we can forgo the repeated adjacency test above, and instead compute all edges reached by words up to this length. It is not clear whether this will be more, or less efficient.

Let g be an element of S_n, and let w be a geodesic word representing it. Let N be the number of states in the largest of the multiplier automata, and

[2](adjacent in the sense that they differ by a generator)

let n_0 be the length of the shortest accepted representative of the identity. Then by lemma 2.3.9 of [ECH+92][3] there is an accepted word for g of length at most $Nn + n_0$.

We may calculate n_0 by first calculating any accepted representative, v, of the identity (see the end of proof 2.3.10 of [ECH+92]). Let M_ε be the equality recogniser (i.e. the automaton accepting pairs of words that represent the same group element). We ennumerate all words equal to id_G by performing a restricted path depth first search of M_ε; we only allow ourselves to travel along paths where the sequence of generators from the left side of each pair spells the word v. The other side then spells a word equal to id_G, and every such word appears in this way. As in the proof of 2.3.9 in [ECH+92], we may ignore paths which take us around a loop in M_ε, and thus obtain a finite list of words. The length of the shortest of these words is n_0.

We could of course skip this last calculation and use the length of v as n_0. However it pays to avoid calculating words of unnecessarily great length, particularly in a group with an automatic structure so awful that the identity element is only represented by long words!

4.2.2.3 Possible efficiency gain

This subsection does not form part of the theory, but deals with some implementation details with which we may speed up the actual program.

[3] In [ECH+92], the N taken is any strictly greater than the N we use here. A brief look at where N is used in the proof shows that this is unnecessary and that our N will do. It is worth taking note of this, as N is a factor in the length we need to go to; this saves us n units of length, and thus possibly a very large number of calculations.

The program uses Derek Holt's program *fsaenumerate*, (part of the package KBMAG). fsaenumerate takes a finite state automaton and performs the path depth-first search search we described, writing the result to file. fsaenumerate was used to save time in coding, but it is unnecessarily inefficient to call this external program:

Firstly, writing to disk is very much slower than manipulating information in the memory. Secondly, in reading the information from the file, it must be translated; strings of characters represent the generators, and this information must be translated into the internal representation.

These first two problems are real, but insignificant compared to the third. We need to calculate balls of increasing radius. Now if we were to perform the path depth-first search in the program, we could freeze the state of the search, and return from where we left off when we needed to increase the radius. fsaenumerate has no such facility, and if we wish to calculate the ball of radius $n+1$, fsaenumerate must start again from the identity.

fsaenumerate allows you to specify a lower as well as upper bound for the length of words you require, but its calculation starts from the identity every time the program is called. The lower bound limits only the output generated.

4.2.3 Deciding which components of a ball complement are infinite

To calculate the components of $\Gamma \setminus B^{\circ}_{n}$, we calculate the Γ-components of the band, $\overline{B}_{n+\kappa} \setminus B^{\circ}_{n}$. Of course the band is finite, but we want to be able to tell which of its Γ-components represent infinite components of $\Gamma \setminus B_n^\circ$.

We may assume without loss of generality that we have a unique accepted word representing each group element,[4] and it is also true that we may assume that the language accepted by WA is prefix closed.[5] We do not know that we can assume without loss of generality that the word acceptor has both properties at once, but we shall make that hypothesis in the following lemma. If the finite state automaton does not have both properties at the same time, we may test whether the group has more than one end using the method outlined in section 4.3.2 on page 112. Note that current software only generates automata with both uniqueness and prefix-closure; in practice, the hypotheses of the lemma are satisfied.

Lemma 4.2.1

Let G be an automatic group with an accepted language which is prefix closed and has a unique representative for each group element.

A component C of $\Gamma \setminus B_n^\circ$ is infinite if and only if there is an infinite accepted word ω, which meets C after it has left B_n° for the last time.[6]

Proof

First we prove the converse. If there is such a word, ω, then it is injective, because it cannot re-visit any vertex of Γ; all its prefixes are accepted and there is only one representative for each group element. Then, by lemma 2.3.1

[4] See section 2.5 of [ECH+92].

[5] See section 2.5 of [ECH+92].

[6] Note that since the accepted language is prefix closed, all states of the word acceptor are accept states. ω is an infinite accepted word in the sense that all subwords of it are accepted.

on page 62, C is infinite.

Secondly, we assume C is infinite, and prove the existence of ω. Since C is infinite, there are words which terminate in it at arbitrarily large distance from the identity.[7]

Only finitely many accepted words may pass into C and then return into B°_n, because each element of the group has only one accepted word passing through it, and there are only finitely many vertices in $C \cap S_n$ for such words to pass through.

Also, there are only finitely many words which terminate in C, but which do not transverse a loop in the word acceptor, WA, because there is a bound on how far you can travel in WA without looping, and there are only finitely many group elements in a ball of finite radius.

Thus there must be infinitely many words which transverse a loop in WA whilst in C, so we may extend infinitely a subword of each. Only finitely many of these extensions can pass back into B°_n, so in fact there are infinitely many words into C which possess infinite extension in C. Thus there is at least one infinite word, ω, which passes into C after it has left B°_n for the last time. (It is possible that all of the finite words we were considering are subwords of this infinite word.)

□

Definition 4.2.2 *Given a finite state automaton* A, *we may remove states from which it is impossible to reach a loop. We call this process* pruning *and call the resultant finite state automaton* pruned. *We denote this automaton*

[7]Note that a path may enter C, then leave it, and then re-enter it.

by pruned(A).

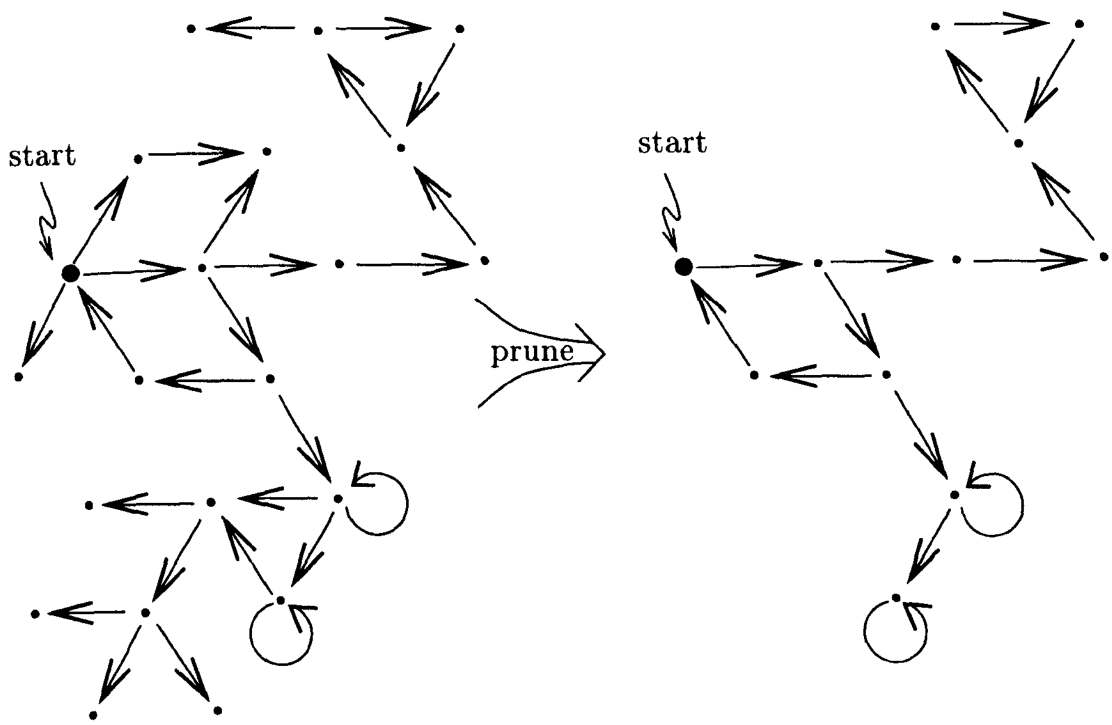

Diagram 4.2.1: Pruning an automaton

The pruned FSA accepts exactly those words in the original language which possess infinite extension. Thus an accepted word in Γ can be infinitely extended if and only if the state in WA in which it is accepted is also present in the pruned acceptor, pruned(WA).

We now make a definition to distinguish between accepted words and subwords of accepted words.

Definition 4.2.3 *Let* A *be a finite state automaton. We call a word* travellable *in* A *if it can be extended to a word accepted by* A, *i.e. if it is a subword of an accepted word. We call an infinite word travellable if every initial finite subword of it is travellable.*

Note that an infinite travellable word could in theory have no accepted subword.

Corollary 4.2.4

Let G *be an automatic group with a word acceptor* $\mathcal{WA}$ *which has every accepted word geodesic.*

Suppose we have constructed $\overline{\mathsf{B}}_{n+\kappa}$*, and that we have additionally labelled each vertex on the Cayley graph with the set of states from* $\mathcal{WA}$ *in which we may be when passing through that vertex.*[8] *This involves no extra computation, because we record the information as we create the ball.*

Then a component C *of* $\Gamma \setminus \mathsf{B}^{\circ}_{n}$ *is infinite if and only if there is an element* $g \in \mathsf{C} \cap (\overline{\mathsf{B}}_{n+\kappa} \setminus \mathsf{B}^{\circ}_{n})$ *which is labelled with a state which is also in the pruned version of the word acceptor.*

Proof

If there is such an element, it may be extended infinitely, and it cannot return into B°_{n} because it is geodesic. So by lemma 2.3.1 on page 62, C is infinite.

Conversely, if C is infinite, there are arbitrarily long accepted words into C, so one of them, ω, must transverse a loop in the word acceptor, so must be infinitely extendable. Thus at every distance from the identity, the state determined by ω is in the pruned automaton. In particular, $\omega(n+\kappa)$ is. Thus there is an element in $\overline{\mathsf{B}}_{n+\kappa} \setminus \mathsf{B}^{\circ}_{n}$ which is labelled by a set of states which includes a state from the pruned automaton.

[8]Note that we list also the non-accept states in our set of states. This listing is possible because all the accepted words are geodesic, and when we constructed the ball we enumerated *all* the travellable words of length $n+\kappa$.

□

Note that, in the program, the finite state automaton given by **autgroup** is a short-lex automaton, so in fact we work with prefix-closed unique geodesics, and so there is only one state in which we may be at any given group element.

We do not need the geodesic assumption on the accepted words, but may need to do more computation if not.

Corollary 4.2.5

Let G *be an automatic group with an accepted language which is prefix closed and has a unique representative for each group element.*

Then there is an algorithm to detect which components of $\Gamma \setminus B_n^\circ$ *are infinite.*

Proof

Construct all words of length up to m and 'edges' between them, with m large enough so that $\overline{B}_{n+\kappa}$ has been constructed.[9] Label each vertex on the Cayley graph with the state from WA in which we were when constructing that vertex.[10]

Consider the group elements determined by the accepted words of length exactly m. Since the whole of B_n° has been calculated, these words may not

[9]See section 2.1 on page 33 for the definition and properties of the connectivity constant, κ.

[10]There is a unique accepted representative for each group element, but also a unique word travellable in the word acceptor for each element; every subword of every accepted word is itself accepted, so all travellable paths passing through a given element must have the same initial subword up to that point.

return to B°_n, because there are unique travellable words to each element. Restrict now to those which possess infinite extension by checking whether the state from WA in which they were accepted also appears in pruned(WA).

There are no infinite geodesics except those extending these words, and these words have all left B°_n for the last time, so a component of the complement of B°_n is infinite if and only if it contains one of these words. We already know how to calculate the components of $\Gamma \setminus B^\circ_n$, so we are done.

□

Definition 4.2.6 *Let* G *be an automatic group. We say that the automatic structure is* good *if one or more of the following is true; either*

- *the accepted words are all geodesics, or*
- *the set of accepted words is prefix-closed and there is a unique accepted word for each group element.*

Algorithm 4.2.7 (detecting more than one end)

Suppose G is a group with a good automatic structure. Then there is an algorithm which answers yes if and only if G has one end. (If G has more than one end, the algorithm will not terminate.)

Let κ be the connectivity constant of G. (See section 2.1 on page 33.)

1. Using one of the algorithms described in section 4.2.2 on page 95, we may calculate the ball of radius $n + \kappa$ for any $n \in \mathbb{N}$.

2. We may calculate the Γ-components of the band, $\overline{B}_{n+\kappa} \setminus B^\circ_n$, and by theorem 2.1.12 on page 51, these represent faithfully the components of $\Gamma \setminus B^\circ_n$.

3. By corollaries 4.2.5 on page 104 and 4.2.4 on page 103, given either of the two types of automatic structure, we may decide which components of $\Gamma \setminus B_n^\circ$ are infinite.

4. Thus if G has a good automatic structure, we may continue calculating how many infinite components there are of $\Gamma \setminus B_n^\circ$ for larger and larger n. If for some $n \in \mathbb{N}$ there is more than one infinite component, terminate, saying so.

If at any stage there is more than one infinite component, then the group has more than one end. Conversely, if the group has more than one end, for some $n \in \mathbb{N}$, B_n° disconnects the Cayley graph into more than one infinite component, and the algorithm detects this.

□

4.2.4 Deciding zero- and two-endedness

If G has zero or two ends, in theory we can always detect this. This is because both finite and two-ended groups are hyperbolic, and thus *strongly geodesically automatic*, by [Pap95]. This means that G has an automatic structure where the set of accepted words is exactly the set of *all* finite geodesics. This means that all pairs of geodesics to pairs of elements differing by a generator fellow-travel. In particular, if we restrict to just the lexicographically first geodesic to each element, this set of geodesics also has the fellow-travelling property. As described in section 2.5 of [ECH+92], we may alter the word acceptor to accept only the lexicographically first representative for each element. Thus a hyperbolic group is necessarily short-lex automatic, and so the program ***autgroup*** from the package KBMAG will in principle calculate the short-lex automatic structure.

Of course, the group may be the monster group, or the monster group cross $\mathbb{Z}$, and the computer does not have the capacity necessary to calculate its automatic structure.

Once we have an automatic structure for a group, we may alter it so that it accepts a unique representative for each element (see section 2.5 of [ECH+92]). Clearly then the group is finite if and only if the word acceptor has no loops.

Recall that the *growth* of a group is the sequence determined by $a_n = |\overline{B}_n \cap G|$, i.e. a_n is the number of elements of the group of distance at most n from the identity in the Cayley graph.

By [Can], a group is two-ended if and only if it has linear growth.

In his M.Sc. dissertation, [Bil96] Stephen Billington described various results relating the growth of an automatic group to the structure of its word acceptor, in particular to the number and nature of strong components. A *strong component* is a subgraph of a finite state automaton which is maximal under the condition that from each state there is a path to each other state.[11] Billington wrote programs which determine the growth of a finite state automaton, and one of them, 2end, tests whether the group has linear growth, and thus whether it is two-ended.

Thus, given a presentation of a group, we may or may not be able to calculate an automatic structure for it using KBMAG. If we do have an automatic structure, we can decide whether or not the group has zero ends (no

[11]Note that this means that between any pair there are paths in both directions.

loops in the word acceptor), and using Billington's program 2end, whether or not it has two ends (linear growth).

4.2.5 Summary

Algorithm 4.2.8 (ends)

Let G be a group given as a finite presentation. Suppose G has a good automatic structure. (See definition 4.2.6 on page 105.)

Then the program *ends* detects if the group has zero, two, or uncountably many ends.

1. First take the presentation and run the KBMAG program autgroup to try to calculate an automatic structure.

2. If this succeeds, check for loops in the word acceptor. If there are none, terminate, stating that the group is finite and hence zero-ended. If there are loops, the group is infinite.

3. Use the program 2end to decide whether or not the group has two ends. If so, terminate, saying that the group has two ends. If not, then the group can only have one or infinitely many ends.

4. Use algorithm 4.2.7 on page 105 to try to detect that the group has more than one end. If it does, terminate, saying that the group has infinitely many ends. (We ruled out two-endedness in the previous step.)

Thus if G has zero, two, or uncountably many ends, ends detects this. If not, it proceeds to calculate larger and larger balls in the Cayley graph, using more and more memory (and with the current implementation using fsaenumerate, more and more disk space).

□

The table on the current page summarises the logical situation.

ends	when we detect this	when we refute this
0	always	always
1	sometimes — see chapter 6	always
2	always	always
∞	always	(sometimes)

Table 4.1: To what extent we can determine the number of ends in a group given a good automatic structure.

The 'sometimes' in brackets in table 4.1 is a simple consequence of the 'sometimes' under detecting one-endedness. Note that it is possible that the algorithm that sometimes detect one-endedness could in fact give an 'always' rather than a 'sometimes'. In the terminology of chapter 6 this would be true if the boundary of an arbitrary one-ended automatic group were always uniformly path-connected, and uniform path-connectedness were equivalent to the exterior paths condition for automatic groups.

4.3 Theoretical: using only the word problem

The program **ends** runs quickly on the examples we have tested, but for simplicity in coding, currently relies on the automatic structure having unique geodesic representatives for each element. As proved in section 4.2 this condition may be relaxed, and we expect some corresponding degradation in speed if so. In this section, we prove that there is a corresponding, but less powerful algorithm for groups which are not necessarily automatic, but which nevertheless have a solution to the word problem. This algorithm is *considerably* slower. The calculation of the ball of radius n is very slow, and the algorithm requires us to calculate a ball of approximately twice the radius needed by the program **ends**. In a group with large growth, this could make the difference between feasible and not feasible.

Thus we consider the results of this section as primarily theoretical results.

Example 4.3.1

To compare the two algorithms on the same group, we take the group $\mathbb{Z} \times \mathbb{Z}_{2n}$ with presentation $\langle a, b \mid aba^{-1}b^{-1}, b^{2n} \rangle$. The open ball of radius $n+1$ is the smallest to disconnect the Cayley graph, and we will need to check this by calculating the ball of radius $n + 1 + \kappa = 2n + 1$.

Using the automatic structure, we calculate the ball of radius $2n + 1$ in time proportional to the total length of all accepted words in the ball, so proportional to $\sum_{i=1}^{n} 4i^2 + \sum_{i=n+1}^{2n+1} 4ni$. [12] Thus the computational complexity

[12]This is because the sphere of radius $i \leqslant n$ has $4i$ elements of length i in it, whereas

is of order $O(n^3)$. The time taken to establish the components of the band $\overline{B}_{n+\kappa} \setminus B^\circ_n$ is the same for the automatic algorithm as for the word problem algorithm, so we ignore this summand. There are at least two infinitely extendible geodesics within the band, and the time taken to find them is at worst proportional to the band's size, i.e. of order $O(n)$.

Using only a solution to the word problem, our algorithm means we must calculate the ball of radius $2(n+1)$. (See section 4.3.2 on the next page.) To calculate this ball we must compare each pair of words from the ball of radius $2(n+1)$ in the free group on a and b for equality. In the free group, we have $1 + 4 \times 3^{2n+1}$ elements in this ball, so we must make comparisons to the order of $O(3^{4n})$, each of which may take non-trivial time.

In this example, it is clear that the automatic algorithm, with a complexity $O(n^3)$ is considerably faster than the word problem algorithm, which has complexity worse than $O(3^{4n})$. This is true despite the quirk that in this case the value of κ meant that $(n+1)+\kappa$ was not much smaller than $2(n+1)$. In general we would expect that κ is somewhat smaller than the radius of the smallest ball disconnecting the Cayley graph, unless we had a particularly nice presentation for the group; in general adding κ to this radius would result in a smaller radius than multiplying it by two, so the automatic algorithm not only calculates balls quicker but need not calculate balls of as large radius as the word problem algorithm.

---

for $i > n$ the spheres have stopped growing, and are all of size $4n$.

4.3.1 Constructing balls

Suppose we are given a presentation for our group, G, and an algorithm which solves the word problem in G. Since we may determine whether a given word represents the identity element, we may test equality between words: $w = u \Leftrightarrow wu^{-1} = \mathrm{id}_G$.

Given a radius n, we wish to calculate the closed ball of radius n, $\overline{B}_n$. Each element at distance at most n from id_G is represented by at least one word of length at most n. We calculate the ball of radius n in the free group on the generators of G,[13] and then find all pairs of words which represent the same group element using the solution to the word problem. This calculation is expensive computationally.

Having done so, we know that we have correctly calculated the structure of the ball of radius n.

4.3.2 Detecting more than one end

As in the **ends** program, we calculate the ball of radius $n+\kappa$ and by theorem 2.1.12, we find the Γ-components of $\Gamma \setminus B^\circ_n$ by finding the components of $\overline{B}_{n+\kappa} \setminus B^\circ_n$.

Now by theorem 2.2.1 on page 52, G has more than one end if at least two components of $\Gamma \setminus B^\circ_n$ have elements of length $2n$ in them. Note that conversely, if G has more than one end, there is an $n \in \mathbb{N}$ such that $\Gamma \setminus B^\circ_n$ has at least two infinite components, which then necessarily contain elements

[13]For example, by running **fsaenumerate** to depth n on the word acceptor for this free group — free groups have a very simple automatic structure.

of length $2n$.

If we find more than one component of $\Gamma \setminus B^{\circ}_{n}$, we test this condition by constructing the ball of radius $2n$. If there are two distinct components of $\Gamma \setminus B^{\circ}_{n}$ containing elements of length $2n$, we stop, and conclude that by theorem 2.2.1 the group has more than one end. If not, we start again with a ball of larger radius.

If the group has more than one end, eventually this algorithm detects the fact.

4.3.3 Finding the number of ends

The algorithm for detecting endedness using only a solution to the word problem is less powerful than that using an automatic structure. The difference is in detecting finiteness and two-endedness.

If G is two-ended, then it is automatic, and so autgroup will calculate the automatic structure and 2end will determine that the group has two ends.

If, however, autgroup does not calculate an automatic structure for the group, we usually do not know whether this is because it stopped too soon or there is no automatic structure for G. In this case we cannot conclude that G does not have two ends. Also, if we find a ball with exactly two infinite components in its complement, we do not know whether this is because the group is two-ended or because the group has infinitely many ends and we need to remove a larger ball to find more than two infinite components.

Similarly, if the group is finite, eventually we can find this out either by calculating the whole Cayley graph or by calculating an automatic structure and finding no loops. However, if we have not succeeded in either of these it may simply be because we have not computed long enough, and we cannot conclude that our group is infinite.

Table 4.2 summarises the logical situation.

ends	when we detect this	when we refute this
0	always	only sometimes
1	never	only sometimes
2	always	only sometimes
∞	always	only sometimes

Table 4.2: To what extent we can determine the number of ends in a group given only a presentation and a solution to the word problem.

4.4 Theoretical: finding an explicit finite subgroup over which we split

If we have found a ball that disconnects the Cayley graph into more than one infinite component, we know that the group splits over a finite subgroup. In this section we show there is an algorithm finding such a subgroup explicitly.

We refer the reader to chapter 3 on page 64 for the definitions and results used in this section.

The method for finding the subgroup derives from Dunwoody's proof of Stallings' theorem using the Bergman norm. See section 3.1 on page 65. In our exposition of the proof we defined a sequence of infinite subsets of

G whose finite coboundaries have strictly decreasing Bergman norm. We used the fact that the lexicographic ordering with the Bergman norm gives a well ordering, to deduce that the sequence is eventually constant, and hence that the constant term is G-nested. G then splits over the stabiliser of the G-nested subset we obtain.

When computing we cannot handle an infinite number of vertices, but in section 3.2.1 on page 71, we proved that the stabiliser of a subset S of G is equal to the stabiliser of its oriented, finite coboundary, $\overrightarrow{\delta S}$.

The philosophy is that an oriented coboundary is equivalent to its defining set, and that we may calculate using the oriented coboundary rather than the infinite set.

4.4.1 Calculating with coboundaries

Recall the following definitions from section 3.5 on page 83:

Definition 4.4.1 (Γ-within) *Let* G *be a finitely presented group with Cayley graph* Γ. *Given a set* S *of elements of* G *and an edge* e *of* Γ, *we say that* e *is* Γ-within S *if it has both its vertices lying inside* S.

Definition 4.4.2 (Γ-connectedness) *Let* G *be a finitely presented group with Cayley graph* Γ. *We say that a set* S *of elements of* G *is* Γ-connected *if for every pair of vertices of* S, *there is an edge path between them which has each edge* Γ*-within* S.

A Γ-component *of* S *is a maximal* Γ*-connected subset of* S.

Lemma 4.4.3 (coboundaries of components)

Let S *be a subset of* G *with finite coboundary* δS, *and suppose* T *is a* Γ-*component of* S. *Then* $\delta T \subset \delta S$, *and* δT *is equal to those edges of* δS *which meet* T.

Proof

Suppose for contradiction that an edge e of δT is not in δS. One of the vertices incident with e is in $T \subset S$. Since e is not in δS, the other vertex, v, must also be in S. Since e *is* in δT, $v \notin T$. Thus $T \cup \{v\} \subset S$, but T was assumed to be a Γ-component of S, and this contradicts the maximality in the definition of Γ-component. Thus $\delta T \subset \delta S$.

Now let $e \in \delta T$. Then $e \in \delta S$, and one of the vertices of e is in T, so e meets T.

Conversely, suppose $e \in \delta S$ and e meets T. Then, one of the vertices, w say, of e, is in T. $T \subset S$, so $w \in S$. The other vertex, v, is not in S, so $v \notin T$. Thus $e = [v, w] \in \delta T$.

□

Lemma 4.4.4 (infinite Γ-components)

Let S *be a subset of* G *with a coboundary* δS *that lies inside* $\overline{B}_n$ *for some* $n \in \mathbb{N}$. *Suppose* T *is a* Γ-*component of* S. *Then* T *is infinite if and only if there is a vertex* v *of* T *at exactly distance* n *from* id_G *which lies in an infinite component of* $\Gamma \setminus B_n^\circ$.

Proof

By the previous lemma, $\delta T \subset \delta S$, and by hypothesis, $\delta S \subset \overline{B}_n$, so $\delta T \subset \overline{B}_n$. So no path in $\Gamma \setminus B_n^\circ$ can cross an edge of δT. Thus a component of $\Gamma \setminus B_n^\circ$ has all its vertices in T or none of them.

If there is a vertex v of T at exactly distance n from id_G which lies in an infinite component of $\Gamma \setminus B_n^\circ$, then all the vertices of this component are in T, so there are infinitely many vertices in T.

Conversely, if there are infinitely many vertices in T, only finitely many of them can lie in finite components of $\Gamma \setminus B_n^\circ$, so there is a vertex of T in an infinite component of $\Gamma \setminus B_n^\circ$. Then all the vertices of this component are in T. Every component of $\Gamma \setminus B_n^\circ$ meets S_n because Γ is path-connected, so there is a vertex v of T at exactly distance n from id_G which lies this infinite component of $\Gamma \setminus B_n^\circ$.

□

Algorithm 4.4.5 (finding an infinite Γ-connected subset)

Let G be a finitely presented group with Cayley graph Γ. Assume we can calculate the ball $\overline{B}_n$ of any radius n about id_G in Γ,[14] and that we can determine which components of $\Gamma \setminus B_n^\circ$ are infinite.

Suppose we are given an oriented finite coboundary $\overrightarrow{\delta S}$ of some infinite subset S of G, and suppose the edges of δS lie inside $\overline{B}_n$ for some $n \in \mathbb{N}$. Then we can calculate the coboundaries of the infinite Γ-components of S. (Note that being given the orientation on δS is equivalent to being given the

[14]This first condition is logically equivalent to having solvable word problem, but we phrase the hypothesis in the form we use it.

vertices of S incident with δS, and so the lemma applies whether we are given the orientation or the vertices.)

Let $\{V_1, \ldots, V_m\}$ be the Γ-components of S, and let κ be the connectivity constant of Γ. (See section 2.1 on page 33.)

1. Mark the edges of δS as temporarily *uncrossable*.[15]

2. Choose a vertex v of S incident with δS.

3. Perform a vertex depth-first search inside $\overline{B}_n$, starting at v and without crossing any uncrossable edges. Mark each vertex encountered as being in the same Γ-component as v.

4. If all vertices of $\overline{B}_{n+\kappa}$ are marked as in the same Γ-component as some vertex, proceed to the next step. Otherwise choose a new vertex v incident with δS and repeat the previous step. In this way we calculate the Γ-components of $S \cap \overline{B}_{n+\kappa}$. By theorem 2.1.10, if there were a path between two vertices of $S \cap \overline{B}_n$, there would also be a path between them inside $S \cap \overline{B}_{n+\kappa}$. Thus the intersection with $\overline{B}_n$ of the Γ-components of $S \cap \overline{B}_{n+\kappa}$ is the same as the intersection with $\overline{B}_n$ of the Γ-components $\{V_1, \ldots, V_l\}$ of S.

5. By lemma 4.4.4 on page 116, a Γ-component V_i is infinite if and only if it contains a vertex of S_n which is in an infinite component of $\Gamma \setminus B_n^\circ$.

[15]In the computer representation of Γ, this means marking pairs of adjacency links unusable. See section 4.2.1 on page 94.

Calculate which are the infinite components of $\Gamma \setminus B_n^\circ$, and examine the vertices of S_n to determine which of the V_i are infinite.

6. Pick one of the infinite components, V_i, of S, and find which edges of δS meet it. These are the edges of δV_i, by lemma 4.4.3 on page 116.

□

Similarly, given δS and δT, we can find the infinite Γ-components of the sets $S \cap T$, $S \cap T^*$, $S^* \cap T$, $S^* \cap T^*$:

Algorithm 4.4.6 (finding infinite Γ-components)

Let G be a finitely presented group with Cayley graph Γ. Assume we can calculate the ball $\overline{B}_n$ of any radius n about id_G in Γ, and that we can determine which components of $\Gamma \setminus B_n^\circ$ are infinite.

Suppose we are given the finite coboundaries δS and δT of some subsets S and T of G, and suppose the edges of both coboundaries lie inside $\overline{B}_n$ for some $n \in \mathbb{N}$. Then we can calculate the intersection with $\overline{B}_n$ of the infinite Γ-components of the sets $S \cap T$, $S \cap T^*$, $S^* \cap T$, $S^* \cap T^*$, and can calculate the coboundaries of these Γ-components.

Let $\{V_1, \ldots, V_m\}$ be the Γ-components of $S \cap T$, $S \cap T^*$, $S^* \cap T$, $S^* \cap T^*$, and let κ be the connectivity constant of Γ. (See section 2.1 on page 33.)

1. Mark the edges of δS and δT as temporarily *uncrossable.*

2. Choose a vertex v incident with δS or with δT.

3. Perform a vertex depth-first search inside $\overline{B}_n$, starting at v and without

crossing any uncrossable edges. Mark each vertex encountered as being in the same Γ-component as v.

4. If all vertices of $\overline{B}_{n+\kappa}$ are marked as in the same Γ-component as some vertex, proceed to the next step. Otherwise choose a new vertex v incident with δS or δT, and repeat the previous step. In this way we calculate the Γ-components of the intersections with $\overline{B}_{n+\kappa}$ of the sets $S \cap T, S \cap T^*, S^* \cap T, \ S^* \cap T^*$. By theorem 2.1.10, when we intersect these with $\overline{B}_n$ we obtain the intersection with $\overline{B}_n$ of the Γ-components of the sets $S \cap T, S \cap T^*, S^* \cap T, \ S^* \cap T^*$.

5. By lemma 4.4.4 on page 116, a Γ-component V_i is infinite if and only if it contains a vertex of S_n which is in an infinite component of $\Gamma \setminus B_n^\circ$. Calculate which are the infinite components of $\Gamma \setminus B_n^\circ$, and examine the vertices of S_n to determine which of the V_i are infinite.

6. For each of the infinite components, V_i, find which edges of δS and δT meet it. These are the edges of δV_i, by lemma 4.4.3 on page 116.

□

4.4.2 Testing whether a coboundary is G-nested

Lemma 4.4.7 (A test for non-nestedness)

Let S and T be subsets of G, and suppose that both S and S^* are Γ-connected. Then the pair (S, T) is not *nested if and only if one of the edges of δT is Γ-within S and another is Γ-within S^*.*

Proof

First we prove that if (S, T) is not a nested pair of subsets of G, then both such edges exist.

Suppose (S, T) is not a nested pair. Then the none of the sets $S \cap T$, $S \cap T^\star$, $S^\star \cap T$, nor $S^\star \cap T^\star$ is empty. Let $x_1 \in S \cap T$, $x_2 \in S \cap T^\star$, $y_1 \in S^\star \cap T$, and $y_2 \in S^\star \cap T^\star$. Since S is Γ-connected, there is an edge path between x_1 and x_2 whose edges are Γ-within S. There must be an edge of this edge path which has one vertex in T and the other in $T^\star$, so it is an edge of δT which is Γ-within S. Similarly, since $S^\star$ is Γ-connected, there is an edge path between y_1 and y_2 whose edges are Γ-within $S^\star$. There must be an edge of this edge path which has one vertex in T and the other in $T^\star$, so it is an edge of δT which is Γ-within $S^\star$. This concludes the proof of the first implication.

Now suppose that an edge e_1 of δT is Γ-within S and another edge e_2 of δT is Γ-within $S^\star$. Each edge has one vertex in T and the other in $T^\star$. The vertices of e_1 show that the sets $S \cap T$ and $S \cap T^\star$ are non-empty, and the vertices of e_2 show that the sets $S^\star \cap T$ and $S^\star \cap T^\star$ are non-empty. Thus (S, T) is not a nested pair.

□

Corollary 4.4.8

Let S *be a* Γ*-connected subset of* G *with* Γ*-connected complement, and finite coboundary* δS*. Let* $d = \mathrm{diam}(\delta S)$*. Then* S *is* not G*-nested if and only if there is an element* $g \in G \cap \overline{B}_{2d}$ *so that one of the edges of* $\delta(gS)$ *is* Γ*-within* S *and another is* Γ*-within* $S^\star$*.*

Proof

Let h be an element of G incident with δS, and let $T = h^{-1}S$. Then for any $g \in G$, the pair (S, gS) is nested if and only if the pair $(T, h^{-1}ghT)$ is nested. Since $\{(T, h^{-1}ghT) \mid g \in G\} = \{(T, gT) \mid g \in G\}$, S is G-nested if and only if T is G-nested. Thus, without loss of generality in our proof, we may assume that id_G is a vertex of an edge of $\overrightarrow{\delta S}$.

By the previous result, all we need show is that there is no element $g \in G$ outside $\overline{B}_{2d}$ for which there is an edge of $\delta(gS)$ Γ-within S and also an edge of $\delta(gS)$ Γ-within $S^\star$.

Now let $g \in G \setminus \overline{B}_{2d}$. Without loss of generality, we may now assume that $g \in S$, because for the case of $g \in S^\star$ we may swap the labels S and $S^\star$.

Assume then that id_G is a vertex of δS and $g \in S$ has length strictly greater than $2d$. Note that $d = \mathrm{diam}(\delta S) = \mathrm{diam}(\delta(gS))$, so every edge e of $\delta(gS)$ is at most distance d from g.

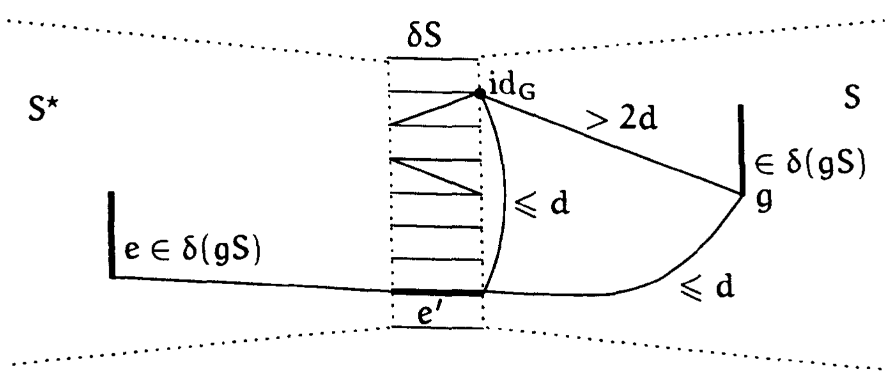

Diagram 4.4.1: Long elements cannot give rise to non-nestedness.

Assume for contradiction that there is an edge, e, of $\delta(gS)$ Γ-within $S^\star$. Take a geodesic edge path from it to the edge of $\delta(gS)$ incident with g. Since $g \in S$, one of the edges, e', of this edge path is in δS, so e' is in δS and is within distance d of g. e' and id_G are both parts of δS, so are within distance d of each other. By the triangle inequality, g and id_G are within distance $2d$ of each other, which contradicts our hypothesis that $g \notin \overline{B}_{2d}$.

□

To algorithmically test for G-nestedness, we need to be able to calculate the action of an element of G on subsets of the Cayley graph.

Algorithm 4.4.9 (G-action)

Let G be a finitely presented group with Cayley graph Γ. Suppose we can calculate the ball of any radius about id_G in Γ, together with a word in the generators representing each element of G within the ball.

Suppose we have calculated $\overline{B}_n$, and let $g \in \overline{B}_n$ be given. For any vertex h of $\overline{B}_n$, we can calculate gh. For any edge e of $\overline{B}_n$, we can calculate ge. For any finite set of edges or vertices, we can calculate its image under the action of the element g of G.

It is enough to demonstrate the algorithm for the action on vertices, because the action on a set of vertices, or an edge, or a set of edges, is determined trivially from the action on the vertices involved.

Let $x_1x_2\cdots x_p$ be a word representing h. We find gh by finding successively gx_1, gx_1x_2, ..., $gx_1x_2\cdots x_p$; to find $gx_1x_2\cdots x_{q+1}$ from $gx_1x_2\cdots x_q$, we move along the edge labelled by x_{q+1}.

Note that it is trivial to find the inverse of a word, so the algorithm applies equally well to calculating the action of g^{-1} on some subset of Γ.

□

Algorithm 4.4.10 (testing nestedness)

Let G be a finitely presented group with Cayley graph Γ. Suppose we can calculate the ball of any radius about id_G in Γ, together with a word in the generators representing each element of G within the ball.

Suppose we are given an oriented coboundary δS, where both S and S^* are Γ-connected, and suppose that we know an upper bound for the diameter of δS. Then we can check whether S is G-nested as follows:

Let d be the upper bound on $\mathrm{diam}(\delta S)$, and suppose we have already calculated $\overline{B}_n$ for some n such that $\delta S \subset \overline{B}_n$.

1. Pick a vertex h incident with one of the edges of δS. Translate δS under the action of h^{-1} so that id_G is incident with an edge of the translated coboundary. Replace S with $h^{-1}S$, so that now id_G is incident with δS.

2. Since $\overline{B}_n$ is connected, there is a path between any two elements of $S \cap \overline{B}_n$ which lies entirely within $\overline{B}_n$. Now let v be an arbitrary vertex of $S \cap \overline{B}_n$, and consider a path from it to some element of S incident with δS, staying inside $\overline{B}_n$. Even if this path does not always stay in S, at some point it meets *some* vertex of S incident with δS; a vertex depth-first search of $S \cap \overline{B}_n$ starting at this vertex will reach v:

   Find all the vertices inside the currently calculated ball that are in S, and mark them. Do this by taking each element v of S incident with an edge of δS, and performing a vertex depth-first search starting at v, without crossing any edge of δS. Mark each vertex met.

3. For each element $g \in \overline{B}_{2d}$, calculate the g-translate of δS, $g(\delta S) = \delta(gS)$.[16] Check whether there are two edges of δS, one of which is Γ-within S, the other of which is Γ-within $S^\star$. If so, stop, and conclude that S is not G-nested by lemma 4.4.7.

   If not, proceed to test other elements g of $\overline{B}_{2d}$. If each of the pairs (S, gS) is found to be nested, then stop and conclude, by corollary 4.4.8, that S is G-nested.

□

4.4.3 Finding a G-nested coboundary

Recall from definition 3.1.2 on page 66 that the Bergman norm $\mu(\delta S)$ of a coboundary δS is the sequence $(\mu_i(\delta S))_{i \in \mathbb{N}}$ where $\mu_i(\delta S)$ is the number of edge paths of length i which start in S and end in $S^\star$.

In the proof of Stallings' theorem[17] we take any infinite set with finite coboundary and infinite complement, and calculate a sequence of such sets with the aim of finding one which is G-nested. The Bergman norm of the terms of the sequence is decreasing while a G-nested set is not found. Since the norm is well-ordered [Ber68], eventually a nested set must be found. We want to be able to calculate the terms in such a sequence, so we need an algorithm which takes a non-G-nested infinite set with finite coboundary and

[16] This equality is proved in corollary 3.2.3 on page 71. We may need to calculate $\overline{B}_{n+i}$ for some i so that $\delta(gS) \subset \overline{B}_{n+i}$. In any case, we need calculate no more than $\overline{B}_{3d}$ because $\mathrm{diam}(\delta S) = d$ and g is taken to be inside $\overline{B}_{2d}$.

[17] Theorem 3.1.7 on page 68.

infinite complement, and calculates another infinite set with strictly smaller Bergman norm and infinite complement.

Lemma 4.4.11

Let S *be an infinite subset of* G *with finite coboundary and infinite complement. Then every infinite* Γ*-component,* T*, of* S *or of* $S^\star$ *has infinite complement, and has finite coboundary* $\delta T \subset \delta S$ *with* $\mu(T) \leqslant \mu(S)$.

Moreover, if S *is* Γ*-connected, and* T *is a* Γ*-component* of $S^\star$ *then in addition, both* T *and* $T^\star$ *are* Γ*-connected.*

Proof

For the first point, it is enough to prove the result for $T \subset S$, since $S^\star$ satisfies the same hypotheses as S.

Let T be an infinite Γ-component of S. Then since $T \subset S$, $S^\star \subset T^\star$, and so $T^\star$ is infinite.

Since T is a Γ-component of S, $\delta T \subset \delta S$ by lemma 4.4.3 on page 116.

If $\delta T = \delta S$, then $\mu(\delta T) = \mu(\delta S)$ i.e. $\mu(T) = \mu(S)$. Otherwise, there are more edges in δS than in δT, so $\mu_1(\delta T) < \mu_1(\delta S)$, so $\mu(T) < \mu(S)$.

Thus T is infinite, has infinite complement, and $\mu(T) \leqslant \mu(S)$.

Now we prove the second point.

Suppose additionally that S is Γ-connected and that $T \subset S^\star$. All that remains is to prove is that $T^\star$ is Γ-connected, since T is itself a Γ-component, and thus Γ-connected.

Let $\{V_1, \ldots, V_n\}$ be the Γ-components of $S^\star$. Without loss of generality, $T = V_1$. For each i, there is an edge between V_i and S (by path-connectivity of the Cayley graph and the maximality in the definition of Γ-component). Thus $V_i \cup S$ is Γ-connected, and so $T^\star = S \cup V_2 \cup \cdots \cup V_n$ is Γ-connected.

Thus each of T and $T^\star$ is infinite and Γ-connected, and $\mu(T) \leqslant \mu(S)$.

□

Corollary 4.4.12

Let S be an infinite subset of G with finite coboundary and infinite complement. Suppose the pair (S, gS) is not nested. Then one of the Γ-components of one of the sets $S \cap gS$, $S \cap gS^\star$, $S^\star \cap gS$, or $S^\star \cap gS^\star$ has strictly smaller Bergman norm than S, and is infinite with infinite complement.

Proof

Dunwoody's result (lemma 3.1.3 on page 67 and its corollary 3.1.4 on page 67) says that one of $S \cap gS$, $S \cap gS^\star$, $S^\star \cap gS$, or $S^\star \cap gS^\star$ is infinite with infinite complement, and has strictly smaller Bergman norm than S. Suppose, without loss of generality, that $S \cap gS$ does.

Not all of the Γ-connecteds of $S \cap gS$ can be finite, because otherwise there would be infinitely many of them, and $S \cap gS$ would not have a finite coboundary.[18] Thus there is an infinite Γ-component of $S \cap gS$. The result follows from the previous lemma.

□

[18]No two Γ-components can share the same edge in their coboundary: Lemma 4.4.3 on page 116 says that if T is a Γ-component of $S \cap gS$, then δT is those edges of $\delta(S \cap gS)$ which meet T. Thus if T and T' were Γ-components of $S \cap gS$ that shared an edge of $\delta(S \cap gS)$, they would have a vertex in common, so would be the same.

Algorithm 4.4.13 (calculating $\mu_n(\delta S)$)

Let G be a finitely presented group with Cayley graph Γ. Suppose we can calculate the ball of any radius about id_G in Γ. Suppose we are given a coboundary δS, and suppose that we know an upper bound for the diameter of δS. Then we can calculate the nth entry in the Bergman norm of δS, $\mu_n(\delta S)$ as follows:

1. Find all vertices at distance at most n from S, but which lie in $S^\star$.[19] (Increase the radius of the calculated ball about the identity if necessary.) Call these vertices *startpoints.*
2. Mark all the other vertices of $S^\star$ that have been calculated as temporarily *unreachable*. No path of length n from $S^\star$ to S meets an unreachable vertex.
3. From each startpoint, perform a path depth-first search to depth n, avoiding unreachable vertices. Each time a path from an element of $S^\star$ ends in S, add one to the count.

□

Algorithm 4.4.14 (finding a nested coboundary)

Let G be a finitely presented group with Cayley graph Γ. Assume we can calculate the ball of any radius about id_G in Γ, together with a word in the

[19] δS and its Bergman norm are independent of which is S and which is $S^\star$, so if necessary, pick one at random and call it S.

generators representing each element of G within the ball. Assume also that we can determine which components of $\Gamma \setminus B_n^\circ$ are infinite.

Suppose we are given a finite coboundary δS of some infinite subset of G which has infinite complement. Then we may calculate a finite G-nested coboundary, δT for some $T \subset G$.

1. S might not be Γ-connected. Use algorithm 4.4.5 to find T_0, an infinite Γ-component of S. Use it again to find T_1, an infinite Γ-component of $T_0^\star$. By lemma 4.4.11 on page 126, each of T_1 and $T_1^\star$ is infinite and Γ-connected, and $\mu(T_1) \leqslant \mu(T_0) \leqslant \mu(S)$.

2. Given δT_i, calculate an upper bound for its diameter by finding paths between each pair of vertices incident with it.[20]

3. Since T_i and $T_i^\star$ are both Γ-connected, we may use algorithm 4.4.10 on page 124 to check whether δT_i is G-nested. If so, stop. If not, we have found $g \in G$ such that the pair (T_i, gT_i) is not nested. (By corollary 4.4.12 on page 127, one of the Γ-components of one of the sets $T_i \cap gT_i$, $T_i \cap gT_i^\star$, $T_i^\star \cap gT_i$, or $T_i^\star \cap gT_i^\star$ has strictly smaller Bergman norm, and is infinite with infinite complement.)

4. Temporarily mark the edges of δT_i and $\delta(gT_i)$ as uncrossable, in order to calculate which Γ-components of $T_i \cap gT_i$, $T_i \cap gT_i^\star$, $T_i^\star \cap gT_i$, and

[20] If we have calculated the ball of sufficient radius, we will find the geodesics among the paths between pairs of vertices, and will have calculated the actual diameter. This is an unimportant point theoretically, but it may improve the speed of the algorithm checking for G-nestedness.

$T_i^\star \cap gT_i^\star$ are infinite with infinite complement. (Use algorithm 4.4.6 on page 119.)

5. Using algorithm 4.4.13 on page 128, calculate the nth element in the Bergman norm of each of the Γ-components, for increasingly large n. Find one of them, U, say, which has strictly smaller Bergman norm.[21] Find a Γ-component of $U^\star$, and call it T_{i+1}. Since U is Γ-connected and infinite with infinite complement, by lemma 4.4.11 on page 126, we know that T_{i+1} has a finite coboundary with $\mu(T_{i+1}) \leqslant \mu(U) < \mu(T_i)$, and that both T_{i+1} and its complement are Γ-connected and infinite.

6. Repeat steps 2 to 5 until T_i is nested. This occurs after finitely many iterations because the Bergman norm and lexicographic order gives a well-ordering [Ber68].

□

4.4.4 Finding a splitting

Algorithm 4.4.15 (finding the stabiliser)

Let G be a finitely presented group with Cayley graph Γ. Assume we can calculate the ball of any radius about id_G in Γ, together with a word in the generators representing each element of G within the ball. Assume also that we can determine which components of $\Gamma \setminus B_n^\circ$ are infinite.

[21] Since one of them has strictly smaller Bergman norm than T_i, we know that for some n, the nth element of the norm must be strictly smaller. We do not know in advance for which $n \in \mathbb{N}$ this will first occur. See section 3.4 on page 77.

Suppose we are given a finite G-nested coboundary δS, of some infinite Γ-connected subset of G which has infinite Γ-connected complement. Then we can calculate $\mathrm{stab}(\overrightarrow{\delta S})$.

Let $\{v_1, \dots, v_n\}$ be the set of vertices of S that are incident with δS (i.e. the terminal vertices of $\overrightarrow{\delta S}$), and let $\{w_1, \dots, w_n\}$ be the set of vertices of $S^\star$ that are incident with δS (the initial vertices of $\overrightarrow{\delta S}$).

1. Find the sets $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_n\}$. If we are not given them, we may calculate them using algorithm 4.4.5 on page 117, since both S and $S^\star$ are infinite and Γ-connected. It does not matter which is which, because $\mathrm{stab}(\overrightarrow{\delta S}) = \mathrm{stab}(\overrightarrow{\delta S^\star})$ (corollary 3.2.6 on page 73).

2. Note that the stabiliser of $\overrightarrow{\delta S}$ consists of exactly those elements that permute the vertices $\{v_1, \dots, v_n\}$, and also permute the vertices $\{w_1, \dots, w_n\}$.

   Thus every element of $\mathrm{stab}(\overrightarrow{\delta S})$ is of the form $v_i^{-1}v_j$. Calculate these elements, using algorithm 4.4.9 on page 123. Call them $\{g_1, \dots, g_m\}$.

3. Since $\mathrm{stab}(\overrightarrow{\delta S}) \subset \{g_1, \dots, g_m\}$, all we need to do is find which of the g_i do not permute the initial vertices $\{w_1, \dots, w_n\}$ or do not permute the terminal vertices $\{v_1, \dots, v_n\}$.

   For each g_i, and for each j, calculate $g_i v_j$ and $g_i w_j$, checking whether they are in $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_n\}$ respectively. This is the case if and only if $g_i \in \mathrm{stab}(\overrightarrow{\delta S})$.

□

We may piece these algorithms together to find a subgroup over which G splits.

Theorem 4.4.16

Let G *be a finitely presented group with Cayley graph* Γ. *Assume we can calculate the ball of any radius about* id_G *in* Γ, *together with a word in the generators representing each element of* G *within the ball. Assume also that we can determine which components of* $\Gamma \setminus B_n^\circ$ *are infinite, for all* $n \in \mathbb{N}$.

Suppose we have found some $n \in \mathbb{N}$ *such that the ball of radius* n *disconnects the Cayley graph into more than one infinite component. Then we can calculate a finite subgroup over which* G *splits.*

Proof

1. Find an infinite component of $\Gamma \setminus B_n^\circ$, and denote the set of elements of G inside it by S. Calculate its (finite) coboundary, by taking the set of edges with one vertex in B_n° and the other in S. $S^\star$ is infinite, since it contains the vertices of another infinite component of $\Gamma \setminus B_n^\circ$.

2. Use algorithm 4.4.14 on page 128 to find a G-nested coboundary, δT.

3. Use algorithm 4.4.15 on page 130 to calculate $H = \mathrm{stab}(\overrightarrow{\delta T})$. H is finite because $H \subset \mathrm{stab}(\delta T)$ and δT is finite. G splits over H by Stallings' theorem 3.1.7 on page 68 (via theorem II.1.8 and I.4.1 of [DD89]).

Note that we can output H in the form of a set of vertices of Γ, or as a list of words in the generators of G.

□

Note that the theorem is certainly not true for arbitrary finitely generated groups, just those in which we may calculate balls and infinite components.

Chapter 5

Gerasimov's algorithm to detect endedness

Recently, Gerasimov provided an algorithm which detects one-endedness in a hyperbolic group. The author finds Gerasimov's paper [Ger] hard to follow. At a meeting between Delzant, Dunwoody and Epstein in Strasbourg, Delzant gave a different exposition of Gerasimov's proof. This representation of Gerasimov's proof was in turn explained to the author by Epstein, and this chapter attempts to record it.

Thus this chapter is expository in nature, with only the definition and test for exterior paths *outside a given radius* contributed by the author. Thus credit for the mathematics lies elsewhere, but any errors and poor explanations are entirely the author's own work.

We do not use the standard notation for the boundary of a hyperbolic group; instead we reserve the symbol for a related construction for arbitrary finitely generated groups described in section 6.2.

5.1 Theoretical background to the algorithm

In their paper *The Boundary of Negatively Curved Groups* [BM91], Bestvina and Mess prove that if a hyperbolic group has one end, and satisfies a combinatorial condition (which later we define, and call the *exterior paths condition for hyperbolic groups*), then its boundary is locally path connected. They also prove that if the hyperbolic group does not satisfy the exterior paths condition then its boundary has a global cut point.

Bowditch [Bow96] and Swarup [Swa96] have both proved that the boundary of a one-ended hyperbolic group has no global cut point. Thus every one-ended hyperbolic group has locally connected boundary.

Infinitely ended hyperbolic groups do not satisfy the hyperbolic exterior paths condition: The result of Bestvina and Mess does not use the hypothesis that the group is one ended, so the hyperbolic exterior paths condition implies locally connected boundary. Infinitely ended hyperbolic groups do not have locally connected boundary.

Thus we may detect whether a hyperbolic group has one end by testing whether it satisfies the hyperbolic exterior paths condition. Gerasimov [Ger] defines an algorithm whereby a sequence of simplicial complexes is calculated. If one of the complexes far enough along in the sequence is connected and satisfies a certain condition, the group has one end.

To explain Gerasimov's result, we outline below a method for testing the exterior paths condition directly.

5.2 An interpretation of Gerasimov's algorithm

As outlined in the previous section, it is enough to prove that we may reliably detect the exterior paths condition if it is present.

5.2.1 The exterior paths condition

Bestvina and Mess define a condition on the Rips complex,[1] which they call ($\ddagger_M$). We make the same definition but apply it to the Cayley graph and call it *exterior paths.*

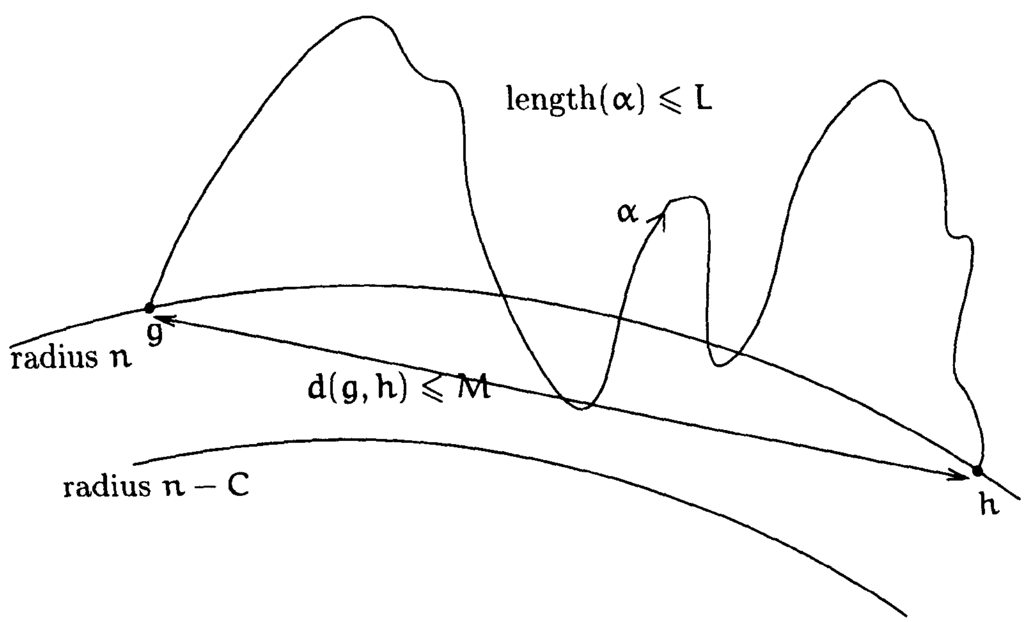

Diagram 5.2.1: Exterior paths

Definition 5.2.1 *Let* G *be a group with an upper bound,* C*, on the distance of an arbitrary element to a geodesic ray (i.e. infinite geodesic) starting at*

[1] A definition can be found at the beginning of [BM91].

id_G. *Let* M *and* L *be natural numbers.*

We say that G *has* (M, L)-exterior-paths *if* $\forall n \in \mathbb{N}$, *and* $\forall g, h \in S_n$, $d(g, h) \leqslant M$ *implies that there is a path,* $\alpha : g \simeq h$ *of length at most* L *such that* α *lies entirely outside the (open) ball of radius* $n - C$, B°_{n-C}. *Such a path is called an* exterior path.

Sometimes we say G has exterior paths everywhere.

Bestvina and Mess only deal with the above case, but it may be that there are only exterior paths between elements far away from id_G. In this case, we use the phrase 'outside radius N', giving the following definition. Note that detecting exterior paths outside some radius is irrelevant to Gerasimov's algorithm, but we will use it in the more generalised setting of chapter 6.

Definition 5.2.2 *Let* G *be a group with an upper bound,* C, *on the distance of an arbitrary element to a geodesic ray starting at* id_G. *Let* M *and* L *be natural numbers.*

We say that a group has (M, L)-exterior-paths outside radius N *if* $\forall n \in \mathbb{N}$ *with* $n > N$, *and* $\forall g, h \in S_n$, $d(g, h) \leqslant M$ *implies that there is a path,* $\alpha : g \simeq h$ *of length at most* L *such that* α *lies entirely outside the (open) ball of radius* $n - C$, B°_{n-C}.

Note that G has (M, L)-exterior-paths if and only of it has (M, L)-exterior-paths outside radius 0, because the case when g and h are id_G is trivial; indeed, the case when $n \leqslant C$ is always trivial.

Definition 5.2.3 (The exterior paths condition for hyperbolic groups) *Let* G *now be a hyperbolic group, and as always, denote its Cayley graph by*

Γ. *Then there is a constant,* C, *depending on the constant of hyperbolicity such that every point of* G *is within distance* C *from a geodesic ray starting at* id_G.

Let δ' *be the constant of hyperbolicity such that any point on any edge of a geodesic triangle is at most distance* δ' *from the union of the other two sides.*

We say that G *satisfies the* exterior paths condition for hyperbolic groups, *or the* hyperbolic exterior paths condition *if for some* $M > 6C + 2\delta' + 3$, *and some* $L \in \mathbb{N}$, G *has* (M, L)*-exterior-paths (everywhere).*

(Bestvina and Mess prove that if G is hyperbolic and satisfies the exterior paths condition for hyperbolic groups, then the boundary of G is locally path connected.)

5.2.2 Testing for exterior paths using the automatic structure

Here we follow the proof of Theorem 6.2 of [EIFZ96]. Epstein, Iano-Fletcher and Zwick prove that the growth function for the number of embeddings of some graph into the Cayley graph of an automatic group is rational, and prove this by showing there is a finite state automaton which accepts tuples of words describing the image of the vertices of the graph we wish to embed.

Let G be an automatic group with all accepted words geodesic.

First we re-examine the automatic structure. Alter it so that it accepts unique word representatives for each group element. We may replace the general multiplier, GM, with a set of multiplier automata, M_a, one for each generator a of G. It is clear that for a given generator, a, these may be calculated from GM by choosing the subset of the accept states corresponding to a. In practice there may be a much simpler automaton accepting the same pairs of words, and size is important for the complexity of the following process.

By convention, we do not include specific fail states in finite state automata; failure occurs when there is no transition corresponding to the next symbol on the input string, or when the input string terminates when we are in a non-accept state. With this convention, we may remove all states from which it is impossible to reach an accept state.

Algorithm 5.2.4 (ext.paths)

Let G be an automatic group with all accepted words geodesic. Then there is an algorithm, *ext.paths*, that determines whether there exist M and L such that G has (M, L)-exterior-paths or (M, L)-exterior-paths outside some radius N, or G does not have (M, L)-exterior-paths at all.

1. For some $m \in \mathbb{N}$, calculate $\overline{B}_m$. (Use the algorithm from section 4.2.2.1 on page 96.)

2. Calculate a spanning tree for $\overline{B}_m$. Call it T_m. Number the vertices of T_m in some order, with the identity element as vertex number 1.

3. Now M_a accepts words w_1 and w_2 such that $w_1 a = w_2$. M_b accepts words w_1 and w_3 such that $w_1 b = w_3$. By theorem 1.4.6 of [ECH+92], we may construct an automaton which accepts triples (w_1, w_2, w_3)

where $w_1 a = w_2$ and $w_1 b = w_3$. It may be necessary to add extra padding symbols ($) to one or more of the words, to adjust for the different lengths. For example, the actual words accepted could be $w_1\$$, w_2, and $w_3\$\$$.

Similarly, the tree T_m gives us the relationships between pairs of words which define a ball of radius m about any vertex of the Cayley graph. Using the construction in theorem 1.4.6 of [ECH+92], we may construct an automaton which accepts $|\overline{B}_m|$-tuples, $(w_1, w_2, \ldots, w_{|\overline{B}_m|})$, of words which together define the ball $\overline{B}_m(w_1)$. Call this automaton A_m.

4. The number of padding symbols $ at the end of one of the words, w_i in the tuple, compared with the number at the end of w_1, gives the relative distance of w_i and w_1 from the identity, since we are assuming that accepted words are geodesic. In this step we alter A_m so that we can deduce from an accept state the relative lengths of the accepted words.

   No padding occurs inside pruned(A_m), because if it did, either padding occurs in the middle of an accepted word, which is false, or the automaton accepts a set of words including some w_i and w_j arbitrarily far apart, which is also false, because the accepted tuples of words are pairwise at most distance $2m$ apart.

   Once we leave pruned(A_m) there are only finitely many states we may pass through. Alter A_m to form $\widehat{A_m}$ by duplicating states and paths if necessary, so that $\widehat{A_m} \setminus \text{pruned}(\widehat{A_m})$ is a forest of trees.

   Each accept state now has a unique sequence of previous states and transitions in which there is padding. Looking back along this edge path we can see how much padding each word has before it is accepted, and may calculate its distance from id_G relative to w_1.

5. For each accept state in $\widehat{A_m}$, create a copy of $\overline{B}_m$, and use T_m and the relative distances calculated from $\widehat{A_m}$ to label this ball with relative distances from id_G. Call these labelled balls *types* of balls.

   Given $g \in G$ with $g \in S_n$ for some n, the ball $\overline{B}_m(g)$ is of one of the finitely many types calculated from the accept states of $\widehat{A_m}$. Using the relative distances from the identity, we can calculate what $S_n \cap \overline{B}_m(g)$ and $\left(\Gamma \setminus B^\circ_{n-C}(\mathrm{id}_G)\right) \cap \overline{B}_m(g)$ are. If $m \geqslant \max\{M, L\}$, we can test the existence of (M, L)-exterior-paths between g and elements in $\overline{B}_M(g)$.

   Thus by checking each of the ball types calculated from $\widehat{A_m}$, we can check exterior paths everywhere by checking exterior paths on these finitely many ball types.

6. It may be that exterior paths fails for some accept states of $\widehat{A_m}$, but that these accept states are not reachable via a loop in $\widehat{A_m}$. This means that these ball types can only occur at a finite distance from id_G. If all the other ball types have exterior paths, G has exterior paths outside some radius. We may calculate this radius by finding the maximum distance from the start state to accept states which fail the exterior paths test.

Test for (M, L)-exterior-paths as above, for larger and larger values of M and L. If G has exterior paths eventually this algorithm detects the fact.

□

5.2.2.1 Possible efficiency gain in hyperbolic groups

In section 8 of [EIFZ96], there is a proof that in a hyperbolic group there is a fixed polynomial such that the growth in the number of injections of any finite graph into the Cayley graph of a hyperbolic group can always be expressed with this denominator.

In the proof, a large automaton, H, is proved to exist with the property that injections of large finite graphs into the Cayley graph correspond to travellable words of this automaton in a finite-to-one way. It may be that H could be used to calculate very large ball types faster than the method above.

It is not explained in [EIFZ96] how to calculate H, but a sketch is given below. This section is designed to be read in conjunction with [EIFZ96], and makes little sense without it. Crucially, the definition and properties of H are omitted.

1. Alter the multiplier automata so that each state is labelled with the pair of states the current words determine in the word acceptor. This will increase the number of states in the multiplier automata.
2. Calculate all accepted words of length up to $2\delta'$, and their δ'-neighbourhoods.
3. Calculate a spanning tree for each such neighbourhood, then combine multiplier automata according to the relationships defined by the edges of the tree, and calculate the relative distances from id_G as we did before. Call these *tree* automata.

Each accept state of each tree automaton is a state in H. Keep the labels from the multipliers (originally from the word acceptor). A state in H is labelled by a δ'-neighbourhood of a word of length $2\delta'$, each vertex of which is labelled by its relative distance to id_G compared to the end of the word, and by its accept state in the word acceptor.

4. Since H is an automaton, there is a radius depending on the number of states of H, inside which all states of H must occur. Calculate a ball of sufficient radius, and locate copies of the states of H in this ball.

5. Calculate the transitions (edges) of H as follows: Each state of H has copy of it inside the ball, and a word of length $2\delta'$. For each generator a find the δ'-neighbourhood of the word determined by adding the generator and shortening the word back to length $2\delta'$. Find accept states on this new neighbourhood, and actual distances of elements from id_G, using the latter to calculate relative distances. Find the state of H with this label, and connect the original state to this new state with a directed edge, labelled by the generator a.

Once we have constructed the automaton H, we may follow the proof of lemma 8.2 of [EIFZ96] to construct the ball types for balls of radius m. For each state, we find all forward paths from the vertices of the state's label up to distance $2m + 2\delta'$, and from this we may determine the relative distances to the identity.

This algorithm for hyperbolic groups is not necessary from a theoretical point of view, and would only be useful if we needed to consider large values of M and L when testing for (M, L)-exterior-paths.

Chapter 6

Towards detecting one-endedness in automatic groups

6.1 Introduction

In this chapter we prove that if the algorithm described in chapter 5 is run on an automatic group G which has geodesic accepted words,[1] and terminates, having found that G has (M, L)-exterior-paths for large enough M, then G has at most two ends.

We do not prove that one-ended automatic groups with geodesic accepted words satisfy the exterior paths condition, so we stop short of proving that

[1]This means that the set of accepted words is a subset of the set of geodesic words, not that every geodesic is accepted — the latter condition is stronger, and is equivalent to G being hyperbolic. See [Pap95] for a proof.

endedness is decidable for such groups.

To show that the exterior paths condition implies having only finitely many ends, we borrow Epstein's construction of a boundary for an arbitrary graph, and relate endedness and exterior paths to connectivity properties of this boundary, following Bestvina and Mess [BM91]. In particular, we show that having (M, L)-exterior-paths for large enough M implies semi-local path-connectedness in ∂G.[2]

ends	algorithm detecting this	exterior paths condition?
0	no loops in word acceptor	trivially
1	*possibly* ext.paths	perhaps
2	2end (see [Bil96])	outside some radius
∞	ends	no

Table 6.1: Ends and algorithms detecting them in automatic groups which have geodesic accepted words.

6.2 The boundary of a Cayley Graph

The definition given here is essentially the same as that explained to the author by Epstein. Epstein's definition will appear in a forthcoming article.

We define the boundary by first defining a notion of distance on the set of geodesic rays from the identity, then adjusting the distance so that it satisfies the triangle inequality. This gives us a set with a pseudometric. Finally we

[2]In fact, we prove the slightly stronger condition of uniform path-connectedness. See definition 6.3.5 on page 154.

identify points of zero distance apart, giving a metric space which we call the boundary of the Cayley graph.

The metric on the boundary depends on the choice of generating set for the Cayley graph, and the choice of a base for exponentiation, and we do not prove that the boundary is a topological invariant of the group; indeed, it might not be topologically invariant under change of generators. We should refer to this boundary as $\partial_b\Gamma$, or the boundary of the Cayley graph with base b, but sometimes we will abuse terminology and call it the boundary of the group, ∂G.

6.2.1 The naive distance on the boundary

Fix $b \geqslant 2$. Let $\widehat{\partial G}$ be the set of geodesic rays from id_G. The *naive distance* between two geodesic rays is small when there is a geodesic between them that stays outside a ball of large radius.

Definition 6.2.1 *Let* $r, r' \in \widehat{\partial G}$ *be two geodesic rays from* id_G. *Define*

$$s_m(r, r') = \max\{d(\operatorname{im}\alpha, \mathrm{id}_G) \mid \alpha \textit{ is a geodesic from } r(m) \textit{ to } r'(m)\}$$

(There are only finitely many geodesics between two points in the Cayley graph because it is locally finite.)

We take the naive distance *between* r *and* r' *to be*

$$\delta_b(r, r') = \inf_{m \in \mathbb{N}} b^{-s_m(r,r')}$$

Note that $0 \leqslant \delta(r, r') \leqslant 1$. Roughly speaking, points are close if you can get between them far away from the identity. Indeed:

Corollary 6.2.2

$\delta(r, r') < \frac{1}{b^n}$ if and only if there is an $m \in \mathbb{N}$ such that $s_m(r, r') > n$. But this, in turn, is true if and only if for some $m \in \mathbb{N}$ there is a geodesic $\alpha : r(m) \simeq r'(m)$ which lies wholly outside the closed ball of radius n, $\overline{B}_n(\mathrm{id}_G)$.

□

Note that this also holds for negative n; in this case, for all $r, r' \in \widehat{\partial G}$, $\delta(r, r') \leqslant 1 < b^{-n}$. Also $\overline{B}_n(\mathrm{id}_G) = \emptyset$, so trivially, for any $r, r' \in \widehat{\partial G}$, there is an $m = 1 \in \mathbb{N}$ and a geodesic $\alpha : r(1) \simeq r'(1)$ outside $\emptyset$.

6.2.2 The pseudometric on the boundary

Now to make the triangle inequality hold, we define the actual distance between r and r':

$$d_b(r, r') = \inf\left\{ \sum_{i=1}^{l} \delta_b(a_{i-1}, a_i) \;\middle|\; \begin{array}{l} (a_i)_{i=0}^{l} \text{ is a finite sequence of points in } \widehat{\partial G} \\ \text{with } a_0 = r \text{ and } a_l = r' \end{array} \right\}$$

For convenience, we denote the set over which we take the infimum as $D(r, r')$. Note that since $\delta(r, r') \leqslant 1$, also $d_b(r, r') \leqslant 1$.

Proposition 6.2.3 ($d_b(\_, \_)$ is a pseudometric)

Let r, r' and r'' be elements of $\widehat{\partial G}$. Then

- $d_b(r, r') \geqslant 0$
- $d_b(r, r') = d_b(r', r)$
- $d_b(r, r') + d_b(r', r'') \geqslant d_b(r, r'')$

Proof

Clearly, $0 \leqslant d_b(r,r') \leqslant \delta_b(r,r')$, because the infimum of a set of non-negative numbers is non-negative.

Now $s_m(r,r') = s_m(r',r)$, so $\delta_b(r,r') = \delta_b(r',r)$ and thus $d_b(r,r') = d_b(r',r)$, because if two sets are the same, so is the infimum.

If we define addition of sets of reals by $X+Y = \{x+y \mid x \in X, y \in Y\}$ then $\inf X + \inf Y = \inf(X+Y)$. Thus $d_b(r,r') + d_b(r',r'') = \inf(D(r,r') + D(r',r'')) = \inf A$, where

$$A = \left\{ \sum_{i=1}^{l} \delta_b(a_{i-1},a_i) + \sum_{i=1}^{l'} \delta_b(b_{i-1},b_i) \;\middle|\; \begin{array}{l} (a_i)_{i=0}^{l} \text{ and } (b_i)_{i=0}^{l'} \text{ are finite sequences of points in } \widehat{\partial G} \\ \text{with } a_0 = r,\ a_l = r',\ b_0 = r' \text{ and } b_{l'} = r'' \end{array} \right\}.$$

But since any finite sequence between r and r' followed by a finite sequence between r' and r'' is essentially a finite sequence between r and r'' via r', we have that $A \subset D(r,r'')$. Thus $\inf A \geqslant \inf D(r,r'')$ i.e. $d_b(r,r') + d_b(r',r'') \geqslant d_b(r,r'')$

□

6.2.3 Definition of ∂G

$\widehat{\partial G}$ is a set with a pseudometric, and so we may construct the quotient metric space, $\partial_b G$ — points of $\partial_b G$ are equivalence classes of elements of $\widehat{\partial G}$, where two elements of $\widehat{\partial G}$ are equivalent iff they are distance zero from each other in the pseudometric. (The quotient metric is well defined because the triangle inequality holds for the pseudometric. Thus the pseudometric axioms hold in the quotient set, and so it really is a metric space.)

We will frequently abuse notation and use a geodesic ray as if it were an element of $\partial_b G$, and also denote the metric on $\partial_b G$ as $d_b(\_,\_)$. If we are not interested in which base b for exponentiation we are using, or it is clear from the context, we drop b from the notation and call the boundary ∂G and the metric $d(\_,\_)$. Note that we have not proven here that $\partial_b G$ is independent in any sense of the choice of b.

Example 6.2.4

Let $G = \langle x, y \mid xyx^{-1}y^{-1}, y^4 \rangle \cong \mathbb{Z} \times \mathbb{Z}_4$.

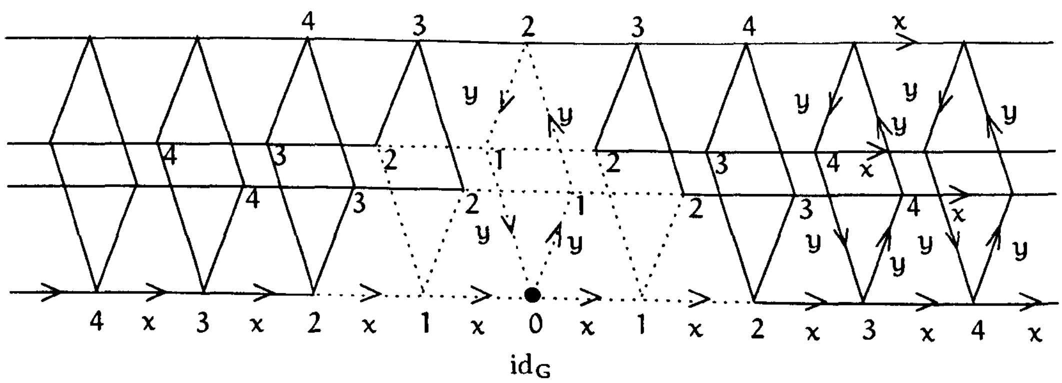

Diagram 6.2.1: $\mathbb{Z} \times \mathbb{Z}_4$

There are many geodesic rays in Γ, but only two elements of ∂G; any geodesic ray terminating with x^∞ is at distance 0 from yyx^∞, and any geodesic ray terminating with $x^{-\infty}$ is at distance 0 from $yyx^{-\infty}$. Since the power of y in a geodesic ray is limited to be between -2 and 2, these are the only two cases.

We pick the representatives $r = yyx^\infty$ and $r' = yyx^{-\infty}$ because they are the closest pair of representatives.

Since a geodesic between $r(m)$ and $r'(m)$ has distance exactly 2 from id_G, except when $m < 2$, when it is closer, we have $\inf_{m \in \mathbb{N}} b^{-s_m(r,r')} = b^{-2}$. Since these are the closest pair of representatives for the boundary points, and there are no other boundary points, we have $d_b(r, r') = b^{-2}$.

Thus $\partial_b G$ is a two-point metric space of diameter b^{-2}.

6.3 Properties of ∂G

Our motivation in using this boundary is to be able to relate the number of ends of the group to connectivity properties in ∂G, and in turn, to be able to detect these connectivity properties by means of an algorithm.

The following table relates the connectivity properties of the Gromov boundary of *hyperbolic* groups to the number of ends of the group.

ends	boundary	connected?	locally connected?
0	$\emptyset$	trivially	trivially
1	?	yes	yes[3]
2	2 points	no	yes
∞	?	no	no

Table 6.2: Ends versus boundary connectivity in *hyperbolic* groups

See [GdlH90] for a definition and discussion of the boundary of hyperbolic groups.

[3]The theorem that one-ended hyperbolic groups have locally connected boundary is a deep result, and has been proved both by Brian Bowditch [Bow96] and by Gadde Swarup [Swa96].

We will prove that if G has infinitely many ends, its boundary is not uniformly path-connected,[4] which is the equivalent of putting a "no" in the bottom right hand corner of a corresponding table for arbitrary finitely generated groups.

We showed in section 4.2.4 on page 106 that the cases of zero and two ends can be decided for automatic groups. If we have ruled these two cases out, and we find that ∂G is uniformly path-connected, then we know that G has one end. This does not mean we can necessarily *always* detect one-endedness in G — we have not proved that one-endedness implies uniform path-connectedness of the boundary. Neither have we proved that uniform path-connectedness implies that G has exterior paths.

Throughout section 6.3 we assume only that G is finitely generated, so that its Cayley graph, Γ, is locally finite.

6.3.1 Shadows of patches are open and closed

Let $n \in \mathbb{N}$, and consider the open ball B_n°. Recall from section 2.2.1 on page 52 that a patch at distance n from id_G is the intersection of the sphere S_n with one of the components of $\Gamma \setminus B_n^\circ$.

Definition 6.3.1 (shadows of patches) *Fix $n \in \mathbb{N}$. Let P be one of the patches at distance n around id_G. By lemma 2.2.3 on page 53, a geodesic ray from the identity passes through exactly one patch about id_G, so P determines a set of geodesic rays which we call the* shadow *of P, or* shadow(P). *Just as the patches at radius n in G partition the sphere S_n, the set of shadows of*

[4]See definition 6.3.5 on page 154.

patches at radius n *partitions* $\widehat{\partial G}$.

Technically speaking, shadow(P) *is a subset of* $\widehat{\partial G}$, *but we shall again abuse notation and use* shadow(P) *to denote the image of this set in* ∂G *as well.*[5]

Lemma 6.3.2

Let r *and* r' *be geodesic rays from* id_G, *and let* P *and* P' *be patches at distance* n *from* id_G. *Suppose that* r *passes through the patch* P *and* r' *passes through* P'. *If* $P \neq P'$ *then* $d_b(r, r') \geqslant \frac{1}{b^n}$.

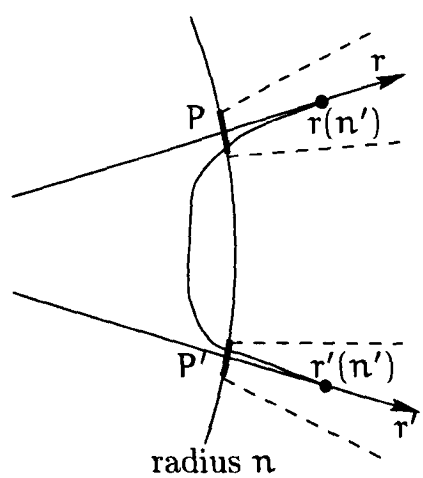

Diagram 6.3.1: Shadows are separated.

Proof

For $n' \geqslant n$, $r(n')$ and $r'(n')$ lie in different components of $\Gamma \setminus B_n^\circ$, so every path between them passes inside B_n°. Thus there is no geodesic between these points which lies wholly outside the closed ball of radius n, $\overline{B}_n$, and by

[5]In fact it will turn out that defining shadow(P) as a set of geodesic rays gives rise to a well-defined subset of ∂G directly. At the moment the definition is equivalent to saying that a point of ∂G is in shadow(P) iff one of its representatives passes through P.

corollary 6.2.2 on page 147, $\delta(r, r') \geqslant \frac{1}{b^n}$.

Thus if two geodesic rays pass through different patches, the naive distance between them is at least b^{-n}. We now need to prove the same result for the pseudometric on $\widehat{\partial G}$.

Let $(a_i)_{i=0}^{l}$ be a sequence of geodesic rays with $a_0 = r$ and $a_l = r'$. Then since r' does not pass through P, there is a smallest i such that a_i does not pass through P. Since a_{i-1} *does* pass through P, $\delta(a_{i-1}, a_i) \geqslant b^{-n}$, so $\sum_{i=1}^{l} \delta_b(a_{i-1}, a_i) \geqslant b^{-n}$. Thus all numbers in $D(r, r')$ are at least b^{-n}, so their infimum is too, i.e. $d_b(r, r') \geqslant \frac{1}{b^n}$.

□

This means that if $d_b(r, r') = 0$, then r and r' pass through the same patch, so shadow(P) is already a well-defined subset of ∂G. Thus the set of shadows of patches at radius n partitions ∂G. Also, since $d_b(-,-)$ is well-defined on ∂G, the lemma above holds for elements of ∂G.

Corollary 6.3.3 (shadows are open sets)

Let P be a patch at distance n from id_G. Let r represent an element of ∂G, with $r \in$ shadow(P). Then by the previous lemma, $d_b(r, r') < \frac{1}{b^n} \Rightarrow r' \in$ shadow(P). Thus shadow(P) is an open subset of ∂G.

□

Corollary 6.3.4 (shadows are closed sets)

Let P be a patch at distance n from id_G. There are only finitely many patches at distance n from the identity, so there are only finitely many shadows of those patches. Every element of ∂G is an element of one of these shadows, *so the complement of* shadow(P) *is the union of finitely many patch shadows, i.e. a union of finitely many open sets. Thus the complement of* shadow(P) *is open, i.e.* shadow(P) *is closed.*

□

6.3.2 Infinitely ended groups have messy boundaries.

Definition 6.3.5 *A metric space* X *is* uniformly path-connected *if there exists* $E > 0$ *such that for all* x *and* y *in* X, $d(x,y) < E$ *implies there is a path between them. We call the number* E *the constant of uniform path-connectedness, and sometimes say that* X *is* E*-uniformly path-connected.*

We prove that if G has infinitely many ends then its boundary is not uniformly path-connected.

Lemma 6.3.6 (Nearby separated rays)

Let G *have more than two ends, and let* N *be the smallest natural number such that* B°_N *separates the Cayley graph (*Γ*) into at least three infinite components.*

Fix $m \geqslant N+2$. *Suppose* r *is a geodesic ray, and let* $g = r(m-(N+1))$. *Then there exists a geodesic ray* r' *such that*

- $r'(m+1)$ *and* id_G *lie in different components of* $\Gamma \setminus B^\circ_N(g)$, *and*
- $r'(m+1)$ *and* $r(m+1)$ *lie in different components of* $\Gamma \setminus B^\circ_N(g)$.[6]

[6] We do not claim that $r(m+1)$ and id_G lie in different components of $\Gamma \setminus B^\circ_N(g)$. It seems implausible that they do, but we don't need to know.

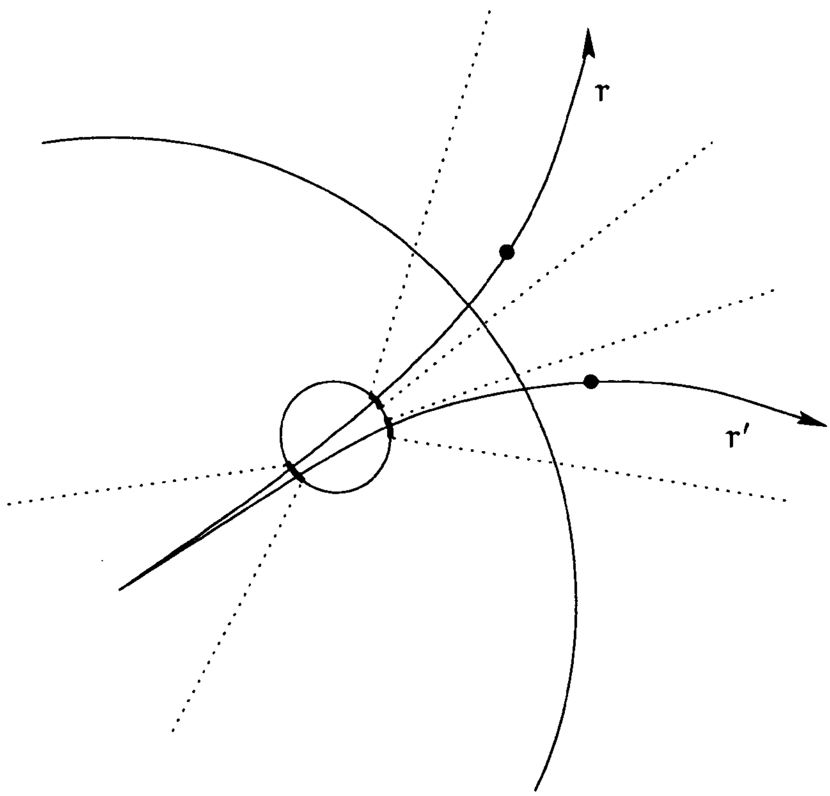

Diagram 6.3.2: Nearby rays in different shadows

Proof

Let $m \geqslant N + 2$, and note that by the triangle inequality, and the fact that $d(g, id_G) = m - (N+1)$, $B^\circ_N(g)$ lies wholly inside $B^\circ_m(id_G)$.

Let $h = r(m+1)$. There are at least three infinite components of $\Gamma \backslash B^\circ_N(g)$, so there is one in which neither id_G nor h lie. Call it C', and call the one in which h lies C.

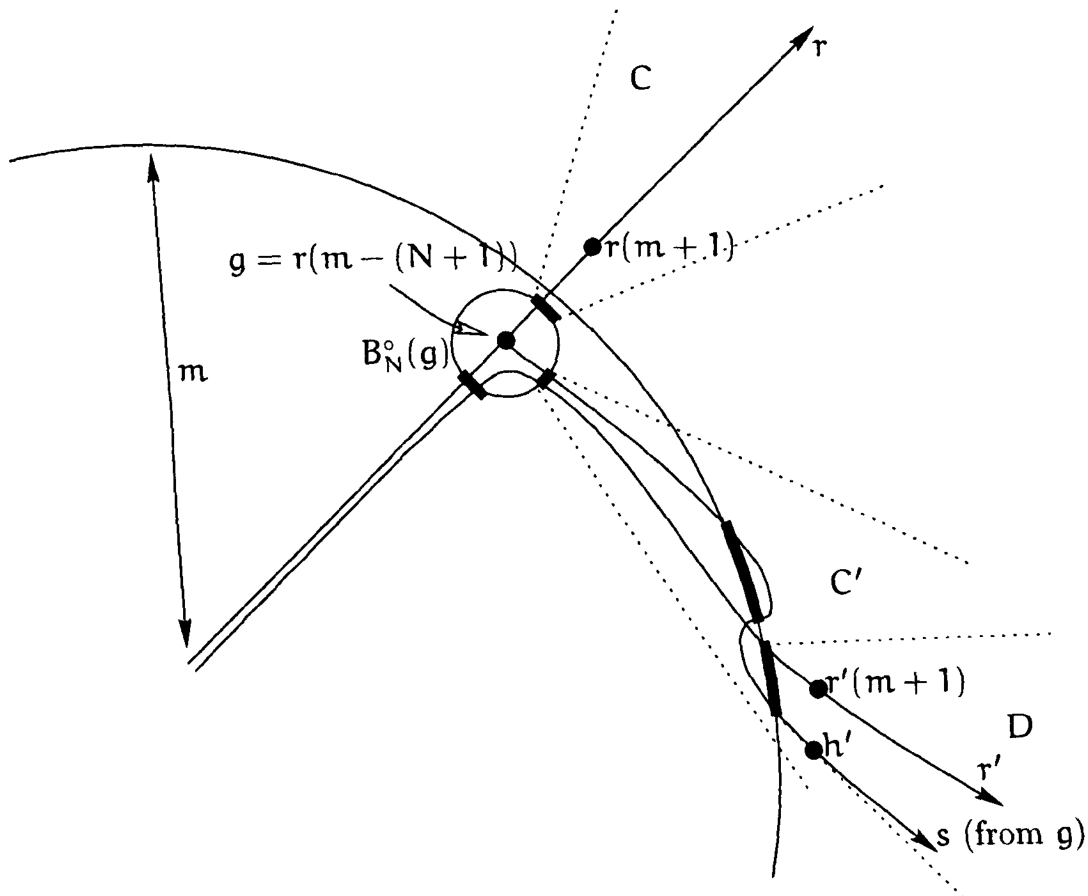

Diagram 6.3.3: Finding a nearby ray in a different shadow

Since C' is infinite, there is a geodesic ray $s : [0, \infty) \to \Gamma$ starting from g and passing into C', by lemma 2.3.1 on page 62. This geodesic can meet $B_m^\circ(\mathrm{id}_G)$ at only finitely many vertices, so after it has done so for the last time, we pick an element of G, $h' \in C' \cap \mathrm{im}(s)$ at distance $m + 1$ from id_G. Thus D, the component of $\Gamma \setminus B_m^\circ(\mathrm{id}_G)$ containing h', is infinite, because it has infinite intersection with s. Therefore there exists a geodesic ray r' starting from id_G, passing into D.

Since both h' and $r'(m+1)$ are in D, there is a path between them outside $B_m^\circ(\mathrm{id}_G)$. Now $B_N^\circ(g) \subset B_m^\circ(\mathrm{id}_G)$, so this path also lies outside $B_N^\circ(g)$, and so $r'(m+1)$ and h' are in the same component of $\Gamma \setminus B^\circ_N(g)$, namely C'.

Since C' was chosen so that neither id_G nor h lie in it, $r'(m+1)$ and id_G lie in different components of $\Gamma \setminus B^\circ_N(g)$, and $r'(m+1)$ and $r(m+1)$ lie in different components of $\Gamma \setminus B^\circ_N(g)$, as required.

□

Lemma 6.3.7 (Close but in different shadows)

Let G *have infinitely many ends, and let* $r \in \widehat{\partial G}$ *be a geodesic ray starting from* id_G*. Let* N *be the smallest natural number such that* B°_N *separates the Cayley graph into at least three infinite components.*

Then $\forall m \geqslant N+2$, $\exists r' \in \widehat{\partial G}$ *such that* $d_b(r, r') \leqslant b^{3N+2-m}$*, but* r *and* r' *pass through different patches at distance* m *from* id_G*, i.e. they lie in different shadows of patches at radius* m*.*

Proof

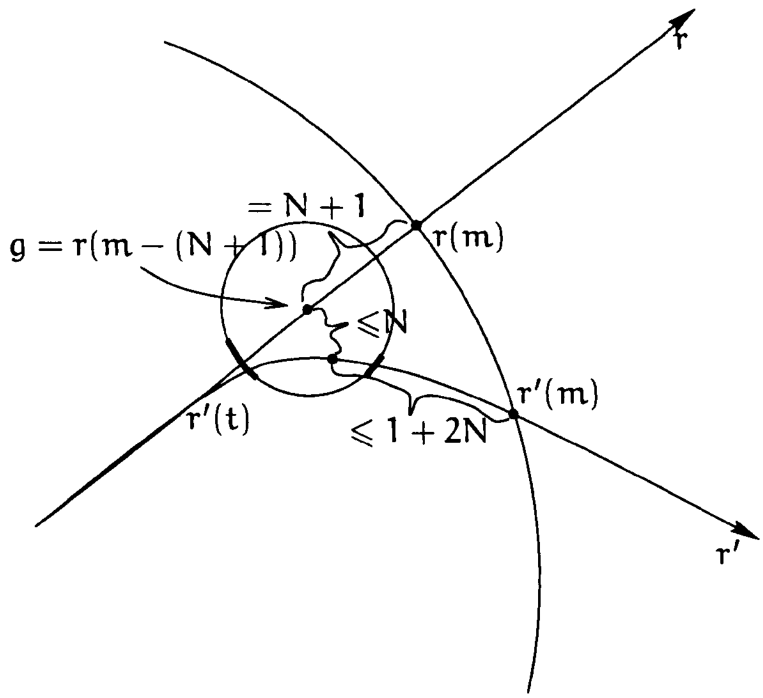

Diagram 6.3.4: The two rays are close.

Let r' be as in the previous lemma, and as before, let $g = r(m-(N+1))$. Then $r'(m+1)$ and id_G lie in different components of $\Gamma \setminus B^\circ_N(g)$, so $r'|_{[0,m+1]}$ must pass through $B^\circ_N(g)$. Thus there exists $t \in \mathbb{N}$ such that $r'(t) \in B^\circ_N(g)$, i.e. $r'(t)$ is within distance N of g.

Thus by the triangle inequality, $t + N \geqslant d(g, id_G) = m - (N+1)$, so $t \geqslant m - (2N+1)$. Thus we have

$$d(r'(m),\ r'(t)) \leqslant 1 + 2N,$$

$$d(r'(t),\ g) \leqslant N, \quad \text{and}$$

$$d(g,\ r(m)) = N + 1, \quad \text{so}$$

$$d(r'(m),\ r(m)) \leqslant 3N + 2.$$

Thus a geodesic between $r'(m)$ and $r(m)$ stays outside $B^{\circ}_{m-(3N+2)}(id_G)$, i.e. $d_b(r, r') \leqslant b^{3N+2-m}$.

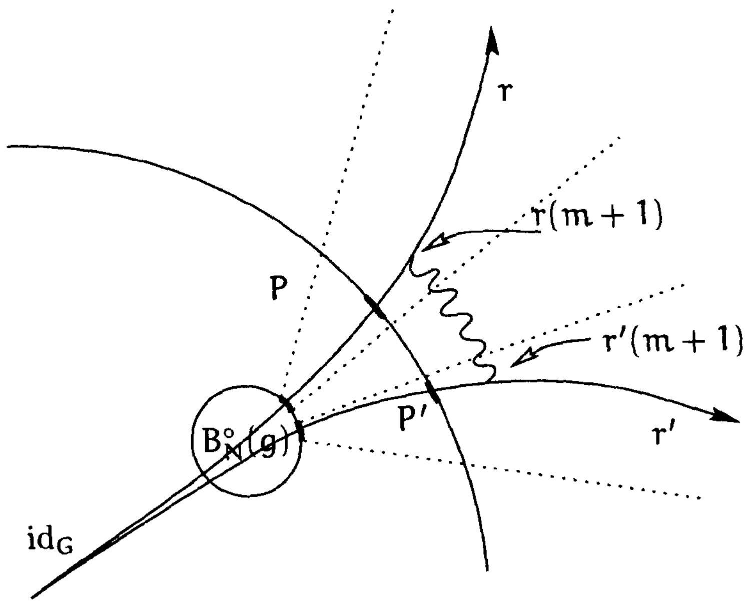

Diagram 6.3.5: The two rays are in different shadows.

All that remains is to prove that r and r' lie in different shadows.

Let P be the patch at distance m from id_G determined by $r(m+1)$, and let P' be the patch at distance m from id_G determined by $r'(m+1)$. Then $r \in \text{shadow}(P)$ and $r' \in \text{shadow}(P')$.

We need to prove that $P \neq P'$. By the patches lemma 2.2.3, this is equivalent to showing that there is no path from $r(m+1)$ to $r'(m+1)$ lying wholly outside $B^{\circ}_m(id_G)$. If there were such a path, it would lie outside $B^{\circ}_N(g)$ as well because $B^{\circ}_N(g) \subset B^{\circ}_m(id_G)$, but this contradicts the previous lemma, which states that $r(m+1)$ and $r'(m+1)$ are in *different* components of $\Gamma \setminus B^{\circ}_N(g)$.

Thus r and r' lie in different shadows.

□

Theorem 6.3.8 (∞ ends implies ∂G not uniformly path-connected)

Let G have infinitely many ends, and let $x \in \partial G$. Then for every open set U in ∂G containing x, there is an element $y \in U$ such that there does not exist a path from x to y. (Thus ∂G is not uniformly path-connected.)

Proof

Let $U \subset \partial G$ be an open set containing x, and let r be a geodesic ray representing x. Then there exists $\varepsilon > 0$ such that $d_b(x, y) < \varepsilon \Rightarrow y \in U$. Let N be the smallest natural number such that B_N° separates the Cayley graph into at least three infinite components.

Pick m large enough so that both $m \geqslant N + 2$ and $b^{3N+2-m} < \varepsilon$. The hypothesis of the previous lemma is satisfied, so let r' be as in the conclusion, and define y to be the element of ∂G determined by r'.

Firstly, $d_b(r, r') \leqslant b^{3N+2-m} < \varepsilon$, i.e. $d_b(x, y) < \varepsilon$ and $y \in U$.

Secondly, r and r' pass through different patches at distance m from id_G, so x and y are in different patch shadows. By corollaries 6.3.3 and 6.3.4, these shadows are both open and closed, so x and y are in different components of ∂G; there is no path between them.

x was an arbitrary element of ∂G, and we have shown that every open neighbourhood of x contains a point from a different path-component. ∂G is not uniformly path-connected.

□

6.3.3 ∂G is compact

Theorem 6.3.9

∂G *is compact.*

The proof of this theorem is similar to the proof of the compactness of the boundary of a hyperbolic group.

Proof

Consider $\partial G = \partial_b G$ for some $b \geqslant 2$.

Let $(x_i)_{i \in \mathbb{N}}$ be an arbitrary sequence of points of ∂G, and let $(r_i)_{i \in \mathbb{N}}$ be a sequence of geodesic rays (i.e. points in $\widehat{\partial G}$) representing them. We show that $(r_i)_{i \in \mathbb{N}}$ has a subsequence which is convergent in the sense that the naive distance between elements of the subsequence and the 'limit' ray tends to zero. Since the distance in ∂G between two points is bounded above by the naive distance between any pair of representatives, this is enough to prove that the corresponding subsequence of $(x_i)_{i \in \mathbb{N}}$ is convergent.

The subsequence will be denoted by $(r_{i_n})_{n \in \mathbb{N}}$, and we will define it element by element, by induction. When we say later that we 'pass to a subsequence' we do not refer to $(r_{i_n})_{n \in \mathbb{N}}$, rather we refer to a subsequence of the original sequence to which we restrict our attention for the purposes of the argument.

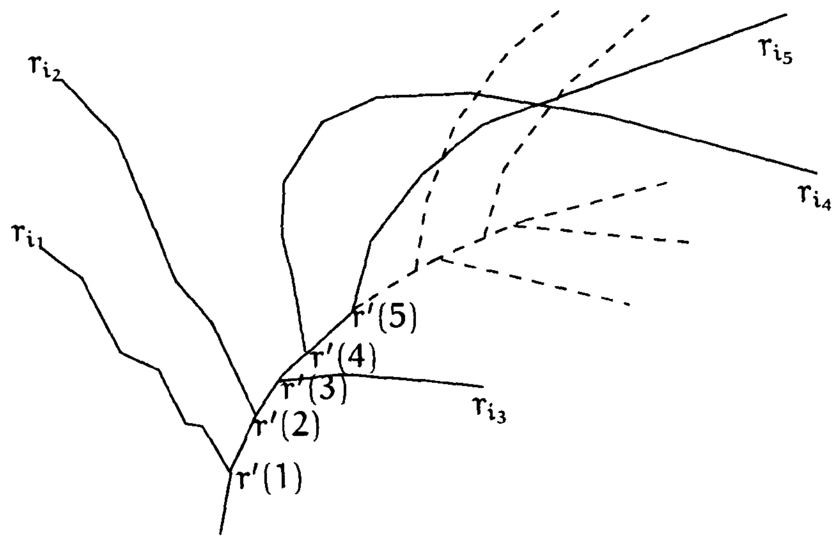

Diagram 6.3.6: Compactness in ∂G

We define $r_{i_1} = r_1$. Assume for induction that we have defined r_{i_m} for all $m \leqslant n$, and have defined a geodesic r' up to length n, with r_{i_m} equal to r' up to the point $r'(m)$. Also assume inductively that we have passed to a subsequence of the original sequence of rays in which all of them are equal to r' up to the point $r'(n)$.

There are finitely many elements at distance $n + 1$ from id_G through which the rays in this subsequence pass, so there is at least one through which infinitely many of them pass. Define this point to be $r'(n+1)$, and pass to a subsequence again so that all the rays in the new subsequence pass through $r'(n+1)$, i.e. are the same as r' up to the point $r'(n+1)$. Pick the first ray in this subsequence as $r_{i_{n+1}}$. This completes the induction.

The induction defines a geodesic ray r' and a subsequence $(r_{i_n})_{n \in \mathbb{N}}$ of the original sequence such that for each $n \in \mathbb{N}$, r_{i_n} is the same as r' up to distance n from id_G. Thus there is a (trivial) geodesic between $r_{i_n}(n)$ and $r'(n)$ which stays outside the closed ball of radius $n-1$, so by corollary 6.2.2

on page 147, $\delta_b(r_{i_n}, r') < b^{-(n-1)}$.

Thus $\delta_b(r_{i_n}, r')$ tends to zero as n tends to infinity, so since $d_b(r_{i_n}, r') \leqslant \delta_b(r_{i_n}, r')$, we know also that $d_b(r_{i_n}, r')$ tends to zero. Let x' be the element of ∂G determined by r'. Since the metric on ∂G is well-defined and equal to the pseudometric on elements of $\widehat{\partial G}$, this proves that the subsequence $(x_{i_n})_{n \in \mathbb{N}}$ determined by $(r_{i_n})_{n \in \mathbb{N}}$ also has the property that $\delta_b(x_{i_n}, x')$ tends to zero as n tends to infinity.

Thus $(x_i)_{i \in \mathbb{N}}$ has a convergent subsequence.

□

6.4 Exterior paths and connectivity at the boundary

We use the phrase 'connectivity at the boundary' for brevity — we in fact refer to uniform path-connectedness, i.e. there is a constant, E, such that if two points of the boundary are within distance E of each other, there is a path between them.

6.4.1 Notation and assumptions

We suppose that G is automatic, and let WA be its word acceptor. We also assume that the accepted words are all geodesics. Thus the fellow-travelling constant, f, is a fellow-travelling constant for accepted geodesics from the identity which terminate within distance 1 of each other.

Recall from definition 4.2.3 on page 102 that an infinite word is travellable iff every initial finite subword can be extended to an accepted word.

6.4.2 Pruning and closeness of infinite words

Recall also (from definition 4.2.2 on page 101) that pruned(WA) accepts exactly those words in the generators which possess infinite extension.

Definition 6.4.1 (global closeness of infinite travellable words) *Each word accepted by* WA *is a word accepted by* pruned(WA) *followed by a word which traces a path wholly outside the* pruned(WA) *automaton; this latter part of the path cannot trace any loops in the automaton. Since there are finitely many states in* WA, *there is a global bound, which we call* C, *on the length of this final subword.*

Note that C *is a global constant depending only on the automatic structure of* G. *The crucial property of* C *is that for any element* $g \in G$, g *is at most distance* C *from an infinite travellable word.*

6.4.3 Exterior paths implies uniform path-connectedness

Recall the following definition from the previous chapter.

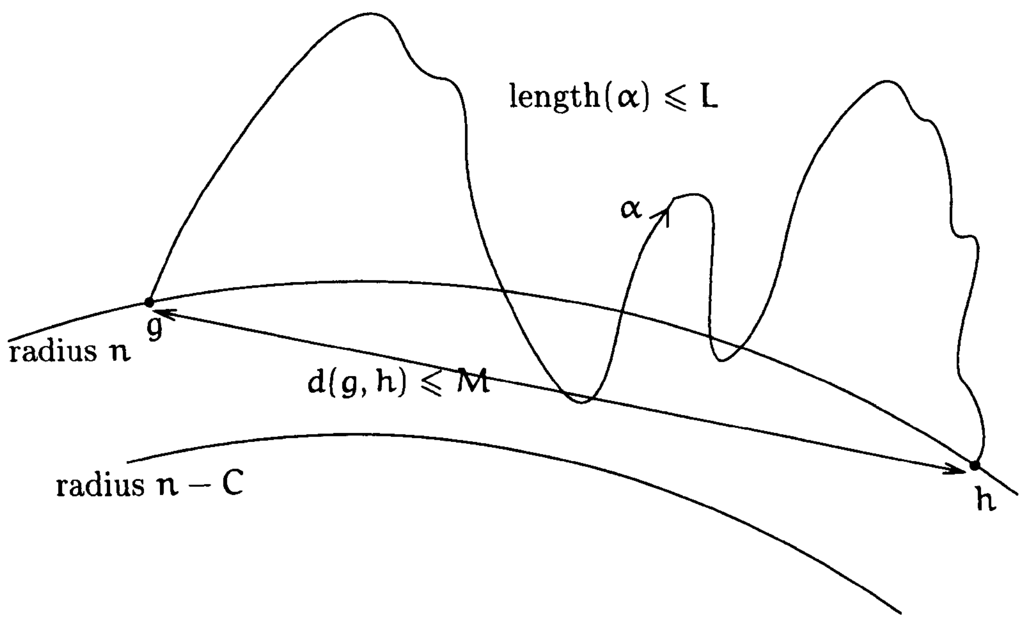

Diagram 6.4.1: Exterior paths

Definition 6.4.2 *Let* G *be a group with an upper bound,* C*, on the distance of an arbitrary element to a geodesic ray starting at* id_G*. Let* M *and* L *be natural numbers.*

We say that G *has* (M, L)-exterior-paths *if* $\forall n \in \mathbb{N}$*, and* $\forall g, h \in S_n$*,* $d(g, h) \leqslant M$ *implies that there is a path,* $\alpha : g \simeq h$ *of length at most* L *such that* α *lies entirely outside the (open) ball of radius* $n - C$*,* B°_{n-C}*.*

It may be that this condition is only satisfied for large n*. In this case, we use the phrase 'outside radius* N*':*

We say that a group has (M, L)-exterior-paths outside radius N *if* $\forall n \in \mathbb{N}$ *with* $n > N$*, and* $\forall g, h \in S_n$*,* $d(g, h) \leqslant M$ *implies that there is a path,* $\alpha : g \simeq h$ *of length at most* L *such that* α *lies entirely outside the (open) ball of radius* $n - C$*,* B°_{n-C}*.*

We remind the reader that G has (M, L)-exterior-paths if and only of it has (M, L)-exterior-paths outside radius 0, because the case when g and h are id_G is trivial; indeed, the case when $n \leqslant C$ is always trivial.

This section leads to the proof of the following theorem. It is stated again and proved as theorem 6.4.8 on page 176 and corollary 6.4.9 on page 178.

Theorem

Let G *be an automatic group with geodesic accepted words that* f*-fellow-travel, and let* C *be the global bound on the distance to some infinite travellable geodesic. Assume that for some* $M \geqslant (2C+1)(\max\{f, 3\})$ *and some* $L \in \mathbb{N}$, G *has* (M, L)*-exterior-paths outside radius* N*. Then for* $b \geqslant \max\{2, L\}$, $\partial_b G$ *is uniformly path-connected, with constant of uniform path-connectedness* $b^{-N+(C+1)}$.

In particular, if $N = 0$ *(i.e.* G *has exterior paths everywhere), then for* $b \geqslant \max\{2, L\}$, $\partial_b G$ *is path connected.*

We defer the proof until later, after some examples and preliminary lemmas.

The ideas for the proof of this theorem are taken from Lemma 3.1 of [BM91]. Bestvina and Mess assume that G is hyperbolic, and perform their calculations in the Rips complex[7] of G. They assume the group has exterior paths everywhere (for some M larger than a constant depending on the constant of hyperbolicity as well as C), and they prove local path connectedness. In this sense their result is quite different to ours.

[7] A definition can be found at the beginning of [BM91].

Example 6.4.3

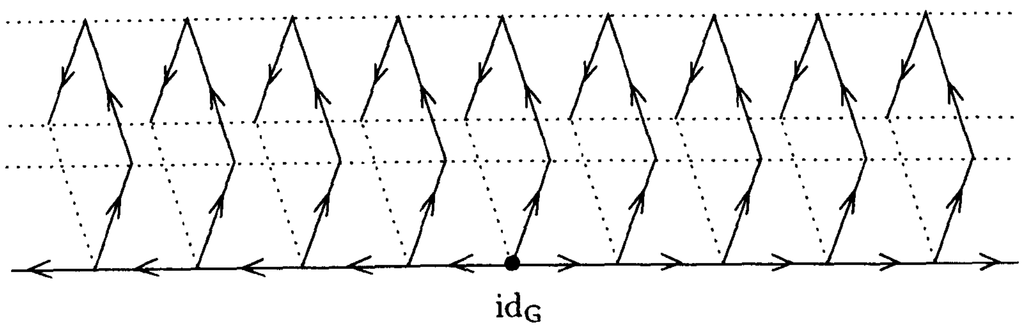

Diagram 6.4.2: $\mathbb{Z} \times \mathbb{Z}_4$ spanning tree with x before y

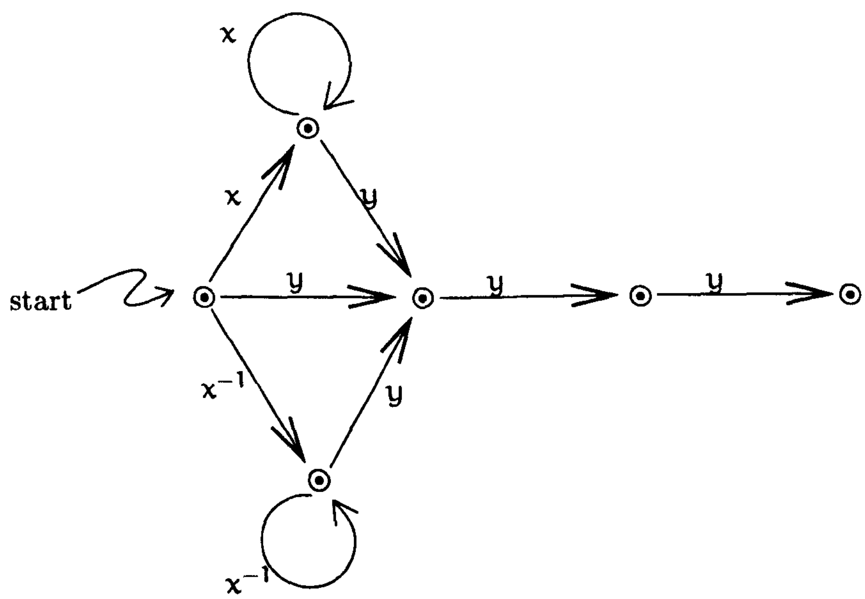

Diagram 6.4.3: $\mathbb{Z} \times \mathbb{Z}_4$ word acceptor with x before y

Let $G = \langle x, y \mid xyx^{-1}y^{-1}, y^4 \rangle \cong \mathbb{Z} \times (\mathbb{Z}/4\mathbb{Z})$, and give it the automatic structure where accepted words are some power of x followed by some non-negative power of y. Thus the constant for closeness of infinite travellable geodesics that we compute from the word acceptor is $C = 3$, the fellow-travelling constant is $f = 2$, and $(2C + 1)(\max\{f, 3\}) = 12$.

The open ball of radius $n - C$ disconnects the Cayley graph for $n > 2 + C$, but there are points in S_n of distance just $2(n-2)$ apart in different components of the complement. Thus we need $M < 2(n-2)$, but also $12 = (2C+1)(\max\{f,3\}) \leqslant M$, so we need $n > 8$, i.e. $N = 8$.

The distance between points in S_n that are in the same component of $\Gamma \setminus B^{\circ}_{n-C}$ is at most 4, and we can choose the geodesic to lie outside $\Gamma \setminus B^{\circ}_{n-C}$.

Thus in this case, G has (12, 4)-exterior-paths outside radius 8, and so for $b \geqslant 4$, $\partial_b G$ is $\mathcal{E}$-uniformly path-connected with $\mathcal{E} = b^{-8+3+1} = b^{-4}$. This is of course trivially true, because in this case ∂G is a two-point set of diameter b^{-2} (see example 6.2.4 on page 149).

Our result gives us a better picture of the structure of the boundary when the automatic structure is such that the values of C and f we calculate from it are as small as possible:

Example 6.4.4

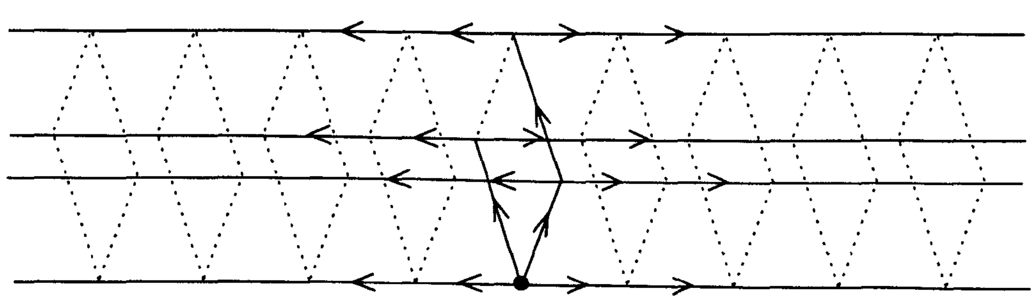

Diagram 6.4.4: $\mathbb{Z} \times \mathbb{Z}_4$ spanning tree with y before x

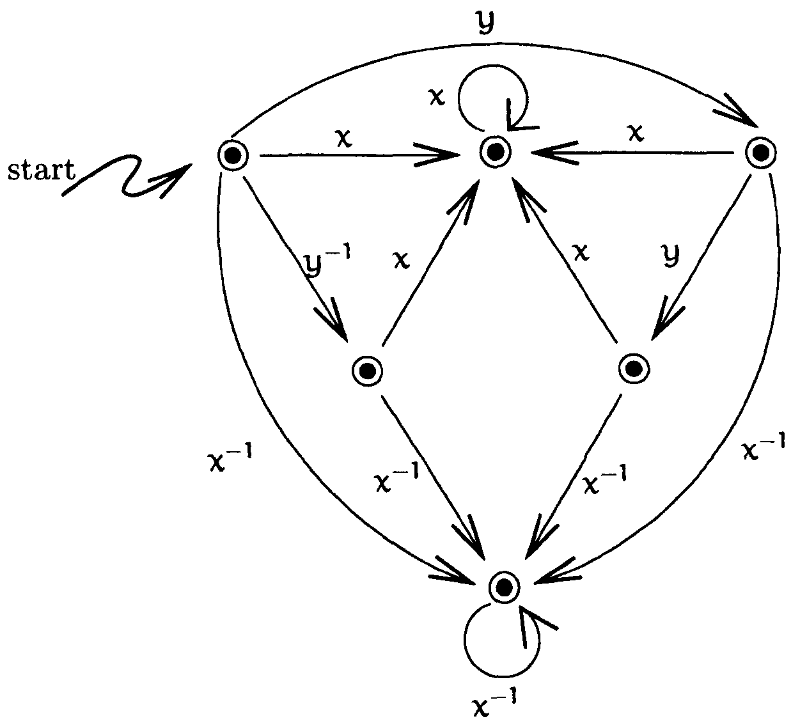

Diagram 6.4.5: $\mathbb{Z} \times \mathbb{Z}_4$ word acceptor with y before x

Again, let $G = \langle x, y \mid xyx^{-1}y^{-1}, y^4 \rangle \cong \mathbb{Z} \times (\mathbb{Z}/4\mathbb{Z})$, but this time give it the automatic structure where accepted words are some power between -1 and 2 of y, followed by some power of x. This time, $C = 0$, $f = 2$, and $(2C+1)(\max\{f, 3\}) = 3$, so we can pick $L = 3$, and we have $(3,3)$-exterior-paths outside radius 3.

Thus for $b \geqslant 3$, $\partial_b G$ is E-uniformly path-connected with $E = b^{-3+1} = b^{-2}$. This is as strong a result as we could hope for, because ∂G is a two-point set of diameter b^{-2}.

Lemma 6.4.5 (Inductively interposing rays)

Let G be an automatic group with geodesic accepted words that f-fellow-travel, and let C be the global bound on the distance to some infinite travellable geodesic. Assume that for some $M \geqslant (2C+1)(\max\{f, 3\})$ and some $L, N \in \mathbb{N}$,

G *has* (M, L)*-exterior-paths outside radius* N.

Suppose we have two geodesic rays r *and* r' *from* id_G*, and that for some* $n > N$, $d(r(n), r'(n)) \leqslant M$. *Then there exist* $L+1$ *geodesic rays,* $\{r_i\}_{i\in\{0,1,\dots,L\}}$ *with* $r_0 = r$ *and* $r_L = r'$*, and* $\forall i \in \{1, 2, \dots, L\}$, $d(r_{i-1}(n+1), r_i(n+1)) \leqslant M$.

Proof

Since $d(r(n), r'(n)) \leqslant M$ and $n > N$, we can find a path of length at most L, $\alpha : r(n) \simeq r'(n)$ that stays outside B°_{n-C}. For notational convenience, we parameterise α by path length, and then compose it with a stationary path so that $\alpha : [0, L] \to \Gamma$.

For each $i \in \{0, 1, \dots, L\}$, we choose a geodesic ray r_i to pass within distance C of $\alpha(i)$, and call the point of closest approach p_i. (For $i = 0$ we pick r, and for $i = L$ we pick r'.) Now p_i and p_{i+1} are within distance $2C+1$, so let β be a geodesic between them with length$(\beta) \leqslant 2C + 1$.

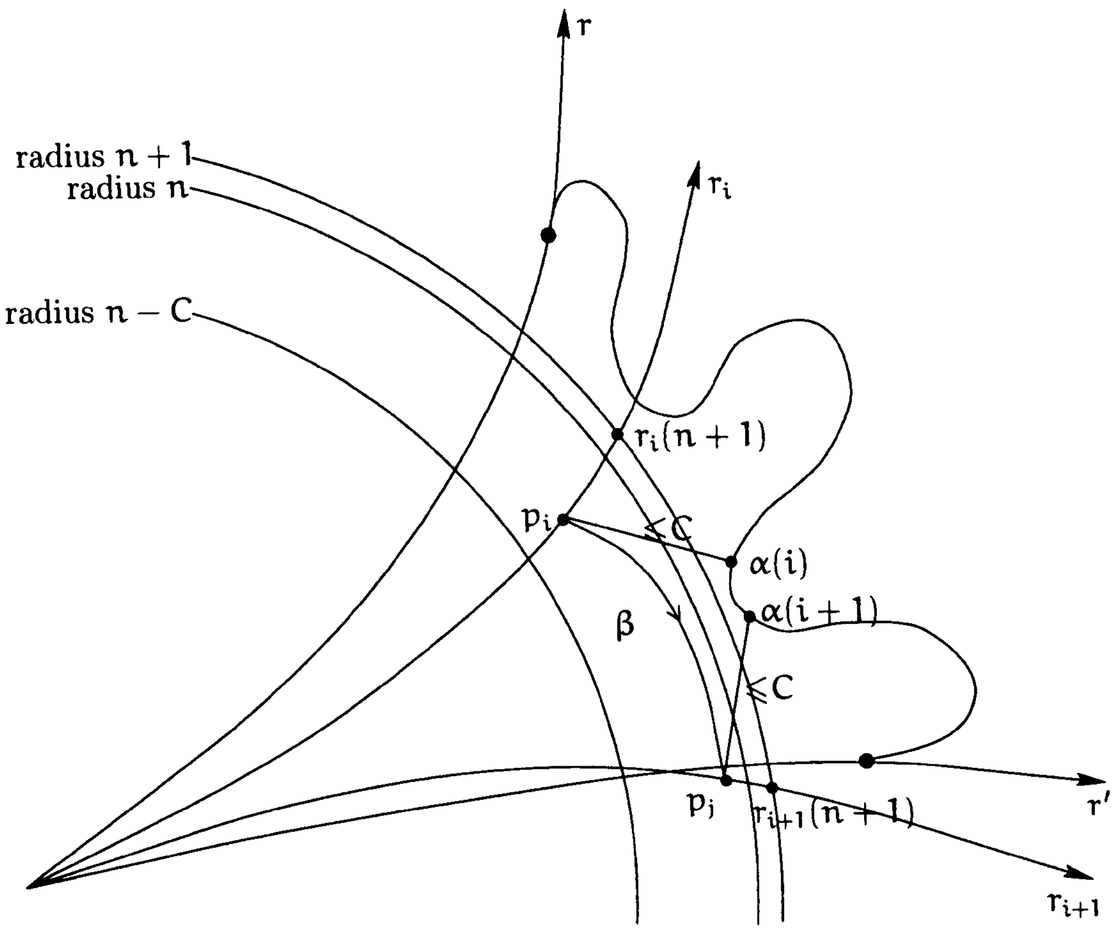

Diagram 6.4.6: Interposing rays

Case (1): β lies entirely outside B°_{n+1}. Construct accepted geodesics to all the vertices on β, using r_i and r_{i+1} for the endpoints. Each of these geodesics is at least of length $n+1$ and they f-fellow-travel, so when we truncate them to length $n+1$, adjacent endpoints are within distance f. Thus $d(r_i(n+1), r_{i+1}(n+1)) \leqslant (2C+1)f$.

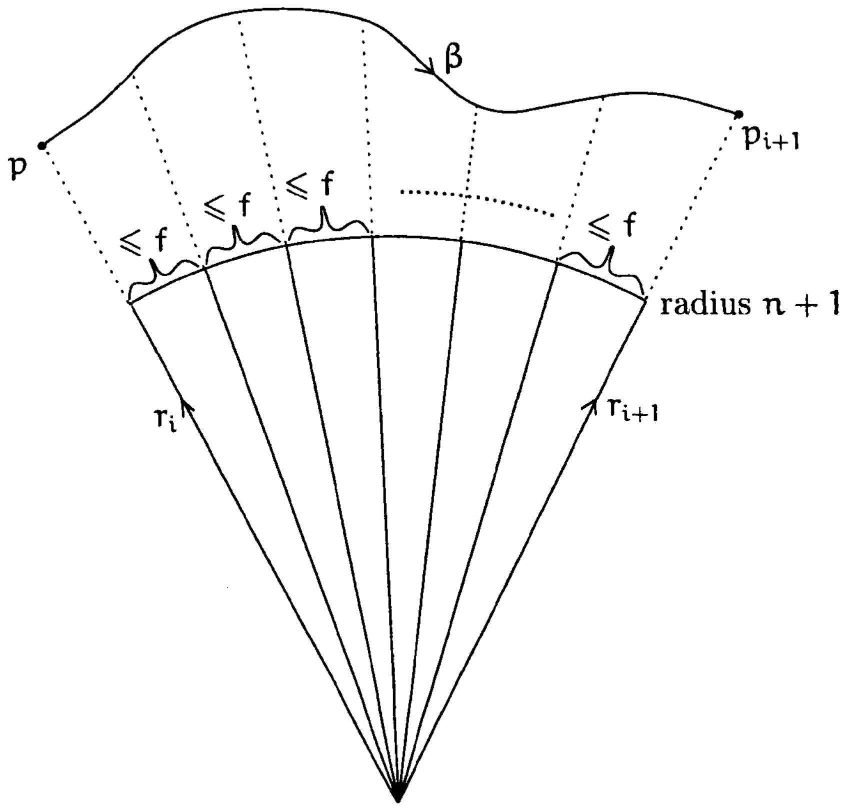

Diagram 6.4.7: Case (1): Truncating geodesics that fellow-travel

Case (2): There is a point on β in $\overline{B}_n$. In this case, the endpoints of β, p_i and p_{i+1} are at most distance $n+(2C+1)$ from the identity. Also, p_i and p_{i+1} lie outside B°_{n-2C} because they are within distance C of α, which lies outside B°_{n-C}. Thus for $j \in \{i, i+1\}$, $p_j \in \overline{B}_{n+1+2C} \setminus B^\circ_{n-2C}$. Thus since p_j is a point on r_j, we have $d(p_j, r_j(n+1)) \leqslant 2C+1$. We also know that $d(p_i, p_{i+1}) \leqslant (2C+1)$, so by the triangle inequality, $d(r_i(n+1), r_{i+1}(n+1)) \leqslant 3(2C+1)$.

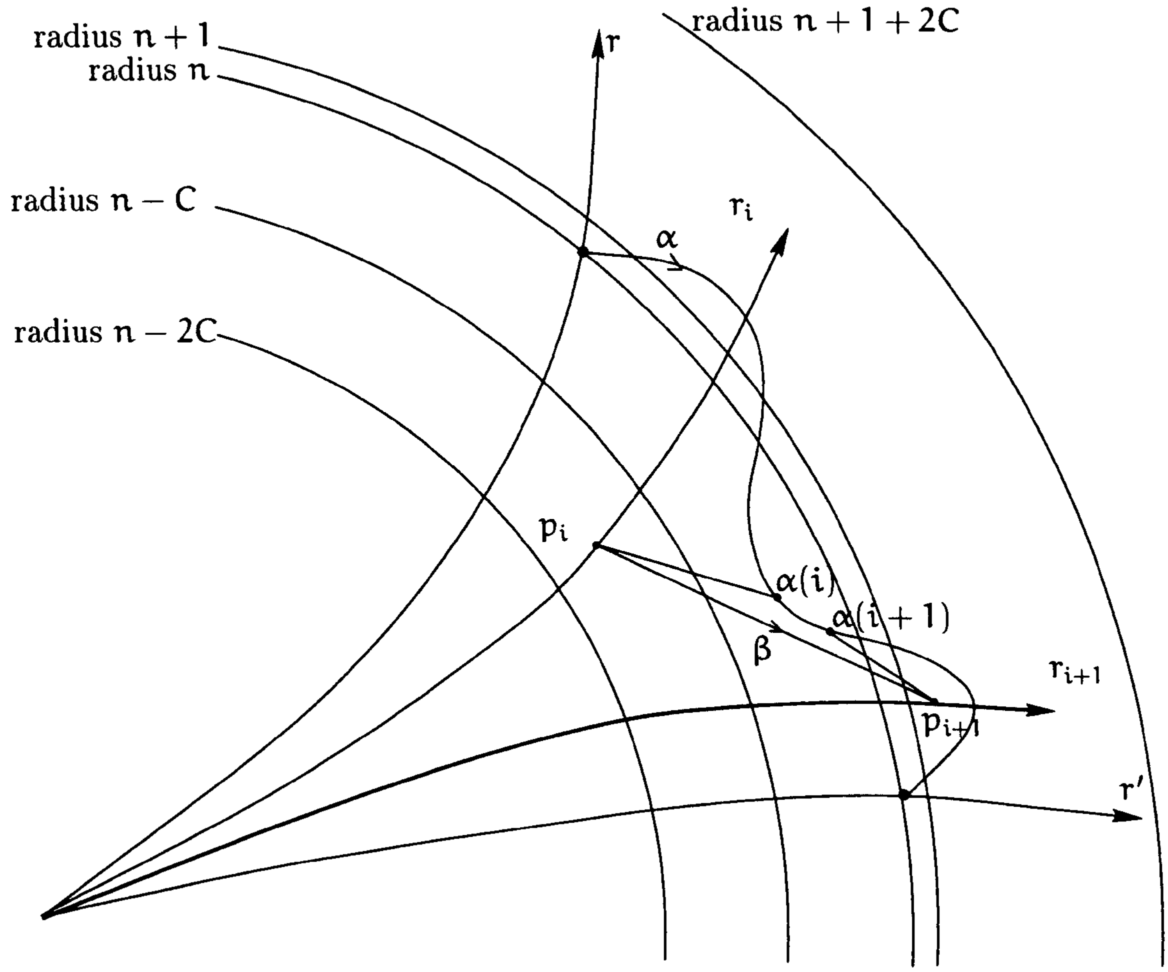

Diagram 6.4.8: Case (2): β lies in a band

Thus in either case, $d(r_i(n+1), r_{i+1}(n+1)) \leqslant (2C+1)(\max\{f,3\}) \leqslant M$ as required.

□

Lemma 6.4.6 (The base case)

Let G *be an automatic group with geodesic accepted words that* f*-fellow-travel, and let* C *be the global bound on the distance to some infinite travellable geodesic. Suppose we have two geodesic rays* r *and* r' *from* id_G*, and that*

$\delta(r, r') < b^{-n+(C+1)}$ *for some* $n \in \mathbb{N}$.

Then for some $L' \in \mathbb{N}$, *there exist* $L' + 1$ *geodesic rays,* $\{r_i\}_{i \in \{0,1,\ldots,L'\}}$ *with* $r_0 = r$ *and* $r'_L = r'$, *and* $\forall i \in \{1, 2, \ldots, L'\}$, $d(r_{i-1}(n+1), r_i(n+1)) \leqslant M$.

The similarity to the previous lemma will not be lost on the reader.

Proof

Since $\delta(r, r') < b^{-n+C+1}$, there is a natural number m and a path $\alpha : r(m) \simeq r'(m)$ that stays outside $\overline{B}_{n-(C+1)}$, i.e. outside B°_{n-C}; we are in exactly the same situation as in the previous lemma, except that this time α is of some unknown length L', and its endpoints are not necessarily at distance exactly $n + 1$ from id_G. This latter difference does not change the argument.

Using exactly the same argument as in the previous lemma, construct this time L' geodesic rays, $\{r_i \mid i \in \{0, 1, \ldots, L'\}\}$ with $d(r_i(n+1), r_{i+1}(n+1)) \leqslant M$ as required.

□

Lemma 6.4.7

Assume that for some $M \geqslant (2C + 1)(\max\{f, 3\})$, *and some* $L, N \in \mathbb{N}$, G *has* (M, L)*-exterior-paths outside radius* N, *and suppose* $b \geqslant \max\{2, L\}$. *Then if* $\varphi(0)$ *and* $\varphi(1)$ *are geodesic rays with* $\delta_b(\varphi(0), \varphi(1)) < b^{-N+(C+1)}$, *then we may extend* φ *to a path between them in* $\partial_b G$.

Proof

By the base case lemma, lemma 6.4.6, we find there are $L' + 1$ geodesic rays $\{r_i\}_{i \in \{0,1,\ldots,L'\}}$ for some $L' \in \mathbb{N}$, with $r_0 = \varphi(0)$ and $r'_L = \varphi(1)$, and $\forall i \in \{1, 2, \ldots, L'\}$, $d(r_{i-1}(N+1), r_i(N+1)) \leqslant M$. If there are paths between adjacent r_i, then we may compose them to obtain a path between $r_0 = \varphi(0)$ and $r'_L = \varphi(1)$, so we may assume without loss of generality that $d(\varphi(0)(N+1), \varphi(1)(N+1)) \leqslant M$.

Claim *We inductively define* $\varphi(\frac{i}{L^j})$ *for each* $j \in \mathbb{N}$ *and for* $i \in \{0, 1, \ldots, L^j\}$, *so that* $d_b(\varphi(\frac{i}{L^j}), \varphi(\frac{i+1}{L^j})) < b^{-N-j+(C+1)}$*.*

Assume for induction on $j \in \mathbb{N}$ that for all $i \in \{0, 1, \ldots, L^j - 1\}$, we have $d(\varphi(\frac{i}{L^j})(N+j), \varphi(\frac{i+1}{L^j})(N+j)) \leqslant M$.

Let $i \in \{0, 1, \ldots, L^j - 1\}$. Then since we have (M, L)-exterior-paths outside radius N, by lemma 6.4.5 on page 169 there are $L + 1$ geodesic rays, $\{r_k\}_{k\in\{0,1,\ldots,L\}}$ with $r_0 = \varphi(\frac{i}{L^j})$, $r_L = \varphi(\frac{i+1}{L^j})$, and $\forall k \in \{1, 2, \ldots, L'\}$, $d(r_{k-1}(N+j+1), r_k(N+j+1)) \leqslant M$.

Define $\varphi(\frac{k}{L^{j+1}}) = r_k$ for $k \in \{0, 1, \ldots, L\}$.

Since $d(\varphi(\frac{i}{L^j})(N+j), \varphi(\frac{i+1}{L^j})(N+j)) \leqslant M$, there is a geodesic between them that lies outside B°_{N+j-C}, so it lies outside $\overline{B}_{N+j-(C+1)}$. Thus $d_b(\varphi(\frac{i}{L^j}), \varphi(\frac{i+1}{L^j})) < b^{-N-j+(C+1)}$.

This completes the induction.

Claim *If* $\left|\frac{i}{L^j} - \frac{i'}{L^{j'}}\right| < \frac{1}{L^j}$ *for some* $j' > j$*, then* $d_b(\varphi(\frac{i}{L^j}), \varphi(\frac{i'}{L^{j'}})) < b^{-N-j-(C+1)}$*.*

We prove this by induction on the difference between j and j'. Now by the earlier claim, for $0 \leqslant k \leqslant L$,

$$\begin{aligned} d_b(\varphi(\frac{i}{L^j}), \varphi(\frac{iL+k}{L^{j+1}})) &< k(b^{-(N+j+1)+(C+1)}) \\ &\leqslant L(b^{-(N+j+1)+(C+1)}) \\ &\leqslant b^{-(N+j)+(C+1)} \end{aligned}$$

because we assumed $b \geqslant \max\{2, L\}$. Extending this result by induction, we complete the proof of the claim.

Claim φ *can be extended to a continuous function* $[0,1] \to \partial G$.

We abuse notation and use $\varphi(\frac{i}{L^j})$ to stand for both the geodesic ray (element of $\widehat{\partial G}$) and the corresponding element of ∂G. Distances remain unaffected.

We have defined φ on a dense subset of $[0,1]$, and by the previous claim it is continuous on this subset. Since ∂G is compact (theorem 6.3.9 on page 161), it is a complete metric space, so φ possesses a unique continuous extension over $[0,1]$.

□

Now we are in a position to prove the main theorem of the section:

Theorem 6.4.8

Let G *be an automatic group with geodesic accepted words that* f*-fellow-travel, and let* C *be the global bound on the distance to some infinite travellable geodesic. Assume that for some* $M \geqslant (2C+1)(\max\{f,3\})$ *and some* $L \in \mathbb{N}$, G *has* (M, L)*-exterior-paths outside radius* N. *Then for* $b \geqslant \max\{2, L\}$, $\partial_b G$ *is uniformly path-connected, with constant of uniform path-connectedness* $b^{-N+(C+1)}$.

Proof

Assume G is automatic with geodesic accepted words that f-fellow-travel, and assume G has (M, L)-exterior-paths outside radius N for some $M \geqslant (2C+1)(\max\{f, 3\})$ and some $L, N \in \mathbb{N}$. Suppose $b > \max\{2, L\}$.

Let $\psi(0)$ and $\psi(1)$ be elements of ∂G with $d_b(\psi(0), \psi(1)) < b^{-N+(C+1)}$. We extend ψ to a continuous function $\psi : [0, 1] \to \partial G$, thus proving that ∂G is $b^{-N+(C+1)}$-uniformly path-connected.

Since $d_b(\psi(0), \psi(1)) < b^{-N+(C+1)}$, it must also be true that there exists $\varepsilon > 0$ so that $d_b(\psi(0), \psi(1)) < b^{-N+(C+1)} - \varepsilon$. Let r and r' be two geodesic rays representing $\psi(0)$ and $\psi(1)$ respectively. Then since the pseudometric is well-defined and equal to the metric on ∂G, $d_b(r, r') = d_b(\psi(0), \psi(1))$. Recall that

$$d_b(r, r') = \inf\left\{ \sum_{i=1}^{l} \delta_b(a_{i-1}, a_i) \;\middle|\; \begin{array}{l} (a_i)_{i=0}^{l} \text{ is a finite sequence of points in } \widehat{\partial G} \\ \text{with } a_0 = r \text{ and } a_l = r'. \end{array} \right\}$$

Since $d_b(r, r')$ is the infimum, given $\varepsilon > 0$ there is an $l \in \mathbb{N}$ so that there exist rays $\{a_i \mid i \in \{0, \ldots, l\}\}$ such that $\sum_{i=1}^{l} \delta_b(a_{i-1}, a_i) < d_b(r, r') + \varepsilon$. We picked ε so that $d_b(\varphi(0), \varphi(1)) < b^{-N+(C+1)} - \varepsilon$, so $\sum_{i=1}^{l} \delta_b(a_{i-1}, a_i) < b^{-N+(C+1)}$. This implies that for each $i \in \{1, \ldots, l\}$, $\delta_b(a_{i-1}, a_i) < b^{-N+(C+1)}$, so by the previous lemma, there are paths $\{\gamma_i \mid i \in \{1, \ldots, l\}\}$ in ∂G, with $\gamma_i : a_{i-1} \simeq a_i$.[8] Composing these gives us a path ψ from $\psi(0)$ to $\psi(1)$

□

[8] Here we abuse notation and use a_i to stand for both the geodesic ray and the point of ∂G which it represents.

Corollary 6.4.9

Let G *be an automatic group with geodesic accepted words that* f*-fellow-travel, and let* C *be the global bound on the distance to some infinite travellable geodesic.*

If, for some $M \geqslant (2C+1)(\max\{f,3\})$, G *has* (M, L)*-exterior-paths, then for* $b \geqslant \max\{2, L\}$, $\partial_b G$ *is path connected:*

G *has* (M, L)*-exterior-paths outside radius* 0*, so by theorem 6.4.8,* $\partial_b G$ *is uniformly path-connected, with constant of uniform path-connectedness* $b^{-0+(C+1)} \geqslant 2 > 1$. *But when we defined the pseudometric on* $\widehat{\partial G}$*, we noted that any two rays are at most distance* 1 *apart, so the fact that any two boundary points that are within distance* 1 *of each other have a path between them means that* ∂G *is path-connected.*

□

6.4.4 The exterior paths condition for automatic groups

Definition 6.4.10 (The exterior paths condition for automatic groups)
Let G *be an automatic group with Cayley graph* Γ.

Let C *be the global bound on the distance to some infinite travellable geodesic. (See 6.4.1.)*

We say that G *satisfies the* exterior paths condition for automatic groups, *or the* automatic exterior paths condition *if there exists* $M \geqslant (2C+1)(\max\{f,3\})$, *and* $L, N \in \mathbb{N}$, *such that* G *has* (M, L)*-exterior-paths outside radius* N.

6.4.4.1 Consequences for computation

Theorem 6.4.8 on page 176 says that if an automatic group G satisfies the exterior paths condition for automatic groups, then its boundary is uniformly path-connected. Theorem 6.3.8 on page 160 proves that such a group cannot have infinitely many ends.

Let G be an automatic group with geodesic accepted words. We can run the algorithms ends (algorithm 4.2.8 on page 108), and ext.paths (algorithm 5.2.4 on page 139) concurrently. G has zero, two or uncountably many ends if and only if ends terminates, saying so, and if ext.paths terminates, we will know that the group has one end (because the case of two ends is decided by ends at the beginning).

6.4.4.2 Two-ended groups satisfy the exterior paths condition

Two-ended groups have two-point, and hence uniformly path-connected boundaries. One would hope, then, that they satisfy the exterior path condition for automatic groups. They do:

Theorem 6.4.11

Let G *be a group with two ends. Then there exists* $L \in \mathbb{N}$ *such that for all* $M \in \mathbb{N}$ *there exists* $N \in \mathbb{N}$ *such that* G *has* (M, L)*-exterior-paths outside radius* N.

Corollary 6.4.12

Two-ended groups satisfy the exterior paths condition for automatic groups.

□

Proof (of theorem 6.4.11)

Let G be 2-ended, and let $M \in \mathbb{N}$. Let κ be the connectivity constant of G.

G is automatic with geodesic accepted words because 2-ended groups are hyperbolic (they have $\mathbb{Z}$ as a subgroup of finite index [SW] and a group is hyperbolic if a subgroup of finite index is hyperbolic [GdlH90]). Let C be the bound on the distance to an infinite travellable geodesic.

By [Can], G has linear growth, so there is a bound, B, say, on the number of elements in the sphere of any radius. (So there can be only $(\kappa + 1)B$ elements in the band $\overline{B}_{n+\kappa} \setminus B^{\circ}_{n}$.)

Let $L = 4C + (\kappa + 1)B$.

Since G is two-ended, there is an $m \in \mathbb{N}$ such that the open ball of radius m disconnects the Cayley graph.

Let $N > M/2 + m$.

Any path between points in different components of $\Gamma \setminus B^{\circ}_{m}$ must pass through B°_{m}. If g and h are both at distance $n > N$ from id_G, and in different components of $\Gamma \setminus B^{\circ}_{m}$, then any path from g to h must first travel to B°_{m} from S_n, then travel across B°_{m}, and then travel out again to distance n from id_G, so is of length at least $2(n-m) > 2(N-m) > M$, (We chose $N > M/2+m$.) Thus if g and h are elements outside radius N, with $d(g,h) \leqslant M$, then g and h lie in the same component of $\Gamma \setminus B^{\circ}_{m}$.

Let $n > N$. Then $n > m$, so the ball of radius n disconnects Γ into two infinite components (and possibly some finite ones).

Any point g of S_n is within distance C of an infinite geodesic, so is within distance C of a ray r into an infinite component of $\Gamma \setminus B^{\circ}_{n}$. Thus there is a path of length at most C between g and $r(n')$, where $n - C < n' < n + C$, and this path must stay outside B°_{n-C}. Composing this with a subpath of r gives us a path of length at most $2C$ between g and $r(n)$, which stays outside B°_{n-C}.

Now let g and g' be two points of S_n that lie in the same component of $\Gamma \setminus B^\circ_m$. There is a path of length $2C$ lying outside B°_{n-C} from g to $r(n)$ and from $r'(n)$ to g', for some rays r and r'. $r(n)$ and $r'(n)$ must lie in infinite components of $\Gamma \setminus B^\circ_n$. These components are subsets of a single component of $\Gamma \setminus B^\circ_m$, and if they were distinct, there would be at least three infinite components of $\Gamma \setminus B^\circ_n$, which is false.

Thus $r(n)$ and $r'(n)$ must lie in the same component of $\Gamma \setminus B^\circ_n$, so there's a path between them that has its vertices inside the band, $\overline{B}_{n+\kappa} \setminus B^\circ_n$. There are at most $(\kappa + 1)B$ vertices in the band, se we can choose this path to be of length at most $(\kappa + 1)B$. Note that it, too, stays outside B°_{n-C}.

Thus any pair of vertices at distance n from id_G that lie in the same component of $\Gamma \setminus B^\circ_m$ have a path between them of length less than or equal to $2C + (\kappa + 1)B + 2C = L$, which stays outside the ball of radius $n - C$.

Thus for any M, we can find sufficiently large N such that for $n > N$, any two elements of S_n that are within distance M of each other are in the same component of $\Gamma \setminus B^\circ_m$ for some m, where B°_m disconnects the Cayley graph into two infinite components (and possibly some finite ones). By the above arguments, there is then a path between them of length at most L that stays outside the ball of radius $n - C$.

Thus G has (M, L)-exterior-paths outside radius N, for $L = 4C + (\kappa + 1)B$, any $M \in \mathbb{N}$, and $N > M/2 + m$. In particular, a 2-ended group satisfies the exterior paths condition for automatic groups.

□

6.5 A generalisation of Gerasimov's result?

Gerasimov's algorithm always detects one-endedness when it occurs in a hyperbolic group, and we would like to do the same for groups with good automatic structures. We have shown that the boundary of an arbitrary group can only be uniformly path-connected if it has finitely many ends, and for any automatic group, we can determine whether we are in the case of zero or two ends. Thus if we detect uniform path-connectedness in a group known to not have two ends, we know the group is one-ended.

To what extent is it true that if an automatic group (or indeed an arbitrary finitely presented group) has one end, its boundary is uniformly path-connected? Is it also true that if a group is uniformly path-connected it has (M, L)-exterior-paths outside radius N for M large enough and some L and N?

In a hyperbolic group we know from Bestvina, Mess, Bowditch and Swarup that having (M, L)-exterior-paths for large enough M is equivalent to having one end, but this relies on some deep results about the boundary of hyperbolic groups. The construction outlined here for the boundary of an arbitrary Cayley graph has a less refined notion of distance. We say that two rays are close if there is a geodesic between them which passes outside a ball of large radius, but the length of this geodesic is not taken into account. This means that the boundary sometimes collapses more than we expect — one might expect the boundary of $\mathbb{Z} \times \mathbb{Z}$ to be a circle, but with this construction it is a single point.

Perhaps with this less subtle boundary, the exterior paths condition is indeed directly equivalent to having a uniformly path-connected boundary. Proving this would prove that our algorithm always detects the number of ends of an automatic group which has geodesic accepted words.

Bibliography

[Ber68] George Bergman, *On groups acting on locally finite graphs*, Annals of Mathematics **88** (1968), 335–340.

[BG95] O Bogopolskii and V Gerasimov, *Finite subgroups of hyperbolic groups*, Algebra i Logika **34** (1995), no. 6, 619–622.

[Bil96] Stephen E Billington, *Growth of finite state automata*, M.Sc. dissertation, Mathematics Institute, University of Warwick, Sept 1996.

[BM91] Mladen Bestvina and Geoffrey Mess, *The boundary of negatively curved groups*, Journal of the American Mathematical Society **4** (1991), no. 3, 469–481.

[Bow95] Brian H Bowditch, *Cut points and canonical splittings of hyperbolic groups*, preprint, 1995.

[Bow96] Brian H Bowditch, *Connectedness properties of limit sets*, Preprint series, no. 270, Faculty of Mathematical Studies, University of Southampton, 1996.

[Can] J W Cannon, *The growth of the closed surface groups and the compact hyperbolic coxeter groups.*

[DD89] Warren Dicks and Martin J Dunwoody, *Groups acting on graphs*, Cambridge University Press, 1989.

[DS99] Martin J Dunwoody and Micah E Sageev, *JSJ- splittings for finitely presented groups over slender groups*, Invent. Math. **135** (1999), no. 1, 25–44.

[Dun98] Martin Dunwoody, *Cutting up graphs*, July 1998, research seminar at Computational and Geometric Aspects of Modern Algebra conference, Heriot-Watt University, Edinburgh.

[ECH+92] David B A Epstein, James W Cannon, Derek F Holt, Silvio V F Levy, Michael S Paterson, and William P Thurston, *Word processing in groups*, Jones and Bartlett, 1992.

[EIFZ96] David B A Epstein, Anthony R Iano-Fletcher, and Uri Zwick, *Growth functions and automatic groups*, Experimental Mathematics **5** (1996), no. 4, 297–315.

[FP97] K Fujiwara and Panos Papasoglu, *JSJ decompositions of finitely presented groups and complexes of groups*, preprint, Apr 1997.

[GdlH90] Etienne Ghys and Pierre de la Harpe (eds.), *Sur les groupes hyperboliques d'après Mikhael Gromov*, Progress in Mathematics, vol. 83, Birkhäuser, 1990.

[Ger] Victor Gerasimov, *Detecting connectedness of the boundary of a hyperbolic group*, preprint.

[GHV91] Etienne Ghys, André Haefliger, and Alberto Verjovsky (eds.), *Group theory from a geometrical viewpoint*, World Scientific, 1991.

[Ly00] Olivier Ly, *Etude algorithmique de complexes simpliciaux infinis*, July 2000, Ph.D. thesis, Computer Science department, University of Bordeaux.

[Pap95] Panos Papasoglu, *Strongly geodesically automatic groups are hyperbolic*, Invent Math **121** (1995), no. 2, 323–334.

[RS95] E Rips and Zlil Sela, *Cyclic splittings of finitely presented groups and the canonical JSJ decomposition*, preprint, 1995.

[Sel95] Zlil Sela, *The isomorphism problem for hyperbolic groups I*, Annals of Mathematics (2) **141** (1995), no. 2, 217–283.

[Ser83] Jean-Pierre Serre, *Arbres, amalgames,* $\mathfrak{sl}_2$, 1983.

[Sta68] John R Stallings, *On torsion-free groups with infinitely many ends*, Annals of Mathematics **88** (1968), 312–334.

[Str90] Ralph Strebel, *Small cancellation diagrams*, pp. 225–273, Vol. 83 of Ghys and de la Harpe [GdlH90], 1990.

[SW] Peter Scott and Terry Wall, *Topological methods in group theory*, Homological Group Theory, London Math Soc Lecture Notes, number 36,Cambridge University Press,1979.

[Swa96] Gadde A Swarup, *On the cut point conjecture*, Electron. Res. Announc. Amer. Math. Soc. **2** (1996), 98–100.

Index of Definitions