Filling Functions

Notes for an advanced course on
The Geometry of the Word Problem
for Finitely Generated Groups

Centre de Recerca Matemàtica
Barcelona

T.R.Riley

July 2005
Revised February 2006
# Contents

Notation vi

1 Introduction 1

2 Filling functions 5
   2.1 Van Kampen diagrams ........................................... 5
   2.2 Filling functions via van Kampen diagrams .................... 6
   2.3 Example: combable groups ....................................... 10
   2.4 Filling functions interpreted algebraically .................. 15
   2.5 Filling functions interpreted computationally ............... 16
   2.6 Filling functions for Riemannian manifolds .................. 21
   2.7 Quasi-isometry invariance ..................................... 22

3 Relationships between filling functions 25
   3.1 The Double Exponential Theorem ............................... 26
   3.2 Filling length and duality of spanning trees in planar graphs 31
   3.3 Extrinsic diameter versus intrinsic diameter .................. 35
   3.4 Free filling length ............................................. 35

4 Example: nilpotent groups 39
   4.1 The Dehn and filling length functions ........................ 39
   4.2 Open questions .................................................. 42

5 Asymptotic cones 45
   5.1 The definition .................................................. 45
   5.2 Hyperbolic groups ............................................... 47
   5.3 Groups with simply connected asymptotic cones .............. 53
   5.4 Higher dimensions ............................................... 57

Bibliography 68
Notation

\[ f, g : [0, \infty) \to [0, \infty) \text{ satisfy } f \preceq g \text{ when there exists } C > 0 \text{ such that } f(n) \leq Cg(Cn + C) + Cn + C \text{ for all } n, \text{ satisfy } f \simeq g \text{ when } g \preceq f, \text{ and satisfy } f \approx g \text{ when } f \leq g \text{ and } g \leq f. \]

\[ \preceq, \succeq, = \]

\[ a^b, a^{-b}, [a, b] \]

\[ \text{Cay}^1(G, X) \]

\[ \text{Cay}^2(\mathcal{P}) \]

\[ \mathbb{D}^n \]

\[ \text{Diam}(\Gamma) \]

\[ d_X \]

\[ \ell(w) \]

\[ \ell(\partial \Delta) \]

\[ \mathbb{N}, \mathbb{R}, \mathbb{Z} \]

\[ R^{-1} \]

\[ \mathbb{S}^n \]

\[ \text{Star}(\Delta_0) \]

\[ (T, T^*) \]

\[ w^{-1} \]

\[ (X | R) \]

\[ X^{-1} \]

\[ (X \cup X^{-1})^* \]

\[ \varepsilon \]

Diagram measurements

\[ \text{Area}(\Delta) \]

\[ \text{DGL}(\Delta) \]

\[ \text{EDiam}(\Delta) \]

\[ \text{FL}(\Delta) \]

\[ \text{GL}(\Delta) \]

\[ \text{IDiam}(\Delta) \]

\[ \text{Rad}(\Delta) \]

the number of 2-cells in \( \Delta \)

\[ \text{min} \left\{ \text{Diam}(T) + \text{Diam}(T^*) \mid T \text{ a spanning tree in } \Delta^{(1)} \right\} \]

the diameter of \( \Delta \) as measured in the Cayley 2-complex

the filling length of \( \Delta \) – see Section 2.2

the diameter of the 1-skeleton of the dual of \( \Delta \)

the diameter of the 1-skeleton of \( \Delta \)

max \{ \rho(a, \partial \Delta) \mid \text{vertices } a \text{ of } \Delta \} \text{ as measured in } \Delta^{(1)}
**Notation**

*Filling functions* – see Section 2.2 unless otherwise indicated

<table>
<thead>
<tr>
<th>Function</th>
<th>Domain</th>
<th>Range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area : $\mathbb{N} \to \mathbb{N}$</td>
<td>the Dehn function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DGL : $\mathbb{N} \to \mathbb{N}$</td>
<td>the simultaneous diameter and gallery length function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDiam : $\mathbb{N} \to \mathbb{N}$</td>
<td>the extrinsic diameter function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FL : $\mathbb{N} \to \mathbb{N}$</td>
<td>the filling length function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GL : $\mathbb{N} \to \mathbb{N}$</td>
<td>the gallery length function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDiam : $\mathbb{N} \to \mathbb{N}$</td>
<td>the intrinsic diameter function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDiam : $\mathbb{N} \to \mathbb{N}$</td>
<td>the upper intrinsic diameter function – see Section 3.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rad : $\mathbb{N} \to \mathbb{N}$</td>
<td>the radius function – see Section 5.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rad : $\mathbb{N} \to \mathbb{N}$</td>
<td>the upper radius function – see Section 5.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The Word Problem was posed by Dehn [33] in 1912. He asked, given a group, for a systematic method (in modern terms, an algorithm) which, given a finite list (a word) of basic group elements (generators and their formal inverses), declares whether or not their product is the identity. One of the great achievements of 20th century mathematics was the construction by Boone [13] and Novikov [74] of finitely presentable groups for which no such algorithm can exist. However, the Word Problem transcends its origins in group theory and rises from defeat at the hands of decidability and complexity theory, to form a bridge to geometry – to the world of isoperimetry and curvature, local and large-scale invariants, as brought to light most vividly by Gromov [62].

So where does geometry enter? Groups act: given a group, one seeks a space on which it acts in as favourable a manner as possible, so that the group can be regarded as a discrete approximation to the space. Then a dialogue between geometry and algebra begins. And where can we find a reliable source of such spaces? Well, assume we have a finite generating set $X$ for our group $G$. (All the groups in this study will be finitely generated.) For $x, y \in G$, define the distance $d_X(x, y)$ in the word metric $d_X$ to be the length of the shortest word in the generators and their formal inverses that represents $x^{-1}y$ in $G$. Then

$$d_X(zx, zy) = d_X(x, y)$$

for all $x, y, z \in G$, and so left multiplication is action of $G$ on $(G, d_X)$ by isometries.

However $(G, d_X)$ is discrete and so appears geometrically emaciated (“boring and uneventful to a geometry’s eye” in the words of Gromov [62]). Inserting a directed edge labelled by $a$ from $x$ to $y$ whenever $a \in X$ and $xa = y$ gives a skeletal structure known as the Cayley graph $\text{Cay}^1(G, X)$. If $G$ is given by a finite presentation $\mathcal{P} = \langle X \mid R \rangle$ we can go further: attach flesh to the dry bones of $\text{Cay}^1(G, X)$ in the form of 2-cells, their boundary circuits glued along edge-loops around which read words in $R$. The result is a simply connected space $\text{Cay}^2(\mathcal{P})$ on which $G$ acts geometrically (that is, properly, discontinuously and cocompactly) that is known as the Cayley 2-complex – see Section 2.1 for a precise definition. Further enhancements may be possible. For example, one could seek to attach
cells of dimension 3 or above to kill off higher homotopy groups or so that the complex enjoys curvature conditions such as the CAT(0) property (see [15] or [22]).

Not content with the combinatorial world of complexes, one might seek smooth models for $G$. For example, one could realise $G$ as the fundamental group of a closed manifold $M$, and then $G$ would act geometrically on the universal cover $\tilde{M}$. (If $G$ is finitely presentable then $M$ can be taken to be four dimensional – see [19, A.3].) Wilder non-discrete spaces, asymptotic cones, arise from viewing $(G,d_X)$ from increasingly distant vantage points (i.e. scaling the metric to $d_X/s_n$ for some sequence of reals with $s_n \to \infty$) and recording recurring patterns using the magic of a non-principal ultrafilter. Asymptotic cones discard all small-scale features of $(G,d_X)$; they are the subject of Chapter 5.

Filling functions, the subject of this study, capture features of discs spanning loops in spaces. The best known is the classical isoperimetric function for Euclidean space $\mathbb{R}^n$ – any loop of length $\ell$ can be filled with a disc of area at most a constant times $\ell^2$. To hint at how filling functions enter the world of discrete groups we mention a related algebraic result concerning the group $\mathbb{Z}^m$, the integer lattice in $m$-dimensional Euclidean space, generated by $x_1,\ldots,x_m$. If $w$ is a word of length $n$ on $\{x_1^\pm 1,\ldots,x_m^\pm 1\}$ and $w$ represents the identity in $\mathbb{Z}^m$ then, by cancelling off pairs $x_i x_i^{-1}$ and $x_i^{-1} x_i$, and by interchanging adjacent letters at most $n^2$ times, $w$ can be reduced to the empty word.

This qualitative agreement between the number of times the commutator relations $x_i x_j = x_j x_i$ are applied and the area of fillings is no coincidence; such a relationship holds for all finitely presented groups, as will be spelt out in Theorem 2.6.1 (The Filling Theorem). The bridge between continuous maps of discs filling loops in spaces and this computational analysis of reducing words is provided by van Kampen (or Dehn) diagrams. The Cayley 2-complex of the presentation

$$\mathcal{P} := \langle x_1, \ldots, x_m \mid [x_i, x_j], \forall i, j \in \{1, \ldots, m\} \rangle$$

of $\mathbb{Z}^m$ is the 2-skeleton of the standard decomposition of $\mathbb{R}^m$ into an infinite array of $m$-dimensional unit cubes. A word $w$ that represents 1 in $\mathcal{P}$ (or, indeed, in any finite presentation $\mathcal{P}$) corresponds to an edge-loop in $\text{Cay}^2(\mathcal{P})$. As Cayley 2-complexes are simply connected such edge-loops can be spanned by filling discs and, in this combinatorial setting, it is possible and appropriate to take these homotopy discs to be combinatorial maps of planar 2-complexes homeomorphic to (possibly singular) 2-discs into $\text{Cay}^2(\mathcal{P})$. A van Kampen diagram for $w$ is a graphical demonstration of how it is a consequence of the relations $\mathcal{R}$ that $w$ represents 1; Figure 2.4 is an example. So the Word Problem amounts to determining whether or not a word admits a van Kampen diagram. (See van Kampen’s Lemma: Lemma 2.4.1.)

Filling functions for finite presentations of groups (defined in Chapter 2) record geometric features of van Kampen diagrams. The best known is the Dehn function (or minimal isoperimetric function) of Madlener & Otto [69] and Gersten [47]; it concerns area – that is, number of 2-cells. In the example of $\mathbb{Z}^m$ this equates to the number of times commutator relations have to be applied to reduce $w$ to the empty word – in this sense (that is, in the Dehn proof system – see Section 2.5) the Dehn function can also be
understood as a non-deterministic TIme complexity measure of the Word Problem for $P$ – see Section 2.5. The corresponding Space complexity measure is called the filling length function of Gromov [62]. It has a geometric interpretation – the filling length of a loop $\gamma$ is the infimal length $L$ such that $\gamma$ can be contracted down to its base vertex through loops of length at most $L$. Other filling functions we will encounter include the gallery length, and intrinsic and extrinsic diameter functions. All are group invariants in that whilst they are defined with respect to specific finite presentations, their qualitative growth depends only on the underlying group; moreover, they are quasi-isometry invariants, that is, qualitatively they depend only on the large-scale geometry of the group – see Section 2.7 for details.

In Chapter 3 we examine the interplay between different filling functions – this topic bares some analogy with the relationships that exist between different algorithmic complexity measures and, as with that field, many open questions remain. The example of nilpotent groups discussed in Chapter 4 testifies to the value of simultaneously studying multiple filling functions. Finally, in Chapter 5, we discuss how the geometry and topology of the asymptotic cones of a group $G$ relates to the filling functions of $G$.

Acknowledgements. These notes build on and complement [19], [48] and [62, Chapter 5] as well as the other two sets of notes in this volume [15, 99]. For Chapter 5 I am particularly indebted to the writings of Drutu [34, 35, 37, 38]. This is not intended to be a balanced or complete survey of the literature, but rather a brief tour heavily biased towards areas in which the author has been involved.

If I have done any justice to this topic then the influence of Martin Bridson and Steve Gersten should shine through. I am grateful to them both for stimulating collaborations, for their encouragement, and for communicating their deep vision for the subject. I thank Emina Alibegovic, Will Dison, Cornelia Drutu Badea, Steve Gersten, Mark Sapir and Hamish Short for comments on earlier drafts. I also thank the Centre de Recerca Matemàtica in Barcelona for their hospitality during the writing of this study, José Burillo and Enric Ventura for organising the workshop, and the NSF for partial support via grant DMS–0540830.
Chapter 2

Filling functions

2.1 Van Kampen diagrams

The presentation 2-complex of

\[ P = \langle X \mid R \rangle = \langle x_1, \ldots, x_m \mid r_1, \ldots, r_n \rangle \]

is constructed as shown in Figure 2.1: take \( m \) oriented edges, labelled by \( x_1, \ldots, x_m \), identify all the vertices to form a rose, and then attach 2-cells \( C_1, \ldots, C_n \), where \( C_i \) has \( \ell(r_i) \) edges, by identifying the boundary circuit of \( C_i \) with the edge-path in the rose along which one reads \( r_i \). The Cayley 2-complex \( \text{Cay}^2(P) \) is the universal cover of the presentation 2-complex of \( P \). The example of the free abelian group of rank 2, presented by \( \langle a, b \mid [a, b] \rangle \) is shown in Figure 2.2.

![Figure 2.1: The presentation 2-complex for \( \langle x_1, \ldots, x_m \mid r_1, \ldots, r_n \rangle \) ](image)

The edges of \( \text{Cay}^2(P) \) inherit labels and orientations from the presentation 2-complex and the 1-skeleton of \( \text{Cay}^2(P) \) is the Cayley graph \( \text{Cay}^1(P) \) (cf. Definition 1.3 in [99]). Identifying the group \( G \) presented by \( P \) with the 0-skeleton of \( \text{Cay}^1(P) \), the path metric on \( \text{Cay}^1(P) \) in which each edge has length 1 agrees with the word metric \( d_X \) on \( G \).

A word \( w \in (X \cup X^{-1})^* \) is null-homotopic when it represents the identity. To such a \( w \) one can associate an edge-circuit \( \eta_w \) in \( \text{Cay}^2(P) \) based at some (and hence any) fixed
vertex \(v\), so that around \(\eta_w\), starting from \(v\), one reads \(w\). A \(\mathcal{P}\)-van Kampen diagram for \(w\) is a combinatorial map \(\pi : \Delta \to \text{Cay}^2(\mathcal{P})\), where \(\Delta\) is \(S^2 \setminus e_{\infty}\) for some combinatorial 2-complex \(S^2\) homeomorphic to the 2-sphere and some open 2-cell \(e_{\infty}\) of \(S^2\), and the edge-circuit \(\pi|_{\partial \Delta}\) based at vertex \(\star\) on \(\partial \Delta\), agrees with \(\eta_w\). (A map \(X \to Y\) between complexes is combinatorial if for all \(n\), it maps the interior of each \(n\)-cell in \(X\) homeomorphically onto the interior of an \(n\)-cell in \(Y\).) The base vertex \(\star\) of \(\Delta\) should not be ignored; we will see it plays a crucial role. Van Kampen’s Lemma (see Section 2.4) says, in particular, that \(w \in (X \cup X^{-1})^*\) is null-homotopic if and only if it admits a \(\mathcal{P}\)-van Kampen diagram.

It is convenient to have the following alternative definition in which no explicit reference is made to \(\pi\). A \(\mathcal{P}\)-van Kampen diagram \(\Delta\) for \(w\) is a finite planar contractible combinatorial 2-complex with directed and labelled edges such that anti-clockwise around \(\partial \Delta\) one reads \(w\), and around the boundary of each 2-cell one reads a cyclic conjugate (that is, a word obtained by cyclically permuting letters) of a word in \(R^{\pm 1}\). From this point of view a \(\mathcal{P}\)-van Kampen diagram for \(w\) is a filling of a planar edge-loop labelled by \(w\) in the manner of a jigsaw-puzzle, where the pieces (such as those pictured in Figure 2.3) correspond to defining relations and are required to be fitted together in such a way that orientations and labels match. The analogy breaks down in that the pieces may be distorted or flipped. Figures 2.4, 2.7, 3.8 and 4.1 show examples of van Kampen diagrams.

The two definitions are, in effect, equivalent because, given the first, edges of \(\Delta\) inherit directions and labels from \(\text{Cay}^2(\mathcal{P})\), and given the second, there is a combinatorial map \(\pi : \Delta \to \text{Cay}^2(\mathcal{P})\), uniquely determined on the 1-skeleton of \(\Delta\), that sends \(\star\) to \(v\) and preserves the labels and directions of edges.

### 2.2 Filling functions via van Kampen diagrams

Suppose \(\mathcal{P} = \langle X | R \rangle\) is a finite presentation for a group and \(\pi : \Delta \to \text{Cay}^2(\mathcal{P})\) is a van Kampen diagram. Being planar, \(\Delta\) has a dual \(\Delta^*\) — the 2-complex with a vertex, edge
2.2. Filling functions via van Kampen diagrams

Figure 2.3: The defining relations in \( \langle a, b, c \mid [a, b], [b, c], [c, a] \rangle \).

Figure 2.4: A van Kampen diagram for \( ba^{-1}ca^{-1}bcb^{-1}ca^{-1}b^{-1}c^{-1}b^{-1}b^{-2}acac^{-1}a^{-1}c^{-1}aba \) in \( \langle a, b, c \mid [a, b], [b, c], [c, a] \rangle \).

and face dual to each face, edge and vertex in \( \Delta \) (including a vertex dual to the face at infinity, the complement of \( \Delta \) in the plane).

For a finite connected graph \( \Gamma \) define

\[
\text{Diam}(\Gamma) := \max \{ \rho(a, b) \mid \text{vertices } a, b \text{ of } \Gamma \}
\]

where \( \rho \) is the combinatorial metric on \( \Gamma \) – the path metric in which each edge is given length 1. Define the area, intrinsic diameter, extrinsic diameter, gallery length, filling
length and DGL of $\pi : \Delta \to \text{Cay}^2(\mathcal{P})$ by

\[
\begin{align*}
\text{Area}(\Delta) & = \# \text{ 2-cells in } \Delta \\
\text{IDiam}(\Delta) & = \text{Diam}(\Delta^{(1)}) \\
\text{EDiam}(\Delta) & = \max \{ d_\mathcal{P}(\pi(a), \pi(b)) \mid \text{ vertices } a, b \text{ of } \Delta \} \\
\text{GL}(\Delta) & = \text{Diam}(\Delta^{(1)} \star (1)) \\
\text{FL}(\Delta) & = \min \left\{ L \mid \exists \text{ a shelling } (\Delta_i) \text{ of } \Delta \text{ such that } \max_i \ell(\partial \Delta_i) \leq L \right\} \\
\text{DGL}(\Delta) & = \min \left\{ \text{Diam}(T) + \text{Diam}(T^*) \mid T \text{ a spanning tree in } \Delta^{(1)} \right\}.
\end{align*}
\]

These are collectively referred to as diagram measurements. Note that IDiam measures diameter in the 1-skeleton of $\Delta$ and EDiam in the Cayley graph of $\mathcal{P}$.

The definitions of FL($\Delta$) and DGL($\Delta$) require further explanation. A shelling of $\Delta$ is, roughly speaking, a combinatorial null-homotopy of $\Delta$ down to its base vertex $\star$. More precisely, a shelling of $\Delta$ is sequence $S = (\Delta_i)$ of diagrams $(\Delta_i)_{i=0}^m$ with $\Delta_0 = \Delta$ and $\Delta_m = \star$ and such that $\Delta_{i+1}$ is obtained from $\Delta_i$ by one of the shelling moves defined below (illustrated in Figure 2.5). Moreover, the base vertex $\star$ is preserved throughout $(\Delta_i)$ -- that is, in every 1-cell collapse $e^0 \neq \star$, and in every 1-cell expansion on $\Delta_i$ where $e^0 = \star$ a choice is made as to which of the two copies of $e^0$ is to be $\star$ in $\Delta_{i+1}$.

- 1-cell collapse. Remove a pair $(e^1, e^0)$ where $e^1$ is a 1-cell with $e^0 \in \partial e^1$ and $e^1$ is attached to the rest of $\Delta_i$ only by one of its end vertices $\neq e^0$. (Call such an $e^1$ a spike.)
- 1-cell expansion. Cut along some 1-cell $e^1$ in $\Delta_i$ that has a vertex $e^0$ in $\partial \Delta_i$, in such a way that $e^0$ and $e^1$ are doubled.
- 2-cell collapse. Remove a pair $(e^2, e^1)$ where $e^2$ is a 2-cell which has some edge $e^1 \in (\partial e^2 \cap \partial \Delta_i)$. The effect on the boundary circuit is to replace $e^1$ with $\partial e^2 \setminus e^1$.

There are natural combinatorial maps $\Delta_i \to \Delta$ whose restrictions to the interiors of $\Delta_i$ are injective, and which map the boundary circuits of $\Delta_i$, labelled by words $w_i$, to a sequence of edge-loops contracting to $\star$ as illustrated schematically in Figure 2.6.
2.2. Filling functions via van Kampen diagrams

Returning to the definition of $DGL(\Delta)$, if $T$ is a spanning tree in the 1-skeleton of $\Delta$ then define $T^*$ to be the subgraph of the 1-skeleton of $\Delta^*$ consisting of all the edges dual to edges of $\Delta^{(1)} \setminus T$. The crucial property of $T^*$ is:

**Exercise 2.2.1.** $T^*$ is a spanning tree in the 1-skeleton of $\Delta^*$.

We now define an assortment of *filling functions* for $\mathcal{P}$.

- The *Dehn function* $\text{Area} : \mathbb{N} \to \mathbb{N}$,
- the *intrinsic diametric function* $\text{IDiam} : \mathbb{N} \to \mathbb{N}$,
- the *extrinsic diametric function* $\text{EDiam} : \mathbb{N} \to \mathbb{N}$,
- the *gallery length function* $\text{GL} : \mathbb{N} \to \mathbb{N}$,
- the function $DGL : \mathbb{N} \to \mathbb{N}$,
- and the *filling length function* $\text{FL} : \mathbb{N} \to \mathbb{N}$

of $\mathcal{P}$ are defined by

\[
\begin{align*}
M(w) &:= \min \{ M(\Delta) \mid \Delta \text{ a van Kampen diagram for } w \}, \\
M(n) &:= \max \{ M(w) \mid \text{words } w \text{ with } \ell(w) \leq n \text{ and } w = 1 \text{ in } G \},
\end{align*}
\]

where $M$ is Area, IDiam, EDiam, GL, DGL and FL respectively. (The meanings of $M(\Delta)$, $M(w)$ and $M(n)$ depend on their arguments: diagram, null-homotopic word, or natural number; the potential ambiguity is tolerated as it spares us from an overload of terminology.)

An *isoperimetric function* (respectively, *isodiametric function*) for $\mathcal{P}$ is any $f : \mathbb{N} \to \mathbb{N}$ such that $\text{Area}(n) \leq f(n)$ (respectively, $\text{IDiam}(n) \leq f(n)$) for all $n$. Chapter 5 of [62] is a foundational reference on filling functions. Many other references to isoperimetric functions, Dehn functions and isodiametric functions appear in the literature; [19] and [48] are surveys. Dehn functions were introduced by Madlener & Otto [69] and, independently, by Gersten [47]. The filling length function of Gromov [62] is discussed extensively in
Exercise 2.2.3. Show that for a finite presentation, \(d(u, v) := \text{Area}(u^{-1}v)\) defines a metric on any set of reduced words, all representing the same group element.

Exercise 2.2.4. Suppose \(u, v \in (X^+)\) represent the same group element and \(\Delta \) is a van Kampen diagram for \(uv^{-1}\) with two distinguished boundary vertices \(\star_1, \star_2\) separating the \(u\)- and \(v\)-portions of the boundary circuit. Define \(F, L(u, v, \Delta)\) to be the least \(L\) such that there is a combinatorial homotopy of \(u\) to \(v\) across \(\Delta\) through paths of length at most \(L\) from \(\star_1\) to \(\star_2\). (More formally, such a combinatorial homotopy is a sequence of van Kampen diagrams \((\Delta_i)_{i=0}^n\) with \(\Delta_0 = \Delta\) and \(\Delta_n\) a simple edge-path along which one reads \(w_2\), and such that \(\Delta_{i+1}\) is obtained from \(\Delta_i\) by 1-cell collapse, 1-cell expansion, or 2-cell collapse in such a way that the \(v\)-portion of the boundary words \(\partial \Delta_i\) is not broken.) Define \(F, L(u, v)\) to be the minimum of \(F, L(u, v, \Delta)\) over all van Kampen diagrams \(\Delta\) for \(uv^{-1}\) and define \(F, L : \mathbb{N} \to \mathbb{N}\) by letting \(F, L(n)\) be the maximum of \(F, L(u, v)\) over all \(u, v\) of length at most \(n\) that represent the same element of the group. Show that

1. \(\text{FL}(2n) \leq F, L(n)\) and \(\text{FL}(2n + 1) \leq F, L(n + 1)\) for all \(n\).

2. \(F, L(u, v)\) defines a metric on any set of words that all represent the same group element.

### 2.3 Example: combable groups

Combable groups form a large and well studied class that includes all automatic groups [42] and (hence) all hyperbolic groups. A normal form for a group \(G\) with finite generating set \(X\) is a section \(\sigma : G \to (X \cup X^{-1})^\ast\) of the natural surjection \((X \cup X^{-1})^\ast \to G\). In other words, a normal form is a choice of representative \(\sigma_g = \sigma(g)\) for each group element \(g\). View \(\sigma_g\) as a continuous path \([0, \infty) \to \text{Cay}^1(G, X)\) from the identity to \(g\) (the “combing line of \(g\)”), travelling at unit speed from the identity until time \(\ell(\sigma_g)\) when it halts for evermore at \(g\).
Following [16, 48], we say $\sigma$ synchronously $k$-fellow-travels when
\[
\forall g, h \in G, \left( d_X(g, h) = 1 \implies \forall t \in \mathbb{N}, d_X(\sigma_g(t), \sigma_h(t)) \leq k \right).
\]
Define a reparametrisation $\rho$ to be an unbounded function $\mathbb{N} \to \mathbb{N}$ such that $\rho(0) = 0$ and $\rho(n + 1) \in \{\rho(n), \rho(n) + 1\}$ for all $n$. We say $\sigma$ asynchronously $k$-fellow travels when for all $g, h \in G$ with $d_X(g, h) = 1$, there exist reparametrisations $\rho$ and $\rho'$ such that
\[
\forall t \in \mathbb{N}, d_X(\sigma_g(\rho(t)), \sigma_h(\rho'(t))) \leq k.
\]
(Note $\rho$ and $\rho'$ both depend on both $g$ and $h$.)

We say $(G, X)$ is (a)synchronously combable when, for some $k \geq 0$, there is an (a)synchronous $k$-fellow-travelling normal form $\sigma$ for $G$. Define the length function $L : \mathbb{N} \to \mathbb{N}$ of $\sigma$ by:
\[
L(n) := \max \left\{ \ell(\sigma_g) \mid d_X(1, g) \leq n \right\}.
\]

**Examples 2.3.1.**

1. **Finite groups.** Suppose $G$ is a finite group with finite generating set $X$. Let $k$ be the diameter of $\text{Cay}(G, X)$. If for all $g \in G$ we take $\sigma_g$ to be any word representing $g$ then $\sigma$ is synchronously $k$-fellow travelling. Moreover, if $\sigma_g$ is a choice of geodesic (i.e. minimal length) word representing $g$ then $\sigma$ has length function satisfying $L(n) \leq \min \{n, k\}$ for all $n$.

2. The (unique) geodesic words representing elements in $F_m = \langle a_1, \ldots, a_m \mid \rangle$ form a synchronous 1-fellow-travelling combing. The words
\[
\{ a_1^{r_1} a_2^{r_2} \ldots a_m^{r_m} \mid r_1, \ldots, r_m \in \mathbb{Z} \}
\]
comprise a synchronous 2-fellow travelling combing of
\[
\mathbb{Z}^m = \langle a_1, \ldots, a_m \mid [a_i, a_j], \forall 1 \leq i < j \leq m \rangle.
\]
In both cases $L(n) = n$ for all $n$.

3. The family of groups $BS(m, n)$ with presentations $\langle a, b \mid b^{-1} a^n b = a^r \rangle$ are named in honour of Baumslag and Solitar who studied them in [7]. $BS(1, 2)$ is often referred to as “the Baumslag–Solitar group” despite repeated public insistence from Baumslag that this is an inappropriate attribution.

The words
\[
\{ b^s a^r \mid r, s \in \mathbb{Z}, u \in \{ab^{-1}, b^{-1}\}^*, \text{ the first letter of } u \text{ is not } b^{-1} \}
\]
define an asynchronous combing of $BS(1, 2)$. In fact, the normal form element for $g \in BS(1, 2)$ results from applying the rewriting rules.
Figure 2.7: A van Kampen diagram for \([a, a^b] \) in \( \langle a, b \mid b^{-1}ab = a^2 \rangle \) and a portion of \( \text{Cay}^2(\mathcal{P}) \).

\[
\begin{align*}
ab &\mapsto ba^2 & a^{-1} &\mapsto e \\
a^{-1}b &\mapsto ba^{-2} & a^{-1}a &\mapsto e \\
a^2b^{-1} &\mapsto b^{-1}a & bb^{-1} &\mapsto e \\
a^{-1}b^{-1} &\mapsto ab^{-1}a^{-1} & b^{-1}b &\mapsto e
\end{align*}
\]

to any word representing \( g \). See [42, Chapter 7] for more details.

4. **Automatic groups.** A group \( G \) with finite generating set \( X \) is **synchronously automatic** when it admits a synchronous combing \( \sigma \) such that the set of normal forms \( \{ \sigma_g \mid g \in G \} \) comprise a regular language. The foundational reference is [42] – there, a group is defined to be synchronous automatic when its multiplication and inverse operations are governed by finite state automata, and agreement of the two definitions is Theorem 2.3.5.

The \( \sigma_g \) are quasi-geodesics and so there is a constant \( C \) such that \( L(n) \leq Cn + C \) for all \( n \). Synchronously automatic groups include all finitely generated virtually abelian groups, all hyperbolic groups, all the braid groups \( B_n \), many 3-manifold groups, and the mapping class groups of closed surfaces with finitely many punctures – see [42, 73].

The larger class of **asynchronously automatic groups**, defined in [42], can be characterised (see Theorems 7.2.8 and 7.3.6 in [42]) as the asynchronously combable groups \( G \) such that \( \{ \sigma_g \mid g \in G \} \) is regular and admits a function \( D : \mathbb{N} \to \mathbb{N} \) such
that for all \( g \in G, r, s \geq 0, t \geq D(r) \), if \( s + t \leq \ell(\sigma_g) \) then \( d_X(\sigma_g(s), \sigma_g(s + t)) > r \). Their length functions satisfy \( L(n) \leq C^n \) for some \( C > 0 \) (see e.g. [20, Lemma 2.3]). Thurston [42, Section 7.4] showed that \( BS(1, 2) \) is asynchronously automatic: the language of the asynchronous combing given above is regular, but, on account of the exponential Dehn function we will establish in Proposition 3.1.6, \( BS(1, 2) \) is not synchronously automatic.

For references of an introductory nature and for open questions see, e.g. [6, 43, 49, 50, 75].

5. CAT(0) groups. Paths in a Cayley graph of a CAT(0) group that run close to CAT(0) geodesics can be used to define a synchronous combing with length function \( L(n) \approx n \). See the proof of Proposition 1.6 in [22, III.1] for more details.

Further classes of groups can be defined by specifying other grammatical, length function or geometric constraints. See, for example, [21, 43, 58] for more details.

Exercise 2.3.2. Check that the words in Example 2.3.1(3) define an asynchronous combing of \( BS(1, 2) \) and show that its length function satisfies \( L(n) \approx 2^n \). Estimate the fellow-travelling constant.

Exercise 2.3.3. Show that if \( X \) and \( X' \) are two finite generating sets for a group \( G \) then \( (G, X) \) is (a)synchronously combable if and only if \( (G, X') \) is. Moreover, show that if \( L(n) \) is a length function for an (a)synchronous combing of \( (G, X) \) then there is an (a)synchronous combing of \( (G, X') \) with length function \( L'(n) \) satisfying \( L(n) \approx L'(n) \).

Theorem 2.3.4 (cf. [42, 48, 55]). Suppose a group \( G \) with finite generating set \( X \) admits a combing \( \sigma \) that (a)synchronously \( k \)-fellow-travels. Let \( L : \mathbb{N} \to \mathbb{N} \) be the length function of \( \sigma \). Then there exists \( C > 0 \) and a finite presentation \( \mathcal{P} = \langle X \mid R \rangle \) for \( G \) for which

\[
\text{Area}(n) \leq Cn L(n) \quad \text{and} \quad \text{EDiam}(n) \leq \text{IDiam}(n) \leq \text{FL}(n) \leq C n
\]

for all \( n \). Moreover, these bounds are realisable simultaneously.

Proof. Suppose \( w \) is a word representing 1 in \( G \). We will construct the cockleshell van Kampen diagram \( \Delta \) for \( w \), illustrated in Figure 2.8. We start with a circle \( C \) in the plane subdivided into \( \ell(w) \) edges, directed and labelled so that anticlockwise from the base vertex \( \star \) one reads \( w \). Then we join each vertex \( v \) to \( \star \) by an edge-path labelled by the word \( \sigma_{g_v} \), where \( g_v \) is the group element represented by the prefix of the word \( w \) read around \( C \) between \( \star \) and \( v \). For each pair \( u, v \) of adjacent vertices on \( C \), we use the fact that \( \sigma_{g_u} \) and \( \sigma_{g_v} \) asynchronously \( k \)-fellow travel (with respect to reparametrisations \( \rho \) and \( \rho' \)) to construct a ladder whose rungs are paths of length at most \( k \); these would be straight if our combing was synchronous, but may be skew in the asynchronous case. A rung connects \( \sigma_{g_u}(\rho(t)) \) to \( \sigma_{g_v}(\rho'(t)) \) for \( t = 1, 2, \ldots \). Duplicate rungs may occur on account of \( \rho \) and \( \rho' \) simultaneously pausing, that is, \( \rho(m) = \rho(m + 1) \) and \( \rho'(m) = \rho'(m + 1) \) for some \( m \); counting each of these only once, there are a total of at most \( \ell_{g_u} + \ell_{g_v} \) rungs and hence at most the same number of 2-cells in the ladder.
Subdivide the rung from $\sigma_{g_u}(\rho(t))$ to $\sigma_{g_v}(\rho'(t))$ into $d_X(\sigma_{g_u}(\rho(t)), \sigma_{g_v}(\rho'(t))) \leq k$ edges and label it by a shortest word representing the same element in $G$ as $[\sigma_{g_u}(\rho(t))]^{-1}\sigma_{g_v}(\rho'(t))$.

![Figure 2.8: The cockleshell diagram filling an edge-loop in an asynchronously combable group.](image)

Obtain $\Delta$ by regarding every simple edge-circuit in a ladder to be the boundary circuit of a 2-cell; words around such edge-circuits have length at most $2k+2$ and represent the identity in $G$. So, defining $R$ to be the set of null-homotopic words of length at most $2k+2$, we find that $\mathcal{P} := \langle X \mid R \rangle$ is a finite presentation for $G$ and $\Delta$ is a $\mathcal{P}$-van Kampen diagram for $w$. We get the asserted bound on $\text{Area}(n)$ from the observation

$$\text{Area}(\Delta) \leq 2 \ell(w) L(\lfloor \ell(w)/2 \rfloor).$$

We will only sketch a proof of the linear upper bound on $\text{FL}(n)$. Consider paths $p$ connecting pairs of points on $\partial \Delta$, that run along successive rungs of adjacent ladders, no two rungs coming from the same ladder. Observe that there is such a $p$ for which the region enclosed by $p$ and a portion of the boundary circuit of $\Delta$ not passing through $\star$ is a topological 2-disc that includes no rungs in its interior. Shell away this region using 1-cell and 2-cell collapse moves. Repeating this process and performing 1-cells collapse moves will shell $\Delta$ down to $\star$ and the boundary loops of the intermediate diagrams will run (roughly speaking) perpendicular to the combing lines.

The inequalities $\text{EDiam}(n) \leq \text{IDiam}(n) \leq \text{FL}(n)$ are straightforward and are not specific to combable groups – see Chapter 3.

**Remark 2.3.5.** Figure 2.8 suggests that $\text{DGL}(\Delta)$ and (hence) $\text{GL}(\Delta)$ are both at most a constant $C$ times $L(\lfloor \ell(w)/2 \rfloor)$, taking $T$ (a spanning tree in the 1-skeleton of $\Delta$) to be the union of all the combing lines, and hence that there is a further simultaneously realisable bound of

$$\text{GL}(n) \leq \text{DGL}(n) \leq C L(n), \quad (2.1)$$

for some constant $C$. But the picture is misleading – some of the rungs could have zero length and hence $T$ could fail to be a tree. A way to circumvent this problem is to fatten...
2.4. Filling functions interpreted algebraically

By definition, when a group \( G \) is presented by \( \langle X \mid R \rangle \) there is a short exact sequence, \( \langle\langle R \rangle\rangle \to F(X) \to G \), where \( \langle\langle R \rangle\rangle \) is the normal closure of \( R \) in \( F(X) \). So a word \( w \in (X \cup X^{-1})^* \) represents \( 1 \) in \( G \) if and only if there is an equality

\[
w = \prod_{i=1}^{N} u_i^{-\epsilon_i} r_i u_i \tag{2.2}
\]

in \( F(X) \) for some integer \( N \), some \( \epsilon_i = \pm 1 \), some words \( u_i \), and some \( r_i \in R \). Recent surveys including proofs of the following key lemma are \([19]\) (Theorem 4.2.2) and \([99]\) (Theorem 2.2).

**Lemma 2.4.1** (van Kampen’s Lemma \([104]\)). *A word \( w \) is null-homotopic in a finite presentation \( \mathcal{P} \) if and only if it admits a \( \mathcal{P} \)-van Kampen diagram. Moreover, \( \text{Area}(w) \) is the minimal \( N \) such that there is an equality in \( F(X) \) of the form \( 2.2 \).*

**Exercise 2.4.2.** Adapt a proof of van Kampen’s Lemma to show that

\[
L \leq \text{IDiam}(w) \leq 2L + \max \{ \ell(r) \mid r \in R \} + \ell(w),
\]

where \( L \) is the minimal value of \( \max_i \ell(u_i) \) amongst all equalities \( 2.2 \) in \( F(X) \). (The additive term \( \ell(w) \) can be discarded when \( w \) is freely reduced.)

**Exercise 2.4.3.** Similarly, relate \( \text{EDiam}(w) \) to the minimal value of

\[
\max \left\{ d_X(1, w_0) \left| \text{prefixes } w_0 \text{ of the unreduced word } \prod_{i=1}^{N} u_i^{-\epsilon_i} r_i u_i \right. \right\}
\]

amongst all equalities \( 2.2 \) in \( F(X) \).
2.5 Filling functions interpreted computationally

Given a finite presentation $P = \langle X \mid R \rangle$ and a word $w_0 \in (X \cup X^{-1})^*$ imagine writing $w_0$ on the tape of a Turing machine, one letter in each of $\ell(w_0)$ adjacent squares. The moves described in the following lemma are ways of altering the word on the tape. Using these moves with the aim of making all squares blank is the Dehn proof system for trying to show $w$ represents the identity in $P$. (The author recalls hearing this terminology in a talk by Razborov, but is unaware of its origins.) We call $(w_i)_{i=0}^m$ a null-$P$-sequence for $w_0$ (or a $P$-sequence if we remove the requirement that the final word $w_m$ be the empty word).

Lemma 2.5.1. Suppose $w_0 \in (X \cup X^{-1})^*$. Then $w_0$ represents $1$ in $P = \langle X \mid R \rangle$ if and only if it can be reduced to the empty word $w_m$ via a sequence $(w_i)_{i=0}^m$ in which each $w_{i+1}$ is obtained from $w_i$ by one of three moves:

1. Free reduction: $w_i = ax^{-1}x \beta \mapsto a\beta = w_{i+1}$, where $x \in X^\pm$.
2. Free expansion: the inverse of a free reduction.
3. Application of a relator: $w_i = \alpha uv \beta \mapsto \alpha \beta = w_{i+1}$, where a cyclic conjugate of $uv^{-1}$ is in $R^\pm$.

Proof. If $w_0$ is null-homotopic then, by Lemma 2.4.1, it admits a van Kampen diagram $\Delta_0$. Let $(\Delta_i)_{i=0}^m$ be any shelling of $\Delta_0$ down to its base vertex $\star$ by 1-cell- and 2-cell-collapse moves. Let $w_i$ be the boundary word of $\Delta_i$ read from $\star$. Then $(w_i)_{i=0}^m$ is a null-sequence for $w_0$. The converse is straightforward.

□

Proposition 2.5.2. Suppose $w_0$ is null-homotopic in $P$. Then, amongst all null-sequences $(w_i)$ for $w_0$, $\operatorname{Area}(w_0)$ is the minimum $A$ of the number of $i$ such that $w_{i+1}$ is obtained from $w_i$ by an application-of-a-relator move, and $\operatorname{FL}(w_0)$ is the minimum of $\max_i \ell(w_i)$.

Sketch proof. The proof of Lemma 2.5.1 shows $A \leq \operatorname{Area}(w_0)$ and $F \leq \operatorname{FL}(w_0)$.

For the reverse inequalities suppose $(w_i)$ is a null-sequence for $w_0$ and that $A$ and $F$ are the number of $i$ such that $w_{i+1}$ is obtained from $w_i$ by an application-of-a-relator move, and of $\max_i \ell(w_i)$, respectively. We seek a van Kampen diagram $\Delta_0$ for $w_0$ such that $\operatorname{Area}(w_0) \leq A$ and $\operatorname{FL}(w_0) \leq F$. We will describe a sequence of planar diagrams $A_i$ that topologically are singular annuli and have the property that the outer boundary of $A_i$ is labelled by $w_0$ and the inner boundary by $w_i$.

Start with a planar, simple edge-circuit $A_0$ labelled by $w_0$. Obtain $A_{i+1}$ from $A_i$ as follows. If $w_i \mapsto w_{i+1}$ is a free reduction $w_i = ax^{-1}x \beta \mapsto a\beta = w_{i+1}$ then identify the two appropriate adjacent edges labelled $x$ and $x^{-1}$ in $A_i$. If it is a free expansion $w_i = a\beta \mapsto ax^{-1}x \beta = w_{i+1}$, then attach an edge labelled by $x$ by its initial vertex to the inner boundary circuit of $A_i$. If it is an application of a relator $w_i = \alpha uv \beta \mapsto \alpha \beta = w_{i+1}$ then attach a 2-cell with boundary labelled by $u^{-1}v$ along the $u$ portion of the inner boundary circuit of $A_i$.

However, the problem with this method of obtaining $A_{i+1}$ from $A_i$ is that there is no guarantee that $A_{i+1}$ will be planar. For example, if $w_i \mapsto w_{i+1}$ is a free expansion
immediately reversed by a free reduction \( w_{i+1} \mapsto w_{i+2} \) then an edge will be pinched off. The resolution is that moves that give rise to such problems are redundant and \((w_i)\) can be amended to avoid the difficulties, and this can be done without increasing \( A \) or \( F \). A careful treatment is in [52].

The proof outlined above actually shows more: if \((w_i)\) is a null-sequence for \( w \) then there is a van Kampen diagram \( \Delta \) for \( w \) that simultaneously has area at most the number of application-of-a-relator moves in \((w_i)\) and has \( FL(\Delta) \leq \max_i f(w_i) \).

Examples 2.5.3.

1. **Hyperbolic groups.** A group \( G \) with finite generating set \( X \) is \((Gromov-) hyperbolic\) if and only if \( \text{Cay}^1(G,X) \) is hyperbolic in the sense we will define in Section 5.2.

   This condition turns out to be independent of the finite generating set \( X \) (as follows from Theorem 5.2.2, for example). Equivalently [22, page 450], a group is hyperbolic if and only if it admits a Dehn presentation – a finite presentation \( \mathcal{P} \) such that if \( w \) is null-homotopic then one can perform an application-of-a-relator move \( w = \alpha u \beta \mapsto \alpha v \beta \) for which \( \ell(v) < \ell(u) \). It follows that \( \text{Area}(n), FL(n) \leq n \), and these bounds are simultaneously realisable.

   So the Dehn and filling length functions of hyperbolic groups grow at most linearly.

2. **Finitely generated abelian groups.** Suppose \( G \) is a finitely generated abelian group

\[
\langle x_1, \ldots, x_m \mid x_i^{\alpha_i}, [x_i, x_j], \forall 1 \leq i < j \leq m \rangle
\]

where each \( \alpha_i \in \{0, 1, 2, \ldots\} \). Any word representing the identity can be reduced to the empty word by shuffling letters past each other to collect up powers of \( x_i \), for all \( i \), and then freely reducing and applying the \( x_i^{\alpha_i} \) relators. So \( \text{Area}(n) \leq n^2 \) and \( FL(n) \leq n \) and these bounds are simultaneously realisable.

Implementing the Dehn Proof System on a Turing machine shows the following. [A Turing machine is symmetric when the transition relation is symmetric: the reverse of any transition is again a transition.]

**Theorem 2.5.4** ([97, Theorem 1.1], [11, Proposition 5.2]). For every finite presentation \( \mathcal{P} \), there is a symmetric Turing machine accepting the language of words \( w \) representing 1 in \( \mathcal{P} \) in non-deterministic time equivalent to the Dehn function of \( \mathcal{P} \) and in non-deterministic space equivalent to the filling length function of \( \mathcal{P} \).

**Sketch proof.** The construction of a non-deterministic Turing machine we give here is based on [97], however we shirk the formal descriptions of machine states and transitions (which are set out carefully in [97]). Let \( \langle X \mid R \rangle \) be a finite presentation of \( \mathcal{P} \) where for all \( r \in R \) we find \( R \) also contains all cyclic permutations of \( r \) and \( r^{-1} \). Our machine has two tapes and two heads; its alphabet is \( X^{\pm 1} \) and it will start with an input word \( w \) written on the first tape and with the second tape blank. At any time during the running of the machine there will be a word (possibly empty, but containing no blank spaces) on each
tape and the remainder of the tapes will be blank; the head on each tape will always be located between two squares, that to its right being blank and that to its left containing the final letter of the word on the tape (assuming that word is not empty). We allow a head to see only the one square to its right and the \( \rho := \max_{r \in R} \ell(r) \) squares to its left (a convenience to make the apply-a-relator transition reversible). The transitions are:

- **Shift.** Erase the final letter \( x \) from the word on the end of one tape and insert \( x^{-1} \) on the end of the word on the other tape.
- **Apply-a-relator.** Append/remove a word from \( R \) to/from the word on the first tape as a suffix.
- **Expand/Reduce.** Append/remove a letter in \( X \pm 1 \) to/from both words.
- **Accept.** Accept \( w \) if the squares seen by the two heads are empty.

Notice that acceptance occurs precisely when both tapes are empty. Consider the words one gets by concatenating the word on the first tape with the inverse of the word on the second at any given time – as the machine progresses this gives a null-sequence; indeed, (modulo the fact that it may take multiple transitions to realise each move) any null-sequence can be realised as a run of the machine. So, as the space of a run of the machine is the maximum number of non-blank squares in the course of the run, the space complexity of this machine is equivalent to the filling length function of \( \mathcal{P} \).

Checking that the time complexity of our machine is equivalent to the Dehn function of \( \mathcal{P} \) requires more care. Suppose \( w \) is a null-homotopic word. It is clear that any run of the machine on input \( w \) takes time at least \( \text{Area}(w) \) because it must involve at least \( \text{Area}(w) \) apply-a-relator transitions. For the reverse bound, we wish to show that there is a run of the machine on input \( w \) in which not only the number of apply-a-relator transitions, but also the number of shift and expand/reduce transitions, can be bounded above by a constant times \( \text{Area}(w) + \ell(w) \). Sapir [96, Section 3.2] suggests using [83, Lemma 1] to simplify the proof of [97]. We outline a similar approach based on the proof of Theorem 3.2.3.

Let \( \Delta \) be a minimal area van Kampen diagram for \( w \) and let \( T \) be a spanning tree in the 1-skeleton of \( \Delta \). Note that the total number of edges in \( T \) is at most \( \rho \text{Area}(w) + \ell(w) \). A shelling of \( \Delta \) is described in the proof of Theorem 3.2.3 in which one collapses 1-cells and 2-cells as one encounters them when following the boundary circuit of a small neighbourhood of \( T \). Realise the null-sequence of this shelling (that is the sequence of boundary words of its diagrams), as the run of our Turing machine. The total number of moves is controlled as required because the shift moves correspond to the journey around \( T \).

The message of these results is that \( \text{Area} : \mathbb{N} \rightarrow \mathbb{N} \) and \( \text{FL} : \mathbb{N} \rightarrow \mathbb{N} \) are complexity measures of the word problem for \( \mathcal{P} \) – they are the Non-Deterministic Time (up to \( \simeq \)-equivalence) and the Non-Deterministic Space of the naïve approach to solving the word problem in the Dehn proof system by haphazardly applying relators, freely reducing and freely expanding in the hope of achieving an empty tape. This theme is taken up by Birget [11] who describes it as a “connection between static fillings (like length and area) and
2.5. Filling functions interpreted computationally

dynamic fillings (e.g., space and time complexity of calculations). It opens up the tantalising possibility of, given a computational problem, translating it, suitably effectively, to the word problem in the Dehn proof system of some finite presentation. Then geometric considerations can be brought to bear on questions concerning algorithmic complexity.

Striking results in this direction have been obtained by Sapir and his collaborators using $S$-machines. See [96] for a recent survey. Their results include the following.

**Theorem 2.5.5** (Sapir-Birget-Rips [97]). If $f(n)$ is the time function of a non-deterministic Turing machine such that $f^4(n)$ is super-additive (that is, $f^4(m + n) \geq f^4(m) + f^4(n)$ for all $m, n$) then there is a finite presentation with $\text{Area}(n) \approx f^4(n)$ and $\text{IDiam}(n) \approx f^3(n)$.

Building on [10, 79, 97], a far-reaching controlled embedding result (Theorem 1.1) is proved in [12], of which the following theorem is a remarkable instance. Theorem 2.5.4 gives the if implication; the hard work is in the converse. [Recall that a problem is NP when it is decidable in non-deterministic polynomial time.]

**Theorem 2.5.6** (Birget-Ol’shanskii-Rips-Sapir [12]). The word problem of a finitely generated group $G$ is an NP-problem if and only if $G \hookrightarrow \hat{G}$ for some finitely presentable group $\hat{G}$ such that the Dehn function of $\hat{G}$ is bounded above by a polynomial. Indeed, $G \hookrightarrow \hat{G}$ can be taken to be a quasi-isometric embedding.

Birget conjectures the following analogue of Theorem 2.5.6 for filling length, relating it to non-deterministic symmetric space complexity. The if part of the conjecture is covered by Theorem 2.5.4.

**Conjecture 2.5.7** (Birget [11]). For a finitely generated group $G$ and function $f : \mathbb{N} \to \mathbb{N}$, there is a non-deterministic symmetric Turing machine accepting the language of words $w$ representing 1 in $G$ within space $\preceq f(\ell(w))$ if and only if $G$ embeds in a finitely presentable group with filling length function $\preceq f$.

Another approach is pursued by Carbone in [27, 28] where she relates van Kampen diagrams to resolution proofs in logic, in such a way that the Dehn function corresponds to proof length.

Our discussions so far in this section get us some way towards a proof of the following theorem. We encourage the reader to complete the proof as an exercise (or to refer to [48] or [99]).

**Theorem 2.5.8** (Gersten [48]). For a finite presentation the following are equivalent.

- The word problem for $\mathcal{P}$ is solvable.
- $\text{Area} : \mathbb{N} \to \mathbb{N}$ is bounded above by a recursive function.
- $\text{Area} : \mathbb{N} \to \mathbb{N}$ is a recursive function.
So, strikingly, finite presentations of groups with unsolvable word problem (which do exist [13, 74]) have Dehn functions out-growing all recursive functions. Moreover, with the results we will explain in Chapter 3, it can be shown that for a finite presentation, one of the filling functions

\[ \text{Area, FL, IDiam, GL, DGL} : \mathbb{N} \to \mathbb{N} \]

is a recursive function if and only if all are recursive functions.

However, in general, Dehn function is a poor measure of the time-complexity of the word problem for a group \( G \). Cohen [31] and Madlener & Otto [69] showed that amongst all finite presentations of groups, there is no upper bound on the size of the gap between the Dehn function and the time-complexity of the word problem (in the sense of the Grzegorczyk hierarchy).

Indeed [96], the constructions of [12, 97] can be used to produce finite presentations with word problems solvable in quadratic time but arbitrarily large (recursive) Dehn functions. The idea is that an embedding of a finitely generated group \( G \) into another group \( \hat{G} \) that has an efficient algorithm to solve its word problem leads to an efficient algorithm for the word problem in \( G \). For example, \( \langle a, b \mid a b a^{-2} = a^2 \rangle \) has an exponential Dehn function (Proposition 3.1.6), but a polynomial time word problem:

**Exercise 2.5.9.** Find a deterministic polynomial time algorithm for \( \langle a, b \mid a b a^{-2} = a^2 \rangle \).

**Hint:** \( \langle a, b \mid a^k = a^2 \rangle \) is subgroup of \( \text{GL}_2(\mathbb{Q}) \) via \( a = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \) and \( b = \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix} \).

In fact, this approach leads to an \( n(\log n)^2(\log \log n) \) time solution – see [12].

The group \( \langle a, b \mid a^k a^{-2} \rangle \), introduced by Baumslag in [4], is an even more striking example. Its Dehn function was identified by Platonov [92] as

\[
\text{Area}(n) \approx \frac{\log n}{2^n},
\]

where \( f(n) := 2^n \). Earlier Gersten [46, 48] had shown it’s Dehn function to grow faster than every iterated exponential and Bernasconi [8] had found a weaker upper bound. However I.Kapovich & Schupp [98] and Miasnikov, Ushakov & Wong [70] claim its word problem is solvable in polynomial time. In the light of this and Magnus’ result that every 1-relator group has solvable word problem [68], Schupp sets the challenge [98]:

**Open Problem 2.5.10.** Find a one-relator group for which there is, provably, no algorithm to solve the word problem within linear time.

In this context we mention Bernasconi’s result [8] that all Dehn functions of 1-relator groups are at most (that is, \( \leq \)) Ackermann’s function (which is recursive but not primitive recursive), and a question of Gersten:

**Open Problem 2.5.11.** Is the Dehn function of every one-relator group bounded above by (that is, \( \leq \)) the Dehn function of \( \langle a, b \mid a^k a^{-2} \rangle \)?
### 2.6 Filling functions for Riemannian manifolds

For a closed, connected Riemannian manifold $M$, define $\text{Area}_M : [0, \infty) \to [0, \infty)$ by

$$
\text{Area}_M(l) := \sup_c \inf_D \{ \text{Area}(D) \mid D : \mathbb{D}^2 \to \tilde{M}, D \mid_{\mathbb{D}^2} = c, \ell(c) \leq l \},
$$

where $\text{Area}(D)$ is 2-dimensional Hausdorff measure and the infimum is over all Lipschitz maps $D$. Some remarks, following Bridson [19]: by Morrey’s solution to Plateau’s problem [72], for a fixed $c$, the infimum is realised; and, due to the regularity of the situation, other standard notions of area would be equivalent here.

One can construct a $\Gamma$-equivariant map $\Phi$ of $\text{Cay}^2(\mathcal{P})$ into the universal cover $\tilde{M}$ of $M$ by first mapping $1$ to some base point $v$, then extending to $\Gamma$ (which we identify with the 0-skeleton of $\text{Cay}^2(\mathcal{P})$) by mapping $\gamma \in \Gamma$ to the translate of $v$ under the corresponding deck transformation, then extending to the 1-skeleton by joining the images of adjacent vertices by an equivariantly chosen geodesic for each edge, and then to the 2-skeleton by equivariantly chosen finite area fillings for each 2-cell. Loops in $M$ can be approximated by group-like paths – the images under $\Phi$ of edge-paths in $\text{Cay}^2(\mathcal{P})$ – and maps of discs $D : \mathbb{D}^2 \to \tilde{M}$ can be related to images under $\Phi$ of van Kampen diagrams into $\tilde{M}$ (or, more strictly, compositions with $\Phi$). This is the intuition for the following result, however a full proof requires considerable technical care.

**Theorem 2.6.1** (The Filling Theorem [19, 26, 60]). For a closed, connected Riemannian manifold $M$, the Dehn function of any finite presentation of $\Gamma := \pi_1 M$ is $\approx$-equivalent to $\text{Area}_M : [0, \infty) \to [0, \infty)$.

Similar results hold for intrinsic diameter, extrinsic diameter and filling length. The **intrinsic diameter** of a continuous map $D : \mathbb{D}^2 \to \tilde{M}$ is

$$
\text{IDiam}_M(D) := \sup \{ \rho(a, b) \mid a, b \in \mathbb{D}^2 \}
$$

where $\rho$ is the pull back of the Riemannian metric:

$$
\rho(a, b) = \inf \{ \ell(D \circ p) \mid p \text{ a path in } \mathbb{D}^2 \text{ from } a \text{ to } b \}.
$$

[Note that there may be distinct points $a, b \in \mathbb{D}^2$ for which $\rho(a, b) = 0$ – that is, $\rho$ may be only a pseudo-metric $\mathbb{D}^2$.] The extrinsic diameter $\text{EDiam}_M(D)$ is the diameter of $D(\mathbb{D}^2)$, as measured with the distance function on $\tilde{M}$. These two notions of diameter give functionals on the space of rectifiable loops in $\tilde{M}$: the **intrinsic** and **extrinsic** diameter functionals $\text{IDiam}_M, \text{EDiam}_M : [0, \infty) \to [0, \infty)$ – at $l$ these functionals take the values

$$
\sup_c \inf_D \{ \text{IDiam}_M(D) \mid D : \mathbb{D}^2 \to \tilde{M}, D \mid_{\mathbb{D}^2} = c, \ell(c) \leq l \} \quad \text{and} \quad \sup_c \inf_D \{ \text{EDiam}_M(D) \mid D : \mathbb{D}^2 \to \tilde{M}, D \mid_{\mathbb{D}^2} = c, \ell(c) \leq l \},
$$

respectively.
The filling length FL(c) of a rectifiable loop \( c : [0, 1] \to \tilde{M} \) is defined in [62] to be the infimal length \( L \) such that there is a null-homotopy \( H : [0, 1]^2 \to \tilde{M} \) where

\[
H(s, 0) = c(s) \quad \text{and} \quad H(0, t) = H(1, t) = H(s, 1)
\]

for all \( s, t \in [0, 1] \), such that for all \( t \in [0, 1] \) we find \( s \mapsto H(s, t) \) is a loop of length at most \( L \). Then the filling length function \( FL_M \) is defined to be the supremum of \( FL(c) \) over all rectifiable loops \( c : [0, 1] \to \tilde{M} \) of length at most \( c \).

The analogue of the Filling Theorem is:

**Theorem 2.6.2.** [23, 24] If \( \mathcal{P} \) is a finite presentation of the fundamental group \( \Gamma \) of a closed, connected Riemannian manifold \( M \), then

\[
\text{IDiam}_M \simeq \text{IDiam}_M, \quad \text{EDiam}_M \simeq \text{EDiam}_M, \quad \text{and} \quad FL_M \simeq FL_M.
\]

### 2.7 Quasi-isometry invariance

The Dehn function is a *presentation invariant* and, more generally, a *quasi-isometry invariant*, in the sense of the following theorem. For a proof and background see [2] or [100, Theorem 4.7].

**Theorem 2.7.1.** If \( G, H \) are groups with finite generating sets \( X, Y \), respectively, if \( (G, d_X) \) and \( (H, d_Y) \) are quasi-isometric, and if \( G \) is finitely presentable, then \( H \) is finitely presentable. Moreover, if \( G \) and \( H \) have finite presentations \( \mathcal{P} \) and \( \mathcal{Q} \), respectively, then the associated Dehn functions \( \text{Area}_P, \text{Area}_Q : \mathbb{N} \to \mathbb{N} \) satisfy \( \text{Area}_P \simeq \text{Area}_Q \). In particular, the Dehn functions of two finite presentations of the same group are \( \simeq \) equivalent.

Thus it makes sense to say that a finitely presentable group \( G \) has a linear, quadratic, \( n^\alpha \) (with \( \alpha \geq 1 \)), exponential etc. Dehn function, meaning

\[
\text{Area}(n) \approx n, n^2, n^\alpha, \exp(n), \ldots
\]

with respect to some, and hence any, finite presentation. Note also, that for \( \alpha, \beta \geq 1 \) and \( f, g : [0, \infty) \to [0, \infty) \) defined by \( f(n) = n^\alpha \) and \( g(n) = n^\beta \) we have \( f \simeq g \) if and only if \( \alpha = \beta \).

Results analogous to Theorem 2.7.1 also apply to EDiam, IDiam and FL – see [23, 24]. In each case the proof involves monitoring diagram measurements is the course of the standard proof that finite presentability is a quasi-isometry invariant – the first such quantitative version was [2]. Proofs of the independence, up to \( \simeq \) equivalence, of \( \text{IDiam} : \mathbb{N} \to \mathbb{N} \) and \( \text{FL} : \mathbb{N} \to \mathbb{N} \) on the finite presentation predating [23, 24] use Tietze transformations and are in [56] and [52]. In [55] it is shown that the gallery length functions of two *fat* finite presentations of the same group are \( \simeq \) equivalent. One expects that *fat* finite presentations of quasi-isometric groups would have \( \simeq \) equivalent gallery length functions, and DGL functions, likewise.
Table 2.1: A summary of the multiple interpretations of filling functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Interpretation</th>
<th>Formula</th>
<th>Non-deterministic</th>
<th>Riemannian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area(w)</td>
<td>$\min_{\Delta} { # \text{ of 2-cells in } \Delta }$</td>
<td>$\min_{\mathcal{P}}$</td>
<td>time</td>
<td>Area</td>
</tr>
<tr>
<td>FL(w)</td>
<td>$\min { \text{FL}(S) \mid \text{shellings } S = (\Delta_i) \text{ of } \Delta }$</td>
<td>?</td>
<td>SPACE</td>
<td>FL</td>
</tr>
<tr>
<td>IDiam(w)</td>
<td>$\min_{\Delta} \text{Diam } \Delta^{(1)}$</td>
<td>$\min_{\mathcal{P}} \max_{i} \ell(u_i)$</td>
<td>?*</td>
<td>IDiam</td>
</tr>
<tr>
<td>EDiam(w)</td>
<td>$\min_{\Delta} \pi(\Delta^{(1)})$</td>
<td>$\min_{\mathcal{P}} \max { d_X(\star, u) \mid u \text{ a prefix of } w }$</td>
<td>?</td>
<td>EDiam</td>
</tr>
<tr>
<td>GL(w)</td>
<td>$\min_{\Delta} (\Delta^*)^{(1)}$</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>DGL(w)</td>
<td>$\min { \text{Diam } T + \text{Diam } T^* \mid T \text{ a spanning tree in } \Delta^{(1)} }$</td>
<td>?</td>
<td>?**</td>
<td>?***</td>
</tr>
</tbody>
</table>

It remains a challenge to supply appropriate interpretations (if they exist) in place of each question mark (modulo * and **).

* Birget [11, Proposition 5.1] relates IDiam(w) to an exponential of the deterministic time of an approach to solving the word problem by constructing a non-deterministic finite automata based on part of the Cayley graph.

** By Theorem 3.2.6 we have DGL $\approx$ FL for fat presentations.
Chapter 3

Relationships between filling functions

This chapter concerns relationships between filling functions that apply irrespective of the group being presented. A more comprehensive account of known relationships is in [54], along with a description of how there are many additional coincidences between filling functions if one only uses van Kampen diagrams whose vertices have valence at most some constant.

We fix a finite presentation $\mathcal{P} = \langle X \mid R \rangle$ of a group $G$ for the whole chapter. Some relationships are easy to find. Suppose $\pi : \Delta \to \text{Cay}^2(\mathcal{P})$ is a van Kampen diagram with base vertex $\star$. Then

$$\text{IDiam}(\Delta) \leq \text{FL}(\Delta)$$

because the images in $\Delta$ of the boundaries $\partial \Delta_i$ of the diagrams in a shelling form a family of contracting loops that at some stage pass through any given vertex $v$ and so provide paths of length at most $\text{FL}(\Delta)/2$ to $\star$. Defining $K := \max \{ \ell(r) \mid r \in R \}$, the total number of edges in $\Delta$ is at most $K \text{Area}(\Delta) + \ell(\partial \Delta)$ and so

$$\text{FL}(\Delta) \leq K \text{Area}(\Delta) + \ell(\partial \Delta).$$

It follows that for all $n$,

$$\text{IDiam}(n) \leq \text{FL}(n) \leq K \text{Area}(n) + n.$$

Similarly, one can show that

$$\text{GL}(n) \leq \text{DGL}(n) \leq K \text{Area}(n) + n \quad \text{and} \quad \text{IDiam}(n) \leq K \text{GL}(n).$$

Next we give the “space-time” bound [52, Corollary 2], [62, 5.C] – an analogue of a result in complexity theory –
Proposition 3.0.2. Define $K := 2 |X| + 1$. Then for all $n$, 
\[
\text{Area}(n) \leq K^{\text{FL}(n)}.
\]

Proof. This result is most transparent in the Dehn proof system of Section 2.5. The number of words of length at most $\text{FL}(n)$, and hence the number of different words occurring in a null-sequence for a null-homotopic word $w$ of length at most $n$, is at most $K^{\text{FL}(n)}$. Every such $w$ has a null-sequence in which no word occurs twice and so there are at most $K^{\text{FL}(n)}$ application-of-a-relator moves in this null sequence. \qed

Proposition 3.0.2 and Theorem 2.3.4 combine to give the following result which may be surprising in that it makes no reference to the length function of the combing.

**Corollary 3.0.3.** [49, 52] The Dehn function of an (a)synchronously combable group grows at most exponentially fast.

### 3.1 The Double Exponential Theorem

Gallery length can be used to prove a theorem of D.E.Cohen [30], known as the Double Exponential Theorem. Cohen’s proof involves an analysis of the Nielsen reduction process. A proof, using Stallings folds, was given by Gersten [45] and was generalised by Papasoglu [90] to the more general setting of filling edge-loops in a simply connected complex of uniformly bounded local geometry. Birget found a proof based on context free languages [11]. The proof below is from [55]. It combines two propositions each of which establish (at most) exponential leaps, the first from $\text{IDiam}$ up to $\text{GL}$, and the second from $\text{GL}$ up to $\text{Area}$.

A key concept used is that of a geodesic spanning tree based at a vertex $\star$ in a graph $\Gamma$ – that is, a spanning tree such that for every vertex $v$ in $\Gamma$, the combinatorial distances from $v$ to $\star$ in $\Gamma$ and in $T$ agree.

**Exercise 3.1.1.** Prove that for every finite connected graph $\Gamma$ and every vertex $\star$ in $\Gamma$ there is a geodesic spanning tree in $\Gamma$ based at $\star$.

**Theorem 3.1.2.** There exists $C > 0$, depending only on $P = \langle X \mid R \rangle$, such that 
\[
\text{Area}(n) \leq n C^{\text{IDiam}(n)}
\]
for all $n \in \mathbb{N}$.

Define the based intrinsic diameter of a diagram $\Delta$ with base vertex $\star$ by 
\[
\text{IDiam}_{\star}(\Delta) := \max \left\{ \rho(\star, v) \mid v \in \Delta^{(0)} \right\},
\]
where $\rho$ denotes the combinatorial metric on $\Delta^{(1)}$. Note that, by definition, $\text{IDiam}_{\star}(\Delta) \leq \text{IDiam}(\Delta)$. 


Proposition 3.1.3. Suppose \( w \) is a null-homotopic word and \( \Delta \) is a van Kampen diagram for \( w \) such that IDiam\(_*\)(\( \Delta \)) is minimal. Moreover, assume \( \Delta \) is of minimal area amongst all such diagrams. For \( A := 2|X| + 1 \),

\[
GL(\Delta) \leq 2A^1 + 2 \text{IDiam}(\Delta).
\]

So \( GL(n) \leq 2A^1 + 2 \text{IDiam}(n) \) for all \( n \).

Proof. Let \( T \) be a geodesic spanning tree in \( \Delta^{(1)} \) based at \( \star \). (See Figure 3.1.) To every edge \( e \) in \( \Delta^{(1)} \setminus T \) associate an anticlockwise edge-circuit \( \gamma_e \) in \( \Delta^{(1)} \), based at \( \star \), by connecting the vertices of \( e \) to \( \Delta^{(1)} \) by geodesic paths in \( T \). Let \( w_e \) be the word one reads along \( \gamma_e \). Then \( \ell(w_e) \leq 1 + 2 \text{IDiam}(\Delta^{(1)}) \).

![Figure 3.1: Illustrations of the proofs of Propositions 3.1.3 and 3.1.4](image)

Let \( T^* \) be the dual tree to \( T \) (as defined in Section 2.2) and \( e^*_\infty \) be the vertex of \( T^* \) dual to the face \( e_\infty \). Suppose that \( e_1 \) and \( e_2 \) are two distinct edges of \( \Delta^{(1)} \setminus T \) dual to edges \( e_1^* \) and \( e_2^* \) of \( T^* \) on some geodesic \( \eta \) in \( T^* \) from \( e^*_\infty \) to a leaf and that \( e_2^* \) is further from \( e^*_\infty \) than \( e_1^* \) along \( \eta \). Then \( \gamma_{e_1} \) and \( \gamma_{e_2} \) bound subcomplexes \( C(\gamma_{e_1}) \) and \( C(\gamma_{e_2}) \) of \( \Delta \) in such a way that \( C(\gamma_{e_2}) \) is a subcomplex of \( C(\gamma_{e_1}) \). Suppose that \( w_{e_1} \) and \( w_{e_2} \) are the same words. Then we could cut \( C(\gamma_{e_1}) \) out of \( \Delta \) and glue \( C(\gamma_{e_2}) \) in its place, producing a new van Kampen diagram for \( w \) with strictly smaller area and no increase in IDiam\(_*\) – a contradiction.

Therefore \( \ell(\eta) \leq A^1 + 2 \text{IDiam}(\Delta) \), as the right-hand side is an upper bound on the number of distinct words of length at most \( 1 + 2 \text{IDiam}(\Delta) \), and the result follows. \( \square \)

Proposition 3.1.4. If \( \Delta \) is a van Kampen diagram for a null-homotopic word \( w \) then

\[
\text{Area}(\Delta) \leq \ell(w)(B + 1)^{GL(\Delta)},
\]

where \( B := \max \{ \ell(r) \mid r \in R \} \). It follows that \( \text{Area}(n) \leq n(B + 1)^{GL(n)} \) for all \( n \).
Proof. Sum over $1 \leq k \leq \text{GL}(\Delta)$, the geometric series whose $k$-th term $n(B + 1)^{k-1}$ dominates the number of vertices at distance $k$ from $e^*_{\infty}$ in a geodesic spanning tree in the 1-skeleton of $\Delta^*$, based at $e^*_{\infty}$, as illustrated in Figure 3.1.

Remark 3.1.5. In fact, the two propositions above establish more than claimed in Theorem 3.1.2. They show that there exists $C > 0$, depending only on $\mathcal{P}$, such that the bounds

$$\text{Area}(n) \leq n C \text{Diam}(n) \quad \text{and} \quad \text{GL}(n) \leq C \text{Diam}(n)$$

are simultaneously realisable on van Kampen diagrams that are of minimal IDiam amongst all diagrams for their boundary words. This together with an inequality of [52] show (as observed by Birget [11]) that IDiam($n$) is always no more than a single exponential of FL($n$).

Proposition 3.1.6. The presentation $\langle a, b \mid a^b = a^2 \rangle$ has filling length and Dehn functions satisfying

$$\text{FL}(n) \approx n \quad \text{and} \quad \text{Area}(n) \approx \exp(n).$$

Proof. That $\text{FL}(n) \approx n$ follows from asynchronous combability and Theorem 2.3.4, and so $\text{Area}(n) \approx \exp(n)$ by Proposition 3.0.2; $\text{FL}(n) \approx n$ by definition of $\geq$, and $\text{Area}(n) \approx \exp(n)$ by the following lemma (see e.g. [53, Section 2.3], [62, Section 4C.2]) applied to fillings for words $[a, a^3]$ as illustrated in Figure 2.7 in the case $n = 6$.

A presentation $\mathcal{P}$ is aspherical when $\text{Cay}^2(\mathcal{P})$ is contractible.

Lemma 3.1.7 (Gersten’s Asphericity Lemma [19, 47]). If $\mathcal{P}$ is a finite aspherical presentation and $\pi : \Delta \to \text{Cay}^2(\mathcal{P})$ is a van Kampen diagram such that $\pi$ is injective on the complement of $\Delta^{(1)}$ then $\Delta$ is of minimal area amongst all van Kampen diagrams with the same boundary circuit.

Exercise 3.1.8. Give another proof that $\text{Area}(n) \geq \exp(n)$ for $\langle a, b \mid a^b = a^2 \rangle$ by studying the geometry of $b$-corridors and the words in $[a^{\pm 1}]$ along their sides in van Kampen diagrams for $[a, a^3]$. (A $b$-corridor is a concatenation of 2-cells, each joined to the next along an edge labelled by $b$, and running between two oppositely oriented edges on $\partial \Delta$ both labelled by $b$. For example, every horizontal strip of 2-cells in the van Kampen diagram in Figure 2.7 is a $b$-corridor.)

Remarkably, no presentation is known for which the leap from IDiam($n$) to Area($n$) is more than the single exponential of $\langle a, b \mid a^b = a^2 \rangle$. So many people (e.g. in print [30, 48, 62], but it is often attributed to Stallings and indeed said that he conjectured a positive answer) have raised the natural question:

Open Question 3.1.9. Given a finite presentation, does there always exist $C > 0$ such that $\text{Area}(n) \leq C \text{IDiam}(n) + n$ for all $n$?
Indeed, one can ask more [55]: is a bound Area(n) ≤ C \( \text{IDiam}(n) + n \) always realisable on van Kampen diagrams of minimal IDiam for their boundary words.

Gromov [62, page 100] asks:

**Open Question 3.1.10.** Given a finite presentation, does there always exist \( C > 0 \) such that FL(n) ≤ C(\( \text{IDiam}(n) + n \)) for all \( n \)?

This question has also been attributed to Casson [62, page 101]. Gromov notes that an affirmative answer to 3.1.10 would imply an affirmative answer to 3.1.9 by Proposition 3.0.2. At the level of combinatorial or metric discs the analogue of 3.1.10 is false – Frankel & Katz [44] produced a family of metric discs \( D_n \) with diameter and perimeter 1 but such that however one null-homotopes \( \partial D_n \) across \( D_n \), one will (at some time) encounter a loop of length at least \( n \); we will describe combinatorial analogues of these discs in Section 3.2.

In [55] Gersten and the author ask 3.1.10 with FL is replaced by GL:

**Open Question 3.1.11.** Given a finite presentation, does there always exist \( C > 0 \) such that GL(n) ≤ C(\( \text{IDiam}(n) + n \)) for all \( n \)?

As mentioned in Remark 3.1.5, like filling length, gallery length sits at most an exponential below the Dehn function and at most an exponential above the intrinsic diameter function, in general, and a positive answer to 3.1.11 would resolve 3.1.9 affirmatively. However, 3.1.11 may represent a different challenge to 3.1.10 because separating the GL and IDiam functions has to involve a family of van Kampen diagrams for which there is no uniform upper bound on the valence of vertices. The features of combinatorial discs \( \Delta_n \) discussed in Section 3.2 may be reproducible in van Kampen diagrams in some group in such a way as to answer 3.1.10 negatively, but those discs have uniformly bounded vertex valences and so will not be of use for 3.1.11.

We now turn to what can be said in the direction of an affirmative answer to 3.1.9. For a null-homotopic word \( w \), define

\[
\text{IDiam}(w) := \max \{ \text{IDiam}(\Delta) \mid \text{minimal van Kampen diagrams } \Delta \text{ for } w \}.
\]

Then, as usual, define a function \( \mathbb{N} \to \mathbb{N} \), the upper intrinsic diameter function, by

\[
\text{IDiam}(n) := \max \left\{ \text{IDiam}(w) \mid \text{words } w \text{ with } \ell(w) \leq n \text{ and } w = 1 \text{ in } G \right\}.
\]

A reason 3.1.9 is a hard problem is that, in general, minimal area and minimal diameter fillings may fail to be realisable simultaneously as illustrated by the example in the following exercise.

**Exercise 3.1.12.** By finding embedded van Kampen diagrams \( \Delta_n \) for the words \( w_n := [x^n, y^n] \) and applying Lemma 3.1.7, show that the Dehn and \( \text{IDiam} \) functions of the aspherical presentation

\[
\langle x, y, s, t \mid [x, y], x^2, y^2 \rangle
\]
Chapter 3. Relationships between filling functions

of Bridson [17, 53] satisfy

\[ \text{Area}(n) \geq 2^n \quad \text{and} \quad \text{IDiam}(n) \geq 2^n. \]

But show that \( \text{IDiam}(w_n) \leq n \) for all \( n \). (Hint. Insert many copies of suitable van Kampen diagrams for \( x^n x^{-2^n} \) into \( \Delta_n \). In fact, it is proved in [53] that \( \text{IDiam}(n) \leq n \).)

Using \( \text{IDiam} \) side-steps the problem of minimal area and diameter bounds not being realised on the same van Kampen diagram and the following single exponential bound on Area can be obtained.

**Theorem 3.1.13.** [53] Given a finite presentation \( P \), there exists \( K > 0 \) such that

\[ \text{Area}(n) \leq nK^{\text{IDiam}(n)}. \]

**Sketch proof.** Suppose \( \Delta \) is a minimal area van Kampen diagram for \( w \). Recall that topologically \( \Delta \) is a singular 2-disc and so is a tree-like arrangement of arcs and topological 2-discs. Consider successive star-neighbourhoods \( \text{Star}^i(\partial \Delta) \) of \( \partial \Delta \) defined by setting \( \text{Star}^0(\partial \Delta) \) to be \( \partial \Delta \) and \( \text{Star}^{i+1}(\partial \Delta) \) to be the union closed cells that share at least a vertex with \( \text{Star}^i(\partial \Delta) \). Define \( A_i := \text{Star}^{i+1}(\partial \Delta) \setminus \text{Star}^i(\partial \Delta) \), a sequence of annuli that partition \( \Delta \) into \( (A_i)_{i=0}^m \) where \( m \leq K_0\text{IDiam}(\Delta) \) for some constant \( K_0 \) depending only on \( P \). This is illustrated in Figure 3.2. (The \( A_i \) will not be topological annuli, in general.)

![Figure 3.2: Successive star-neighbourhoods of the boundary of a diagram](image)

Suppose that for every vertex \( v \) on \( \partial \Delta \) and every topological 2-disc component \( C \) of \( \Delta \), there are at most three edges in the interior of \( C \) that are all incident with \( v \), all oriented away from or all towards \( v \), and all have the same label. Then by considering 2-cells in \( C \) that are incident with \( v \), we find there is constant \( K_1 \), depending only on \( P \), such that \( \text{Area}(A_0) \) and the total length of the inner boundary circuit of \( A_0 \) are both at most \( K_1\ell(\partial \Delta) \).

We then seek to argue similarly for \( \Delta \setminus A_0 \), to get an upper bound on the length of the inner boundary and the area of \( A_1 \). And then continuing inductively at most \( K_0\text{IDiam}(\Delta) \) times and summing an appropriate geometric series we will have our result.

The way we get this control on edges incident with vertices is to use diamond moves. These are performed on pairs of edges that are in the same 2-disc component of a van Kampen diagram, that are oriented towards or away from a common end-vertex \( v \in \partial \Delta \), that
3.2. Filling length and duality of spanning trees in planar graphs

Figure 3.3: A diamond move

have the same label, and whose other vertices, \( v_1 \) and \( v_2 \), are either both in the interior of \( \Delta \) or both on \( \partial \Delta \). The case where \( v_1 \) and \( v_2 \) are in the interior is illustrated in Figure 3.3. Performing a diamond moves when \( v_1, v_2 \in \partial \Delta \) creates a new 2-disc component.

Do such diamond moves until none remain available – the process does terminate because diamond moves either increase the number of 2-disc components (which can only happen finitely many times) or decreases the valence of some boundary vertex without altering the number of 2-disc components. Then peal off \( A_0 \) and repeat the whole process on the boundary vertices of \( \Delta \setminus A_0 \).

Continue likewise until the diagram has been exhausted. This takes at most \( \overline{IDiam(w)} \) iterations because, whilst the diamond moves may change \( IDiam \), they preserve Area. \( \Box \)

Exercise 3.1.14. Use \( \langle a, b \mid a^b = a^2 \rangle \) to show that the result of 3.1.13 is sharp, in general.

Open Problem 3.1.15. [53] Give a general upper bound for \( IDiam : \mathbb{N} \to \mathbb{N} \) in terms of \( IDiam : \mathbb{N} \to \mathbb{N} \). Theorem 3.1 implies a double exponential bound holds; one might hope for a single exponential.

3.2 Filling length and duality of spanning trees in planar graphs

For a while, the following question and variants set out in [54], concerning planar graphs and dual pairs of trees in the sense of Section 2.2, constituted an impasse in the study of relationships between filling functions. [A multigraph is a graph in which edges can form loops and pairs of vertices can be joined by the multiple edges.]

Question 3.2.1. Does there exist \( K > 0 \) such that if \( \Gamma \) is a finite connected planar graph (or multigraph) then there is a spanning tree \( T \) in \( \Gamma \) such that

\[
\begin{align*}
\text{Diam}(T) & \leq K \text{Diam}(\Gamma) \quad \text{and} \\
\text{Diam}(T^*) & \leq K \text{Diam}(\Gamma^*)
\end{align*}
\]

It is easy to establish either one of the equalities with \( K = 2 \): take a geodesic spanning tree based at some vertex (see Section 3.1). The trouble is that \( T \) and \( T^* \) determine each
other and they fight – altering one tree to reduce its diameter might increase the diameter of the other. The following exercise sets out a family of examples in which taking one of $T$ and $T^*$ to be a geodesic tree based at some vertex, does not lead to $T$ and $T^*$ having the properties of Question 3.2.1.

**Exercise 3.2.2.** Figure 3.4 shows the first four of a family of connected (multi)graphs $\Gamma_n$. The horizontal path through $\Gamma_n$ of length $2^n$ is a spanning tree $T_n$ in $\Gamma_n$ and $T_n^*$ is a geodesic spanning tree in $\Gamma_n^*$ based at the vertex dual to the face at infinity. Calculate $\text{Diam}(T_n^*)$, $\text{Diam}(\Gamma_n)$ and $\text{Diam}(\Gamma_n^*)$. Find a spanning tree $S_n$ in $\Gamma_n$ such that $\text{Diam}(S_n)$ and $\text{Diam}(S_n^*)$ are both at most a constant times $n$.

![Figure 3.4: The graphs $\Gamma_1, \ldots, \Gamma_4$ of Exercise 3.2.2.](image)

A **diagram** is a finite planar contractible 2-complex – in effect, a van Kampen diagram bereft of its group theoretic content. Question 3.2.1 can be regarded as a question about diagram measurements because a multigraph is finite, planar and connected if and only if it is the 1-skeleton of a diagram. Question 3.2.1 was recently resolved negatively in [95] as we will explain. Filling length plays a key role via the following result (compare [62, 5.C]).

**Theorem 3.2.3.** For all diagrams $\Delta$

$$\text{FL}(\Delta) \leq C(\text{DGL}(\Delta) + \ell(\partial\Delta)),$$

where $C$ is a constant depending only on the maximum length of the boundary cycles of the 2-cells in $\Delta$.

**Proof (adapted from [55]).** It suffices to show that given any spanning tree $T$ in $\Delta^{(1)}$,

$$\text{FL}(\Delta) \leq \text{Diam}(T) + 2 \lambda \text{Diam}(T^*) + \ell(\partial\Delta),$$

(3.1)

where $\lambda$ is the maximum length of the boundary circuits of 2-cells in $\Delta$.

Let the base vertex $\star$ of $\Delta$ be the root of $T$ and the vertex $\star_{\infty}$ dual to the 2-cell at infinity be the root of $T^*$. Let $m$ be the number of edges in $T$. Let $\gamma : [0, 2m] \to T$ be the edge-circuit in $T$ that starts from $\star$ and traverses every edge of $T$ twice, once in each
3.2. Filling length and duality of spanning trees in planar graphs

direction, running close to the anticlockwise boundary loop of a small neighbourhood of \(T\). For \(i = 1, 2, \ldots, 2m\) let \(\gamma_i\) be the edge traversed by \(\gamma_{[i-1,i]}\) and consider it directed from \(\gamma(i-1)\) to \(\gamma(i)\). Let \(\tau_i\) be the geodesic in \(T^*\) from \(\star_{\infty}\) to the vertex \(v_i^*\) dual to the 2-cell to the right of \(\gamma_i\) (possibly \(v_i^* = \star_{\infty}\)). Let \(T_i\) be the union of all the 2-cells dual to vertices on \(\tau_i\).

There may be some 2-cells \(C\) in \(\Delta \setminus \bigcup_i T_i\). For such a \(C\), let \(e\) be the edge of \((\partial C) \setminus T\) whose dual edge is closest in \(T^*\) to \(\star_{\infty}\). Call such an edge \(e\) *stray*. Then \(e\) must be an edge-loop because otherwise the subdiagram with boundary circuit made up of \(e\) and the geodesics in \(T\) from the end vertices of \(e\) to \(\star\) would contain some \(v_i^*\). It follows that \(e\) is the boundary of a subdiagram that is of the form of Figure 3.5 save that shelling away a 2-cell reveals at most \((\lambda - 1)\) more 2-cells instead of two. We leave it to the reader to check that such a diagram has filling length at most \((\lambda - 1)\text{Diam}(T^*)\).

Figure 3.5: Subdiagrams arising in the proof of Theorem 3.2.3.

To realise (3.1), shell the 2-cells of \(\Delta\) in the following order: for \(i = 1\), then \(i = 2\), and so on, shell all the (remaining) 2-cells along \(T_i\), working from \(\star_{\infty}\) to \(v_i^*\). Except, whenever a stray edge \(e\) appears in the boundary circuit, pause and entirely shell away the diagram it contains. In the course of all the 2-cell collapses, do a 1-cell-collapse whenever one is available.

One checks that in the course of this shelling, aside from detours into subdiagrams enclosed by stray edges, the anticlockwise boundary-circuit starting from \(\star\) follows a geodesic path in \(T\), then a path embedded in the 1-skeleton of some \(T_i\), before returning to \(\star\) along \(\partial \Delta\). These three portions of the circuit have lengths at most \(\text{Diam}(T)\), \(\lambda \text{Diam}(T^*)\) and \(\ell(\partial \Delta)\), respectively. Adding \((\lambda - 1)\text{Diam}(T^*)\) for the detour gives an estimate within the asserted bound. \(\square\)

A family \((\Delta_n)\) of diagrams, essentially combinatorial analogues of metric 2-discs from [44], is constructed as follows to resolve Question 3.2.1. One inductively defines trees \(\mathcal{T}_n\) by taking \(\mathcal{T}_1\) to be a single edge and obtaining \(\mathcal{T}_n\) from three copies of \(\mathcal{T}_{n-1}\) by identifying a leaf vertex of each. (It is not important that this does not define \(\mathcal{T}_n\) uniquely.) Fatten \(\mathcal{T}_n\) to an \(n\)-thick diagram as shown in Figure 3.6 in the case \(n = 3\). Then attach a combinatorial annulus around the fattened tree by identifying the outer boundary of the annulus (in the sense illustrated) with the boundary of the fattened tree. The annulus is made up of concentric rings, each with twice as many faces as the next, and is constructed in such a way that the length of its outer boundary circuit is the same as the length of the boundary.
Figure 3.6: Attaching the outer boundary of the annulus to the boundary of the fattened tree gives the third of a family \((\Delta_n)\) of diagrams with filling length outgrowing diameter and dual diameter.

circuit of the fattened tree; it serves as hyperbolic skirt bringing the diameters of \(\Delta^{(1)}_n\) and its dual down to \(\leq n\).

The filling length of \(\Delta_n\) grows \(\leq n^2\) – the reason being that a curve sweeping over \(T_n\) will, at some time, meet \(n + 1\) different edges of \(T_n\); more precisely:

**Lemma 3.2.4.** Suppose \(T_n\) is embedded in the unit disc \(D^2\). Let \(*\) be a basepoint on \(\partial D^2\). Suppose \(H : [0, 1]^2 \to D^2\) is a homotopy satisfying \(H(0, t) = H(1, t) = *\) for all \(t\), and \(H_0(s) = e^{2\pi i s}\) and \(H_1(s) = *\) for all \(s\), where \(H_t\) denotes the restriction of \(H\) to \([0, 1] \times \{t\}\). Further, assume \(H([0, 1] \times [0, t]) \cap H([0, 1] \times [t, 1]) = H([0, 1] \times \{t\})\) for all \(t\). Then \(H_t\) meets at least \(n + 1\) edges in \(T_n\) for some \(t \in [0, 1]\).

The final details of the proof that \((\Delta_n)\) resolves Question 3.2.1 negatively are left to the following exercise.

**Exercise 3.2.5.**

1. Show that \(\text{Diam}(\Delta_n), \text{GL}(\Delta_n) \leq n\).

2. Prove Lemma 3.2.4 (hint: induction on \(n\)) and deduce that \(\text{FL}(n) \geq n^2\).

3. Using Theorem 3.2.3, deduce a negative answer to Question 3.2.1.

The lengths of the boundary circuits in a \(P\)-van Kampen diagram are at most the length of the longest defining relator in \(P\). So Theorem 3.2.3 shows that \(\text{FL} : \mathbb{N} \to \mathbb{N}\) and \(\text{DGL} : \mathbb{N} \to \mathbb{N}\) satisfy \(\text{FL} \leq \text{DGL}\). The reverse, \(\text{DGL} \leq \text{FL}\), is also true for fat presentations (defined in Section 2.3). The proof in [55] is technical but, roughly speaking, the idea is that a shelling of \(\Lambda\) gives a family of concentric edge-circuits in \(\Lambda\) contracting down to \(*\)
3.3. Extrinsic diameter versus intrinsic diameter

which we can fatten and then have the arcs of both $T$ and $T^*$ follow these circuits. Thus we get

**Theorem 3.2.6** ([55]). *The filling functions* $GL$, $DGL$ and $FL$ for any finite fat presentation satisfy $GL \leq DGL \approx FL$.

Despite the negative answer to Question 3.2.1, whether the inequality in this theorem can be replaced by a $\approx$, rendering all three functions equivalent, remains an open question.

3.3 Extrinsic diameter versus intrinsic diameter

It is clear that for a van Kampen diagram $\pi : \Delta \to Cay^2(P)$ we have $EDiam(\Delta) \leq IDiam(\Delta)$ because paths in $\Delta^{(1)}$ are sent by $\pi$ to paths in $Cay^2(P)$. But, in general, one expects a shortest path between $\pi(a)$ and $\pi(b)$ in the 1-skeleton of $Cay^2(P)$ to take a shortcut through the space and so not lift to a path in $\Delta^{(1)}$, so $EDiam(\Delta)$ could be strictly less than $IDiam(\Delta)$. In [23] finite presentations are constructed that confirm this intuition; indeed, the examples show that (multiplicative) polynomial gaps of arbitrarily large degree occur:

**Theorem 3.3.1.** [23] Given $\alpha > 0$, there exists a finite presentation with

$$n^\alpha EDiam(n) \leq IDiam(n).$$

This result prompts the question of how far the gap can be extended:

**Open Question 3.3.2.** [23] What is the optimal upper bound for $IDiam : \mathbb{N} \to \mathbb{N}$ in terms of $EDiam : \mathbb{N} \to \mathbb{N}$, in general?

3.4 Free filling length

Recall from Section 2.2 that $FL(\Delta)$ is defined with reference to a base vertex $\star$ on $\partial \Delta$. If we remove the requirement that $\star$ be preserved throughout a shelling then we get $FFL(\Delta)$: the minimal $L$ such that there is a free combinatorial null-homotopy of $\partial \Delta$ across $\Delta$ – see Figure 3.7. We then define $FFL(w)$ for null-homotopic $w$, and the free filling length function $FFL : \mathbb{N} \to \mathbb{N}$ in the usual way.

**Exercise 3.4.1.** Re-interpret $FFL(w)$ in terms of the Dehn Proof System. (Hint. Allow one extra type of move to occur in null-sequences.)

It is questionable whether the base point plays a significant role in the definition of $FL$:

**Open Problem 3.4.2.** [24] Does there exist a finite presentation for which $FL(n) \neq FFL(n)$?

However, the following example gives null-homotopic words $w_n$ for which $FL(w_n)$ and $FFL(w_n)$ exhibit dramatically different qualitative behaviour. None-the-less the filling functions agree: $FL(n) \approx FFL(n) \approx 2^n$. (Exercise 3.4.4 provides some hints towards proof of these remarks; details are in [24].)
Chapter 3. Relationships between filling functions

Example 3.4.3. [24] Let $G$ be the group with presentation

$$
\langle a, b, t, \tau \mid b^{-1}aba^{-2}, [t, a], [\tau, at], [T, t], [\tau, T] \rangle
$$

and define $w_n := [T, a^{b^n} \tau a^{b^n}]$. Then $w_n$ is null-homotopic, has length $8n + 8$, and $\text{FFL}(w_n) = n$, but $\text{FL}(w_n) \approx 2^n$.

Exercise 3.4.4.

1. Show that $\text{FL}(w_n) \approx 2^n$. (*Hint. The van Kampen diagram $\Delta_n$ for $w_n$ shown in Figure 3.8 in the case $n = 3$, has $2^n$ concentric $t$-rings – concatenations of $2$-cells forming an annular subdiagram whose internal edges are all labelled by $t$. Show that every van Kampen diagram for $w_n$ has $t$-rings nested $2^n$ deep.*)
2. Show that $\text{FFL}(w_n) \approx n$. (Hint. Produce a diagram $\Delta'_n$ from $\Delta_n$ such that $\text{FFL}(\Delta'_n) \leq n$ by making cuts along paths labelled by powers of $a$ and gluing in suitable diagrams over the subpresentation $\langle a, b \mid b^{-1}aba^{-2} \rangle$.)

3. Find a family of words $w'_n$ with $\text{FFL}(w'_n) \approx 2^n$ and $\ell(w'_n) = n$. 

3.4. **Free filling length**
Chapter 4

Example: nilpotent groups

Recall that a group $G$ is nilpotent of class $c$ if it has lower central series

$$G = \Gamma_1 \geq \Gamma_2 \geq \cdots \geq \Gamma_{c+1} = \{1\},$$

defined inductively by $\Gamma_1 := G$ and $\Gamma_{i+1} := [G, \Gamma_i]$.

4.1 The Dehn and filling length functions

The following theorem and its proof illustrate the value of two key ideas: studying not only Area but how it interacts with other filling functions, and using the Dehn proof system (see Section 2.5) to manipulate words and monitor the Dehn and filling length function.

**Theorem 4.1.1.** If $G$ is a finitely generated nilpotent group of class $c$ then the Dehn and filling length functions of $G$ satisfy

$$\text{Area}(n) \preceq n^{c+1} \quad \text{and} \quad \text{FL}(n) \preceq n.$$ 

Moreover, these bounds are simultaneously realisable.

The combinatorial proof in [51] is the culmination of a number of results: [32, 46, 65]. A more general result was proved by Gromov [62, 5, $\Lambda^*_c$], [63] using masterful but formidable analytical techniques. The linear upper bound on filling length was proved prior to Theorem 4.1.1 by the author in [93] via asymptotic cones: finitely generated nilpotent groups have simply connected (indeed contractible) asymptotic cones by work of Pansu [85], and so the result follows as we will see in Theorem 5.3.3. Pittet [91], following Gromov, proved that a lattice in a simply connected homogeneous nilpotent Lie group of class $c$ admits a polynomial isoperimetric function of degree $c + 1$. (A nilpotent Lie group is called homogeneous if its Lie algebra is graded.)

No better general upper bound in terms of class is possible because, for example, the Dehn function of a free nilpotent group of class $c$ is $\approx n^{c+1}$ – see [99] or [5, 48].
However, it is not best possible for individual nilpotent groups: the $2k + 1$ dimensional integral Heisenberg groups are of class 2 but have Dehn functions $\approx n^2$ when $k > 1$ [1], [62, 5.A.1], [82]. By way of contrast, the 3-dimensional integral Heisenberg group $\mathcal{H}_3$ is free nilpotent of class 2 and so has a cubic Dehn function.

A full proof of Theorem 4.1.1 is beyond the scope of this text. We will illustrate the ideas in the proof in [51] in the case of $\mathcal{H}_3$, presented by

$$\mathcal{P} := \langle x, y, z \mid [x, y] z^{-1}, [x, z], [y, z] \rangle.$$

The proof is via the Dehn proof system. We will show that there is a sequence $(w_i)$ that uses $\leq n^{i+1}$ application-of-a-relator moves and has $\max_i \ell(w_i) \leq n$. This suffices -- see Section 2.5.

The following lemma gives a means of compressing powers of $z$. It is a special case of [51, Corollary 3.2], which concerns non-identity elements in the final non-trivial group $\Gamma_i$ in the lower central series. We use the terminology of Section 2.5. Fix $n \in \mathbb{N}$ (which in the proof will be the length of the word we seek to fill). For integers $s \geq 0$, define compression words $u(s)$ representing $z^s$ in $\mathcal{H}_3$; if $0 \leq s \leq n^2 - 1$ then

$$u(s) := z^s [x^n, y^n],$$

where $s = s_0 + s_1 n$ for some $s_0, s_1 \in \{0, 1, \ldots, n - 1\}$; define

$$u(n^2) := [x^n, y^n]$$

and

$$u\left( A + Bn^2 \right) := u(A) u\left( n^2 \right)^B$$

for all integers $A, B$ with $0 \leq A \leq n^2 - 1$ and $B > 0$. Note that $\ell(u(A + Bn^2)) \leq K_0 n$ where $K_0$ depends only on $B$.

**Lemma 4.1.2.** Fix $K_1 > 0$. There exists $K_2 > 0$ such that for all integers $K_1 n^2 \geq s \geq 0$, there is a concatenation of $\mathcal{P}$-sequences

$$z^s \rightarrow z^s u(0) \rightarrow z^{s-1} u(1) \rightarrow z^{s-2} u(2) \rightarrow \cdots \rightarrow z u(s - 1) \rightarrow u(s),$$

that converts $z^s$ to $u(s)$, and uses at most $K_2 n^2$ application-of-a-relator moves.

We leave the proof of the lemma as an exercise except to say that the key is absorbing $z^n$ subwords into commutators $[x^n, y^k]$ in which $k < n$, via:

$$z^n [x^n, y^k] = z^n x^{-n} y^{-k} x^n y^k = z^n x^{-n} y^{-k} yx^n y^k = z^n x^{-n} y^{-(k+1)} (xz^{-1}) y^{k+1} = [x^n, y^{k+1}].$$

A diagrammatic understanding of this calculation can be extracted from Figure 4.1 – the diagonal line running across the diagram from $\star$ is labelled by $z^{16}$ and the half of the diagram below this diagonal demonstrates the equality $z^{16} = [x^4, y^4]$. 
4.1. The Dehn and filling length functions

Figure 4.1: A van Kampen diagram for $[x^{-4}, y^{-4}][y^4, x^4]$ in the presentation $\langle x, y, z \mid [x, y]z^{-1}, [x, z], [y, z] \rangle$ for the 3-dimensional integral Heisenberg group $H_3$. The diagonal line beginning at $\star$ is labelled by $z^{16}$.

The proof of Theorem 4.1.1 is an induction on the class. The base case of class 1 concerns finitely generated abelian groups $G$ and is straightforward – see Example 2.5.3 (2). For the induction step one considers $G/\Gamma_c$, which is a finitely generated nilpotent group of class $c - 1$. Let $\overline{P} = \langle \overline{X} \mid \overline{R} \rangle$ be a finite presentation for $G/\Gamma_c$. Let $X = \overline{X} \cup X'$ where $X'$ is a finite set of length $c$ commutators that span $\Gamma_c$. The relators $\overline{R}$ represent elements of $\Gamma_c$ in $G$, so there is a finite presentation for $\langle X \mid R \rangle$ in which the elements of $R$ express the equality of elements of $\overline{R}$ with words on $X'$.

Given a length $n$ null-homotopic word $w$ in $P$, the word $\overline{w}$ obtained from $w$ by deleting all letters in $X'$ is null-homotopic in $\overline{G}$ and so admits a null-$\overline{P}$-sequence $(\overline{w}_i)_{i=0}^{m}$ which, by induction, involves words of length $\leq n$ and uses $\leq nc$ application-of-a-relator moves.

We seek to lift $\overline{w}$ to a null-$\overline{P}$-sequence for $w$ in which the words have length $\leq n$ and $\leq n^{c+1}$ application-of-a-relator moves are used. We illustrate how this works in the case of $G = H_3$. We have

$$\overline{G} = \Gamma_1/\Gamma_2 = H_3/\langle z \rangle,$$

which is presented by $\overline{P} = \langle x, y \mid [x, y] \rangle$. Start with $w$ and use commutator relations to shuffle all $z$ and $z^{-1}$ to the right and left ends of the word, respectively, leaving $\overline{w} = \overline{w}_0$ in the middle of the word. Each move $\overline{w}_i \mapsto \overline{w}_{i+1}$ that applies the relator $[x, y]$ lifts to a move
that introduces a $z^{±1}$ because $[x, y] \in \overline{R}$ lifts to $[x, y]z^{-1} \in R$. Whenever a $z$ or $z^{-1}$ appears in this way, immediately shuffle it to the right or left end of the word, respectively. As they arrive, compress the letters $z$ at the right end of the word as per Lemma 4.1.2, and compress the $z^{-1}$ at the left end by using the $P$-sequence obtained by inverting every word in the $P$-sequence of the lemma. Once $(\overline{w})_{m=0}$ has been exhausted we have a compression word at the right end of the word and its inverse at the left end, and the empty word $\overline{w}$ in between. Freely reduce to obtain the empty word.

The total number of $z^{±1}$ being sent to each end in this process is $\leq n^2$; the lengths of the words stay $\leq n$ on account of the compression process, Lemma 4.1.2, and the induction hypothesis; and, as required, the total number of application-of-a-relator moves used is $\leq n^3$, the dominant term in the estimate coming from Lemma 4.1.2. Note that the reason for collecting $z$'s and $z^{-1}$'s at different ends of the word is that the compression process of Lemma 4.1.2 has some steps that are more expensive that others, and having the power of $z$ (or $z^{-1}$) monotonically increasing avoids crossing these thresholds unduly often.

### 4.2 Open questions

By Theorem 2.3.4, an affirmative answer to the following question would be a stronger result than Theorem 4.1.1.

**Open Question 4.2.1.** Do all class $c$ finitely generated nilpotent groups admit (a)synchronous combings with length functions $L(n) \leq n^c$? (Cf. [62, 91].)

The 3-dimensional integral Heisenberg group cannot have a synchronous combing with length function growing at most linearly fast because that would contradict the $n^3$ lower bound on its Dehn function [5, 48, 99]. This makes it a candidate for a natural example of a group exhibiting the behaviour of the example of Bridson discussed in Remark 2.3.6.

**Open Question 4.2.2.** Is the 3-dimensional integral Heisenberg group synchronously combable? More generally, which nilpotent groups are synchronously combable?

In reference to the second part of 4.2.2 we mention that Holt [42] showed that the finitely generated nilpotent groups that are asynchronously combable are those that are virtually abelian. Our next question shows how little we yet know about Dehn functions of nilpotent groups.

**Open Question 4.2.3.** [6, Question N2] Does there exist a finitely generated nilpotent group whose Dehn function is not $\simeq$-equivalent to $n^\alpha$ for some integer $\alpha$?

Young [105, Section 5] describes a nilpotent group put forward by Sapir as a candidate to have Dehn function $\simeq n^2 \log n$. The upper bound has been established; the challenge of proving it to be sharp remains.

**Open Question 4.2.4.** Do linear upper bounds on the filling length functions for nilpotent groups extend to finite presentations of polycyclic groups or, more generally, linear groups?
Progress (in the affirmative direction) towards 4.2.4 would, by Proposition 3.0.2, increase the evidence for the following.

**Conjecture 4.2.5. (Gersten [54])** There is a recursive function $f : \mathbb{N} \to \mathbb{N}$ (e.g. $f(n) = 2^n$) such that finite presentations of linear groups all have $\text{Area}(n) \leq f(n)$. 
Chapter 4. Example: nilpotent groups
Chapter 5

Asymptotic cones

5.1 The definition

In [62] Gromov says of a finitely generated group $G$ with a word metric $d$:

"This space may at first appear boring and uneventful to a geometer’s eye since it is discrete and the traditional local (e.g. topological and infinitesimal) machinery does not run in $G.""

Imagine viewing $G$ from increasingly distant vantage points, i.e. scaling the metric by a sequence $s = (s_i)$ with $s_i \to \infty$. An asymptotic cone is a limit of $(G, \frac{1}{s_i}d)$, a coalescing of $G$ to a more continuous object that is amenable to attack by “topological and infinitesimal machinery”, and which “fills our geometer’s heart with joy” ([62] again).

The asymptotic cones $\text{Cone}_{\omega}(X, e, s)$ of a metric space $(X, d)$ are defined with reference to three ingredients: the first is a sequence $e = (e_i)$ of basepoints in $X$ (if $X$ is a group then, by homogeneity, all the $e_i$ may as well be the identity element); the second is a sequence of strictly positive real numbers $s = (s_i)$ such that $s_i \to \infty$ as $i \to \infty$; and the third is a non-principal ultrafilter on $\mathbb{N}$, the magic that forces convergence. Van den Dries and Wilkie [103] recognised the usefulness of non-principal ultrafilters for cutting through delicate arguments concerning extracting a convergent subsequence (with respect to the Gromov-Hausdorff distance) from a sequence of metric spaces in Gromov’s proof [59] that groups of polynomial growth are virtually nilpotent.

Definition 5.1.1. A non-principal ultrafilter is a finitely additive probability measure on $\mathbb{N}$ that takes values in $\{0, 1\}$ and gives all singleton sets measure 0.

The existence of non-principal ultrafilters is equivalent to the Axiom of Choice (see Exercises 5.1.2) and so we can be barred from describing asymptotic cones explicitly and most proofs involving asymptotic cones are non-constructive. Generally, asymptotic cones are wild beasts: the only finitely generated groups whose asymptotic cones are locally compact metric spaces are virtually nilpotent groups. [En route to his Polynomial Growth
Chapter 5. Asymptotic cones

Theorem [59], Gromov proves that the asymptotic cones of virtually nilpotent groups are proper. Drut¸u proves the converse in [38] and she adds [36] that by the Hopf-Rinow Theorem (see, for example, [22, Proposition 3.7]) proper can be replaced by locally compact as asymptotic cones of groups are complete (see [103]) geodesic spaces (see Exercise 5.1.4). Indeed, some asymptotic cones contain π1-injective copies of the Hawaiian earring and (so) have uncountable fundamental groups [25]. But rather than being put off, one should see these features as part of the spice of the subject.

The way a non-principal ultrafilter ω : P(ℕ) → {0, 1} forces convergence is by coherently (see Exercise 5.1.2 (4), below) choosing limit points from sequences of real numbers: given a sequence (a_i) in ℝ we say that a ∈ ℝ is an ω-ultralimit of (a_i) when ∀ε > 0, ω {i : |a - a_i| < ε} = 1, that ∞ is an ω-ultralimit of (a_i) when ω {i : a_i > N} = 1 for all N > 0, and that −∞ is an ω-ultralimit when ω {i : a_i < -N} = 1 for all N > 0.

Exercise 5.1.2.

1. Show that if ω is a non-principal ultrafilter on ℕ and A, B ⊆ ℕ satisfy ω(A) = ω(B) = 1 then ω(A ∩ B) = 1.

2. Establish the existence of non-principal ultrafilters ω on ℕ. (Hint. Consider the set of functions ω : P(ℕ) → {0, 1} such that ω^{-1}(1) is closed under taking intersections and taking supersets, and does not include the empty set, but does include all A ⊆ ℕ for which ℕ \ A is finite. Use Zorn’s Lemma.)

3. Let ω be a non-principal ultrafilter. Prove that every ω-ultralimit is also a limit point in the usual sense.

4. Show that every sequence (a_i) of real numbers has a unique ω-ultralimit in ℝ ∪ {±∞} denoted limω a_i. (Hint. First show that if ±∞ or −∞ is an ultralimit then it is the unique ultralimit. Next assume ±∞ is not an ultralimit and prove the existence of an ultralimit by suitably adapting a proof that every bounded sequence of real numbers has a limit point. Finally prove uniqueness.)

5. Let limω a_i denote the ω-ultralimit of (a_i). Prove that for all λ, μ ∈ ℝ and sequences (a_i), (b_i)

   \[ \lim_{\omega} (\lambda a_i + \mu b_i) = \lambda \lim_{\omega} a_i + \mu \lim_{\omega} b_i, \]

   and limω a_i < limω b_i if and only if a_i < b_i for all n in a set of ω-measure 1.

6. Suppose f : ℝ → ℝ is continuous. Show that for sequences (a_i) of real numbers,

   \[ f(\lim_{\omega} a_i) = \lim_{\omega} f(a_i). \]

Points in Coneω(X, e, s) are equivalence classes of sequences (x_i) in X such that limω d(e_i, x_i)/s_i < ∞, where two sequences (x_i), (y_i) are equivalent if and only if limω d(x_i, y_i)/s_i = 0. The metric on Coneω(X, e, s) is also denoted by d and is given by d(x, y) = limω d(x_i, y_i)/s_i where (x_i) and (y_i) are sequences representing x and y.

Exercise 5.1.3. Check d is well-defined and is a metric on Coneω(X, e, s).
The next exercise gives some important basic properties of asymptotic cones. In particular, it shows that if a metric space is \((1, \mu)\)-quasi-isometric to a geodesic metric space (as is the case for a group with a word metric, for example) then its asymptotic cones are all geodesic metric spaces. [A geodesic in a metric space \((X, d)\) is an isometric embedding \(\gamma : I \to X\), where \(I \subseteq \mathbb{R}\) is a closed interval (possibly infinite or bi-infinite). We say that \((X, d)\) is a geodesic space if every two points in \(X\) are connected by a geodesic.]

Exercise 5.1.4.

1. Show that a \((\lambda, \mu)\)-quasi-isometry \(\Phi : X \to Y\) between metric spaces induces \(\lambda\)-bi-Lipschitz homeomorphisms

\[\text{Cone}_\omega(X, (e_i), s) \to \text{Cone}_\omega(Y, (\Phi(e_i)), s),\]

for all \(\omega, (e_i)\) and \(s\).

2. Show that if \(a = (a_i)\) and \(b = (b_i)\) represent points in \(\text{Cone}_\omega(X, e, s)\) and \(\gamma_i : [0, 1] \to X\) are geodesics from \(a_i\) to \(b_i\) parametrised proportional to arc length, then \(\gamma : [0, 1] \to \text{Cone}_\omega(X, e, s)\) defined by \(\gamma(t) = (\gamma_i(t))\) is a geodesic from \(a\) to \(b\).

It is natural to ask how asymptotic cones depend on \(\omega\) and \(s\). Varying one is similar to varying the other (see [94, Appendix B]) but the precise relationship is not clear. In this chapter we will discuss results about statements about all the asymptotic cones of a group for a fixed \(\omega\) but \(s\) varying. (Other authors fix \(s\) as \((s_i) = (i)\) and vary \(\omega\), and others allow both to vary.) This will be crucial because another quirk of the subject is that the topological type of the asymptotic cones of a group, even, may depend on \(\omega\) or \(s\). Thomas & Velickovic [102] found the first example of a finitely generated group with two non-homeomorphic asymptotic cones – they used an infinite sequence of defining relations satisfying small cancellation to create holes in the Cayley graph on an infinite sequence of scales, and then, depending on whether or not its ultrafilter caused the asymptotic cone to see a similar sequence of scales, it either has non-trivial fundamental group, or is an \(\mathbb{R}\)-tree. Later Kramer, Shelah, Tent & Thomas [67] found an example of a finitely presented group with the mind-boggling property of having \(2^{2^{2^0}}\) non-homeomorphic cones if the Continuum Hypothesis (CH) is false but only one if it is true. Also they showed that under CH a finitely generated group has at most \(2^{2^{2^{2^0}}}\) non-homeomorphic asymptotic cones, a result Druţu & Sapir [39] proved to be sharp when they found a finitely generated group which (independent of CH) has \(2^{2^{2^{2^0}}}\) non-homeomorphic cones. Most recently, Sapir & Ol’shanskii [80, 76] constructed a finitely presented group for which the vanishing of the fundamental groups of their asymptotic cones depends on \(s\) (independent of CH).

### 5.2. Hyperbolic groups

A metric space \((X, d)\) is \(\delta\)-hyperbolic in the sense of Gromov [61] if for all \(w, x, y, z \in X\),

\[d(x, w) + d(y, z) \leq \max \{d(x, y) + d(z, w), d(x, z) + d(y, w)\} + \delta.\]
We say \((X,d)\) is \textit{hyperbolic} when it is \(\delta\)-hyperbolic for some \(\delta \geq 0\).

This \textit{four-point condition} for hyperbolicity is one of a number of equivalent formulations. A geodesic metric space \((X,d)\) is \textit{hyperbolic} if and only if it satisfies the \textit{thin-triangles condition}: there exists \(\delta > 0\) such that, given any geodesic triangle in \(X\), any one side is contained in a \(\delta\)-neighbourhood of the other two.

\textit{Exercise 5.2.1.} Relate the \(\delta\) that occurs in the thin-triangles condition to the \(\delta\) in the four-point condition.

We say that a group \(G\) with finite generating set \(X\) is \textit{hyperbolic} when \(\text{Cay}^1(G,X)\) is a hyperbolic metric space or, equivalently, \((G,d_X)\) satisfies the four-point condition for some \(\delta \geq 0\).

For a van Kampen diagram \(\Delta\) over a finite presentation \(\mathcal{P}\), define

\[\text{Rad}(\Delta) := \max \{ \rho(a, \partial \Delta) \mid \text{vertices } a \text{ of } \Delta \},\]

where \(\rho\) is the combinatorial metric on \(\Delta^{(1)}\). For a null-homotopic word define \(\text{Rad}(w)\) be the minimum of \(\text{Rad}(\Delta)\) amongst all van Kampen diagrams for \(w\) in the usual way, but also define \(\overline{\text{Rad}}(w)\) to be the maximum of \(\text{Rad}(\Delta)\) amongst all minimal area diagrams for \(w\). Then, as usual, define the corresponding filling functions, the \textit{radius} and \textit{upper radius} functions, \(\text{Rad}, \overline{\text{Rad}} : \mathbb{N} \to \mathbb{N}\) for \(\mathcal{P}\) to be the maxima of \(\text{Rad}(w)\) and \(\overline{\text{Rad}}(w)\), respectively, over all null-homotopic words of length at most \(n\). If \(\text{Rad}(n) \geq n\) then \(\text{IDiam}(n) \approx \text{Rad}(n)\), but in hyperbolic groups the extra sensitivity of \(\text{Rad}\) for low radius diagrams is useful, as we will see in the following theorem, which gives formulations of hyperbolicity for groups in terms of Dehn functions, radius functions, and asymptotic cones. A geodesic metric space \(X\) is an \(\mathbb{R}\)-tree when each pair of points \(a, b \in X\) is joined by a unique geodesic segment \([a,b]\), and if \([a,b]\) and \([b,c]\) are two geodesic segments with unique intersection point \(b\), then \([a,c]\) is the concatenation of \([a,b]\) and \([b,c]\).

\textbf{Theorem 5.2.2.} Let \(\omega\) be a non-principal ultra-filter on \(\mathbb{N}\). For a group \(G\) with finite generating set \(X\), the following are equivalent.

\begin{enumerate}
\item \(G\) is hyperbolic.
\item For all sequences of real numbers \(s = (s_i)\) with \(s_i \to \infty\), the cones \(\text{Cone}_\omega(G, 1, s)\) are \(\mathbb{R}\)-trees.
\item \(G\) admits a Dehn presentation (as defined in Section 2.5).
\item There exists \(C > 0\) and a finite presentation for \(G\) with Dehn function satisfying \(\text{Area}(n) \leq Cn\) for all \(n\).
\item If \(\mathcal{P}\) is a finite presentation for \(G\) then the Dehn function of \(\mathcal{P}\) satisfies \(\text{Area}(n)/n^2 \to 0\) as \(n \to \infty\).
\item There exists \(C > 0\) and a finite presentation for \(G\) with \(\overline{\text{Rad}}(n) \leq C \log n\) for all \(n\).
\item There exists \(C > 0\) and a finite presentation for \(G\) with \(\text{Rad}(n) \leq C \log n\) for all \(n\).
\end{enumerate}
5.2. Hyperbolic groups

(8). If \( \mathcal{P} \) is a finite presentation for \( G \) then the radius function of \( \mathcal{P} \) satisfies \( \text{Rad}(n)/n \to 0 \) as \( n \to \infty \).

In fact, all the asymptotic cones of non-virtually-cyclic hyperbolic groups are the same: they are the universal \( \mathbb{R} \)-trees with \( 2^{2^k} \) branching at every vertex [40]. The example [102], discussed earlier, shows that quantifying over all \( s \) is necessary in Theorem 5.2.2 (2). However, M.Kapovich & Kleiner [66] have shown that if a finitely presented group \( G \) has one asymptotic cone that is a \( \mathbb{R} \)-tree then \( G \) is hyperbolic.

Proof of Theorem 5.2.2. We begin by proving the equivalence of (1) and (2), which is due to Gromov [61]; subsequent accounts are [34], [37], and [57, Chapter 2, §1]. The proof here is based closely on that of Drutu and, in fact, amounts to not just a group theoretic result but to a characterisation of hyperbolic geodesic metric spaces.

(1) \( \implies \) (2). Assume \( G \) is hyperbolic. From Exercise 5.1.4 we know \( \text{Cone}_\omega(G, 1, s) \) is a geodesic space. One can show that any four points \( w, x, y, z \in \text{Cone}_\omega(G, 1, s) \) satisfy the four point condition with \( \delta = 0 \) by applying the four point condition to \( w_i, x_i, y_i, z_i \), where \( (w_i), (x_i), (y_i), (z_i) \) are representatives for \( w, x, y, z \), and using properties of \( \lim_w \) such as those in Exercise 5.1.2 (4). It follows that all geodesic triangles are in \( \text{Cone}_\omega(G, 1, s) \) are 0-thin and from that one can deduce that \( \text{Cone}_\omega(G, 1, s) \) is an \( \mathbb{R} \)-tree.

(2) \( \implies \) (1). Assume \( \text{Cone}_\omega(G, 1, s) \) are \( \mathbb{R} \)-trees for all \( s \). Let \( d \) denote the path metric on \( \text{Cay}^1(G, X) \). Suppose, for a contradiction, there is no \( \delta > 0 \) such that all the geodesic triangles in \( \text{Cay}^1(G, X) \) are \( \delta \)-thin. Then there are geodesic triangles \( [x_i, y_i, z_i] \) in \( \text{Cay}^1(G, X) \) such that, defining \( s_i \) to be the infimal distance such that every point on any one side of \( [x_i, y_i, z_i] \) is in an \( s_i \)-neighbourhood of the other two, we have \( s_i \to \infty \).

By compactness and by relabelling if necessary, we can assume that \( s_i = d(a_i, b_i) \) for some \( a_i \in [x_i, y_i] \) and \( b_i \in [y_i, z_i] \). Then

\[
t_i := d(a_i, [x_i, z_i]) \geq s_i
\]

and \( t_i = d(a_i, c_i) \) for some \( c_i \in [x_i, z_i] \) – see Figure 5.1. Let \( a = (a_i) \) and \( s = (s_i) \). Define \( C := \text{Cone}_\omega(\text{Cay}^1(G, X), a, s) \), which is an \( \mathbb{R} \)-tree by hypothesis.

\[\begin{array}{c}
\text{Figure 5.1: The geodesic triangle } [x_i, y_i, z_i].
\end{array}\]
Let \( y_i \) be a geodesic running from \( y_i \) to \( x_i \) via \( a_i = y_i(0) \) at constant speed \( s_i \) — that is, \( d(a_i, y_i(r)) = |r| s_i \). Define \( k_1 := \lim_{\omega} d(a_i, x_i)/s_i \) and \( k_2 := \lim_{\omega} d(a_i, y_i)/s_i \), which will be \( \infty \) if \( \omega \) or \( \gamma \) (respectively) fail to define a point in \( C \). If \( k_1, k_2 < \infty \) then let \( \gamma : [-k_2, k_1] \to C \) be the unit speed geodesic defined by

\[
\gamma(r) := (\gamma(r)) \quad \text{for} \quad -k_2 < r < k_1, \\
\gamma(k_1) := x, \quad \gamma(-k_2) := y.
\]

(For \( -k_1 < r < k_2 \), the expression \( \gamma_r \) is well defined for all \( n \) in a set of \( \omega \)-measure 1, and this is enough for \( \gamma_r \) to define a point in \( C \).) If either or both of \( k_1 \) and \( k_2 \) is \( \infty \) then let \( \gamma \) be an infinite or bi-infinite unit-speed geodesic \( \gamma(r) = (\gamma_r(r)) \). Let \( A \) denote the image of \( \gamma \). In the same way, define \( B \) to be (the image of) the geodesic through \( b = (b_i) \) and between \( y \) and \( z \).

Suppose \( c = (c_i) \) is a well defined point in \( C \) (that is, \( \lim_{\omega} t_i/s_i < \infty \)). Then, as before, define \( C \) to be the geodesic through \( c \) and between \( x \) and \( z \). We claim that every point \( p = (p_i) \) on one of \( A, B, C \) is in a \( 1 \)-neighbourhood of the other two. Then, as \( C \) is an \( \mathbb{R} \)-tree, the only way for \( A, B \) and \( C \) to remain close in this way, will be for \( A \cup B \in C \) to form a tripod (whose sides may be infinite), but that contradicts \( d(a, B \cup C) = 1 \).

Suppose \( p \) is on \( A \). In other words, \( p = \gamma(r) = (\gamma_r(r)) \) for some \( r \). We will show \( d(p, B \cup C) \leq 1 \). Similar arguments, which we omit, apply to \( p \) in other locations. Let \( q_i \) be a point on \([x_i, z_i] \cup [y_i, z_i] \) closest to \( p_i \). Then \( d(p_i, q_i) \leq s_i \) and \( d(p, q) \leq 1 \) where \( q = (q_i) \). We will show that \( q \in B \cup C \).

Let \( S = \{ i \mid q_i \in [z_i, x_i] \} \) and \( S' = \{ i \mid q_i \in [y_i, z_i] \} \). Suppose \( S \) has \( \omega \)-measure 1. Let \( \gamma' \) be the (possibly infinite or bi-infinite) unit speed geodesic in \( C \) with \( \gamma'(0) = c \), running between \( z \) and \( x \), and defined in the same way as \( \gamma \). Now \( d(p_i, q_i) \leq s_i \) and

\[
d(c_i, q_i) \leq d(c_i, a_i) + d(a_i, p_i) + d(p_i, q_i) \leq t_i + |r| s_i + s_i.
\]

So for all \( i \in S \) we find \( q_i = \gamma'(r_i) \) for some \( r_i \) with \( |r_i| \leq t_i/s_i + |r| + 1 \) and therefore \( \lim_{\omega} r_i < \infty \). It follows that \( \gamma'(r_i) = (\gamma'(|\lim_{\omega} r_i|)) \) in \( C \) and hence that \( q \in C \), as required. If \( S \) has \( \omega \)-measure 0 then \( S' \) has \( \omega \)-measure 1 and a similar argument shows \( q \in B \).

On the other hand, suppose \( c \) is not a well-defined point in \( C \) or, equivalently, \( \lim_{\omega} t_i/s_i = \infty \). Then \( A \) and \( B \) are either both infinite or both bi-infinite as \( x \) and \( z \) cannot define points in \( C \). We will show that every point on \( A \) is a distance at most 1 from \( B \). In an \( \mathbb{R} \)-tree this is not compatible with \( d(a, B) = 1 \).

Suppose \( p = (p_i) \) is a point on \( A \). Then \( p = \gamma(r) = (\gamma_r(r)) \) for some \( r \in \mathbb{R} \). As before, let \( q_i \) be a point on \([x_i, z_i] \cup [y_i, z_i] \) closest to \( p_i \). For all \( i \) such that \( q_i \in [c_i, x_i] \cup [c_i, z_i] \) we have \( d(a_i, q_i) \geq t_i \). So as \( d(p, q) = 1 \) and \( \lim_{\omega} t_i/s_i = \infty \), there is a set \( S \subseteq \mathbb{N} \) of \( \omega \)-measure 1 such that \( q_i \in [b_i, y_i] \cup [b_i, z_i] \) for all \( i \in S \). A similar argument to that used earlier shows that \( q \) is a point on \( B \), and completes the proof.

(1) \( \implies \) (3), (6), (7). There is an account of an elegant proof due to N. Brady that (1) \( \implies \) (3) in H. Short’s notes [99, Proposition 4.11]. That (1) \( \implies \) (7) is proved in Drutu [34] in the course of establishing equivalence of (1), (7) and (8). A proof of (1) \( \implies \) (7) can also be found in [99, Lemma 6.2] – the idea is to use the thin-triangles condition to find a finite
presentation for $G$ to construct suitable van Kampen diagrams. The stronger result $(1) \implies (6)$ that the logarithmic radius upper bound is realised on all minimal area van Kampen diagrams is stated in [62, §5.C] and proved in [53] using an estimate of the radius of a minimal area diagram in terms of the areas and of the annular regions and the lengths of the curves separating them in the decomposition of the proof of Theorem 3.1.13.

$(3) \implies (4), (4) \implies (5), (6) \implies (7),$ and $(7) \implies (8)$ are all immediate (modulo standard arguments about changing the finite presentation).

$(5) \implies (2).$ The following argument, combined with $(2) \implies (1),$ provides an alternative to the better known proofs of $(5) \implies (1),$ given by Bowditch [14], Ol’shanskii [77] and Papasoglu [86]. The following is a sketch of the proof of Drutu [35]. We begin with two lemmas:

**Lemma 5.2.3** (Bowditch [14]). Suppose we express an edge-circuit $\gamma$ in the Cayley 2-complex $\text{Cay}^2(\mathcal{P})$ of a finite presentation as a union of four consecutive arcs: $\gamma = a_1 \cup a_2 \cup \alpha^3 \cup a^4.$ Define $d_1 := d(a_1^4, a_2^3)$ and $d_2 := d(a_3^2, a_4^1)$ where $d$ is the combinatorial metric on the 1-skeleton of $\text{Cay}^2(\mathcal{P}).$ Then $\text{Area}(\gamma) \geq K d_1 d_2,$ where $K$ is a constant depending only on $\mathcal{P}.$

Bowditch’s lemma can be proved by taking successive star-neighbourhoods $S^i$ of $a_j$ in a $\mathcal{P}$-van Kampen diagram $\Delta$ filling $\gamma.$ That is, $S^0 = a_j$ and $S^{i+1} = \text{Star}(S^i).$ [For a subcomplex $A$ of a complex $B,$ the star neighbourhood of $A$ is the union of all the (closed) cells that intersect $A.$] For some constant $K^\prime,$ we have $S^i \cap a_3 = \emptyset$ for all $i \leq K d_1,$ and for all such $i$ the portion of $\delta S^i$ in the interior of $\Delta$ has length at least $d_2.$ The result then follows by summing estimates for the areas of $S^{i+1} \setminus S^i.$

Our second lemma is straight-forward.

**Lemma 5.2.4.** If a geodesic space is not a $\mathbb{R}$-tree then it contains a geodesic triangle whose sides only meet at their vertices.

Now assume $(5)$ and suppose $\text{Cone}_{w}(G, \mathbf{1}, \mathbf{s})$ is not an $\mathbb{R}$-tree for some $\mathbf{s}.$ So there is a geodesic triangle $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$ satisfying the condition of Lemma 5.2.4. Let $l$ be the sum of the lengths of the sides of $[\mathbf{x}, \mathbf{y}, \mathbf{z}].$ Decompose $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$ into four consecutive arcs $a_1^1, a_2^2, a_3^3, a_4^4$ such that $d(a_1^1, a_2^2), d(a_2^2, a_3^3) > \epsilon$ for some $\epsilon > 0.$

Write $\mathbf{x} = (x_i), \mathbf{y} = (y_i)$ and $\mathbf{z} = (z_i)$ where $x_i, y_i, z_i$ are vertices in $\text{Cay}^2(\mathcal{P}).$ The four points $a_j = a_j^1 \cap a_j^{i+1}$ (indices $j = 1, 2, 3, 4$ modulo 4) are represented by sequences $a_j = (a_j^l)$ of vertices. Connect up $x_i, y_i, z_i, a_j^0, a_j^1, a_j^2, a_j^3$ (in the cyclic order of the corresponding points $\mathbf{x}, \mathbf{y}, \mathbf{z}, a_1^i, a_2^i, a_3^i$ around $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$) by geodesics to make an edge-circuit $\gamma_i$ in $\text{Cay}^2(\mathcal{P}).$ Let $w_i$ be the word one reads around $\gamma_i$ and let $l_i := \ell(w_i).$ Then

$$N_1 := \{ i \mid |l - l_i/s_i| < l/2 \}$$

satisfies $o(N_1) = 1.$ In particular, $s_i > 2l_i/l$ for all $i \in N_1.$ Also

$$N_2 := \{ i \mid d(a_j^0, a_j^1) > \epsilon s_i \}$$

$$N_3 := \{ i \mid d(a_j^2, a_j^3) > \epsilon s_i \}$$
Chapter 5. Asymptotic cones

have $\omega(N_2) = \omega(N_3) = 1$. So $N := N_1 \cap N_2 \cap N_3$ has measure 1 and so is infinite, and by Lemma 5.2.3 and the inequalities given above,

$$\text{Area}(w_i) \geq (\epsilon s_i)^2 \geq \epsilon^2 \left(\frac{2L_i}{l_i} \right)^2 \geq l_i^2 = \ell(w_i)^2$$

for all $i \in N$. This contradicts (5).

(8) $\implies$ (2). This is the final step in our proof of the equivalence of the eight statements in the theorem. Assume (8). Let $s = (s_i)$ be a sequence of real numbers with $s_i \to \infty$. It is straightforward to check that every geodesic triangle $[x, y, z]$ in $\text{Cone}_\omega(G, 1, s)$ is a tripod: lift to a sequence of geodesic triangles $[x_i, y_i, z_i]$ in $\text{Cay}^1(G, X)$; each admits a van Kampen diagram with sublinear radius; deduce that each side of $[x, y, z]$ is in a $0$-neighbourhood of the other two sides.

Remarkably, weaker conditions than (5) and (8) imply hyperbolicity. If $P$ is a finite presentation for which there exist $L, N, \epsilon$ such that every null-homotopic word $w$ with $N \leq \text{Area}(w) \leq LN$ has $\text{Area}(w) \leq \epsilon \ell(w)^2$, then $P$ presents a hyperbolic group [34, 61, 78, 87]. And if there exists $L > 0$ such that $\text{Rad}(n) \leq n/73$ for all $n \leq L$ then $P$ presents a hyperbolic group [61, 89]; it is anticipated that 8 could replace 73 in this inequality and the result still hold [89].

Exercise 5.2.5. For $x, y, z, t$ in a metric space $(X, d)$, define $\delta(x, y, z; t)$ to be

$$\max\{ d(x, t) + d(t, y) - d(x, y), d(y, t) + d(t, z) - d(y, z), d(z, t) + d(t, x) - d(z, x) \}.$$ 

We say that $(X, d)$ enjoys the $L_\delta$ property of Chatterji [29] if for all $x, y, z \in X$ there exists $t \in X$ such that $\delta(x, y, z; t) \leq \delta$.

1. Which of the following spaces are $L_\delta$ for some $\delta \geq 0$?

   (a) $\mathbb{Z}^n$ with the word metric associated to the presentation $\langle x_1, \ldots, x_n \mid [x_i, x_j], \forall I \leq i < j \leq n \rangle$.

   (b) $\mathbb{R}^n$ with the Euclidean metric.

   (c) A hyperbolic group with a word metric.

2. Prove that if $(X, d)$ is an $L_\delta$ metric space for some $\delta \geq 0$ then all its asymptotic cones are $L_0$. (In contrast to Theorem 5.2.2, the converse is false for general metric spaces; however it is an open problem for spaces admitting cocompact group actions.)
5.3 Groups with simply connected asymptotic cones

We now give a characterisation of finitely generated groups with simply connected asymptotic cones. The implication \((I) \Rightarrow (2)\) was proved by R. Handel [64] and by Gromov [62, 5.F]; an account is given by Drutu [35]. The reverse implication appears in [88, page 793]. More general arguments in [94] develop those in [35, 62, 64, 88].

**Theorem 5.3.1.** Let \(G\) be a group with finite generating set \(X\). Fix any non-principal ultrafilter \(\omega\). The following are equivalent.

1. The asymptotic cones \(\text{Cone}_\omega(G, 1, s)\) are simply connected for all \(s = (s_n)\) with \(s_n \to \infty\).

2. Let \(\lambda \in (0, 1)\). There exist \(K, L \in \mathbb{N}\) such that for all null-homotopic words \(w\) there is an equality

\[
 w = \prod_{i=1}^{K} u_i w_i u_i^{-1}
\]

in the free group \(F(X)\) for some words \(u_i\) and \(w_i\) such that the \(w_i\) are null-homotopic and \(\ell(w_i) \leq \lambda \ell(w) + L\) for all \(i\).

3. Let \(\lambda \in (0, 1)\). There exist \(K, L \in \mathbb{N}\) such that for all null-homotopic words \(w\) there is a diagram around the boundary of which reads \(w\), and that possesses at most \(K\) \(2\)-cells, the boundary circuits of which are labelled by null-homotopic words of length at most \(\lambda \ell(w) + L\).

**Sketch proof.** The equivalence of (2) and (3) when the \(\lambda\)'s are each the same is proved in the same way as Lemma 2.4.1. We leave the task of extending this to the case where the \(\lambda\) differ as an exercise.

We sketch a proof by contradiction, essentially from [18], that shows that (I) implies (3). Fix \(\lambda \in (0, 1)\). Suppose there are null-homotopic words \(w_n\) with \(\ell(w_n) \to \infty\) and such that if \(\Delta_n\) is a diagram with \(\partial \Delta_n\) labelled by \(w_n\) and whose \(2\)-cells have boundaries labelled by null-homotopic words of length at most \(\lambda \ell(w)\), then \(\text{Area}(\Delta_n) \geq n\). (Note that, despite the lack of mention of \(L\) in this last statement, its negation implies (3).) Let \(\gamma_n : \partial[0, 1] \to \text{Cay}^1(G, X)\) be an edge-circuit based at the identity that follows a path labelled by \(w_n\). Define \(s_n : = \ell(w_n)\) and define \(\gamma : \partial[0, 1] \to \text{Cone}_\omega(G, 1, s)\) by \(\gamma(r) = \gamma_n(r)\). As \(\text{Cone}_\omega(G, 1, s)\) is simply connected, \(\gamma\) can be extended to a continuous map \(\overline{\gamma} : [0, 1]^2 \to \text{Cone}_\omega(G, 1, s)\). By uniform continuity there exists \(\epsilon > 0\) such that for all \(a, b \in [0, 1]^2\), if \(d(a, b) < \epsilon\) (in the Euclidean metric) then \(d(\overline{\gamma}(a), \overline{\gamma}(b)) \leq \lambda/16\).

For convenience, we can assume \(1/\epsilon\) is an integer. Consider the images \(x^{i,j}\) of the \((1 + 1/\epsilon)^2\) lattice points of the subdivision of \([0, 1]^2\) into squares of side \(1/\epsilon\). The images \(x^{i,j}, x^{i+1,j}, x^{i+1,j+1}, x^{i+1,j+1}\) of four corners of a square of side \(1/\epsilon\) are represented by sequences \((x_n^{i,j}), (x_n^{i+1,j}), (x_n^{i+1,j+1}), (x_n^{i,j+1})\) with

\[
 \lim_{s_n} d(x_n^{i,j}, x_n^{i+1,j}) + d(x_n^{i+1,j}, x_n^{i+1,j+1}) + d(x_n^{i+1,j+1}, x_n^{i,j+1}) + d(x_n^{i,j+1}, x_n^{i,j}) \leq \frac{\lambda}{4}
\]
and hence
\[ d(x_i^j, x_n^j) + d(x_n^j, x_{i+1}^j) + d(x_{i+1}^j, x_n^{j+1}) + d(x_n^{j+1}, x_i^{j+1}) < \frac{\lambda^2(w_n)}{2} \]
for all $n$ in a set $S$ of $\omega$-measure 1 (and hence infinite). Moreover $S$ can be taken to be independent of the choice of square, as there are only finitely many in the subdivision of $[0,1]^2$. All adjacent $x_n^{i,j}$ can be joined by geodesics to give fillings of infinitely many $w_n$ with diagrams with $1/e^2$-2-cells, each with boundary length at most $\lambda L(w_n)$. This is the contradiction we seek.

Finally, we assume (3) and sketch a proof of (1). Fix $\lambda \in (0,1)$ and let $K, L$ be as per (3). Let $s_n$ be a sequence of real numbers with $n \to \infty$. We will show that $C := \text{Cone}_\omega(G, 1, s)$ is simply connected. We begin by proving that all rectifiable (that is, finite length) loops $\gamma : \partial \mathbb{D}^2 \to C$ are null-homotopic.

Assume $\gamma$ is parametrised proportional to arc length. Inscribe a regular $m$-gon in $\mathbb{D}^2$ with $m > 2/\lambda$. Let $a_0^m, \ldots, a^{m-1}$ be the images under $\gamma$ of the vertices of the $m$-gon. For $i = 0, \ldots, m-1$, let $(a_n^i)_{n \geq 0}$ be sequences of vertices in $\text{Cay}^1(G, X)$ representing $a^i$. For fixed $n$ and for all $i$, join $a_n^i$ to $a_n^{i+1}$ (upper indices modulo $m$) by a geodesic and fill the resulting edge-circuit as in (3). Roughly speaking, as there are no more than $K 2$-cells in each filling, one topological configuration occurs for all $n$ in some set of $\omega$-measure 1, and the fillings converge in $C$ to a filling of the geodesic $m$-gon by a diagram of that type in which the lengths of the boundaries of the $\leq K 2$-cells are all at most $\lambda L(\gamma)$ (the additive $L$ term disappears in the limit). Adding in the regions between the $m$-gon and $\gamma$, we have a filling with at most $K + m$ regions each with boundary length at most $\lambda L(\gamma)$. Now iterate the process, refining the filling further, each successive time decreasing the length of the boundaries of the regions by a factor of $\lambda$. Asymptotic cones are complete metric spaces [103], and this fact is used in defining a limit that is a continuous extension of $\gamma$ across $\mathbb{D}^2$.

Figure 5.2: A tessellation of $\mathbb{D}^2$ by ideal triangles.

To show that arbitrary loops $\gamma : \partial \mathbb{D}^2 \to C$ are null-homotopic, first tessellate $\mathbb{D}^2$ by ideal triangles as illustrated in Figure 5.2, and then extend $\gamma$ to the 1-skeleton of the tessellation by mapping each edge to a geodesic connecting its endpoints. Each geodesic triangle is a rectifiable loop of length at most three times the diameter of the image of $\gamma$ (which is finite, by compactness). So each triangle can be filled in the way already
explained, and the result is an extension of $\gamma$ to a continuous map $\overrightarrow{\gamma} : \mathbb{D}^2 \to C$. This extension can be proved to be continuous by an argument that uses continuity of $\gamma$ to estimate the perimeter of geodesic triangles in the tessellation whose domains are close to $\partial \mathbb{D}^2$. Details are in [94]. □

Exercise 5.3.2. Recall that if $X$ and $X'$ are two finite generating sets for $G$ then $(G, d_X)$ and $(G, d'_X)$ are quasi-isometric and so their asymptotic cones are Lipschitz-equivalent. It follows that conditions (2) and (3) of Theorem 5.3.1 do not depend on the particular finite generating set $X$. Prove this result directly, i.e. without using asymptotic cones.

The polynomial bound (5.2) of the following theorem is due to Gromov [62, 5F'']. Druţu [35, Theorem 5.1] also provides a proof. The upper bound on IDiam appears as a remark of Papasoglu at the end of [88] and is proved in detail in [18]. The idea is to repeatedly refine the 2-cells of a diagram as per Theorem 5.3.1 (3) until the lengths of the boundary loops of the 2-cells are all at most a constant, but there is a technical issue of maintaining planarity during each refinement. Below, we use the Dehn proof system to avoid this issue (or, to be more honest, push it elsewhere: to the proof of Proposition 2.5.2, essentially).

Theorem 5.3.3. Fix any non-principal ultrafilter $\omega$. Suppose, for all sequences of scalars $s_n \to \infty$, the asymptotic cones $\text{Cone}_\omega(G, 1, s)$ of a finitely generated group $G$ are simply connected. Then there exists a finite presentation $\langle X \mid R \rangle$ for $G$ with respect to which

\begin{align*}
\text{Area}(n) & \leq n^\alpha, \\
\text{IDiam}(n) & \leq n, \\
\text{FL}(n) & \leq n,
\end{align*}

where $\alpha = \log_{1/\lambda} K$ and $K$ is as in Theorem 5.3.1. Further, these bounds are realisable simultaneously.

Proof. Fix $\lambda \in (0, 1)$ and let $K, L \in \mathbb{N}$ be as in Theorem 5.3.1 (3). Suppose $w$ is null-homotopic. Then, for similar reasons to those used in Section 2.5, there is a sequence $(w_i)_i$ with each $w_{i+1}$ obtained from $w_i$ by either a free reduction/expansion or by

$$w_i = \alpha \beta \mapsto \alpha u \beta = w_{i+1}$$

(5.5)

where $u$ is a null-homotopic word with $\ell(u) \leq \lambda \ell(w) + L$. Moreover, the number of moves of type (5.5) is at most $K$ and

$$\max_i \ell(w_i) \leq (K + 1)(\lambda n + L) + n.$$

This bound on $\max_i \ell(w_i)$ holds because the total number of edges in the diagram is at most $K(\lambda n + L) + n$, and a further $\lambda n + L$ is added because the moves of type (5.5) insert the whole of $u$ rather than exchanging one part of a relator for another like an application-of-a-relator move.
Let \( R \) be the set of null-homotopic words on \( X \) of length at most \( 1 + (L/(1 - \lambda)) \). We will show by induction on \( \ell(w) \) that \( w \) admits a null-\( P \)-sequence where \( P = \langle X \mid R \rangle \). This is clearly the case if \( \ell(w) \leq 1 + (L/(1 - \lambda)) \). Assume \( \ell(w) > 1 + (L/(1 - \lambda)) \). Then for \( u \) as in (5.5)

\[
\ell(u) \leq \lambda \ell(w) + L < \ell(w),
\]

and so by the induction hypothesis and Proposition 2.5.2, each \( u \) admits a null-sequence involving at most \( \text{Area}(u) \) application-of-a-relator moves and words of length at most \( FL(u) \). In place of each move (5.5) insert such a null-sequence, run backwards, to create the \( u \). The result is a null-\( P \)-sequence for \( w \). Moreover, this null-\( P \)-sequence shows that

\[
\text{Area}(n) \leq K \text{Area}(\lambda n + L) \quad \text{and} \quad \text{FL}(n) \leq \text{FL}(\lambda n + L) + (K + 1)(\lambda n + L) + n. \tag{5.6}
\]

Let \( k = 1 + \lfloor \log_\lambda(1/n) \rfloor \). Repeatedly apply (5.6) until, after \( k \) iterations, we have \( \text{Area}(n) \) in terms of

\[
\text{Area}\left(\lambda^k n + \frac{L}{1 - \lambda}\right).
\]

But this is 1 (or 0) because the argument is at most \( 1 + (L/(1 - \lambda)) \). So

\[
\text{Area}(n) \leq K^k \leq K^{1 - \log_\lambda n} = K^{\log_\lambda 1/K},
\]

which proves (5.2). We get (5.4) by summing a geometric series arising from iteratively applying (5.7), and (5.3) then follows. \( \square \)

**Example 5.3.4.** Finitely generated groups \((G,X)\) satisfying the \( L_\delta \) condition of Exercise 5.2.5 enjoy conditions (2) and (3) of Theorem 5.3.1 with \( K = 3 \), with \( L = \delta \) and \( \lambda = 2/3 \). So all the asymptotic cones of \( G \) are simply connected, \( G \) is finitely presentable, and

\[
\text{Area}(n) \leq n^{\log_{3/2} 3}, \quad \text{FL}(n) \leq n, \quad \text{and} \quad \text{IDiam}(n) \leq n.
\]

This upper bound on Dehn function is subcubic \((\log_{3/2} 3 \approx 2.71)\) and is due to Elder [41]. It is an open problem to find a group satisfying the \( L_\delta \) condition for some \( \delta \) but not having \( \text{Area}(n) \leq n^2 \).

Having Dehn function growing at most polynomially fast does not guarantee that the asymptotic cones of a group \( G \) will all be simply connected: there are groups \([18, 97]\) with such Dehn functions but with IDiam growing faster than linearly. Indeed, Ol’shanskii and Sapir have recently constructed a group with with \( \text{IDiam}(n) \approx n \) and \( \text{Area}(n) \approx n^3 \) but no asymptotic cones simply connected \([84]\):

\[
\langle a, b, c, k \mid [a, b], [a, c], k^b = ka, k^c = ka \rangle.
\]

They claim that \( S \)-machines from \([81]\) could be be used to obtain an example with \( \text{IDiam}(n) \approx n \) and \( \text{Area}(n) \approx n^2 \log n \). So the following theorem of Papasoglu \([88]\) is near sharp.
Theorem 5.3.5. If a group has quadratic Dehn function then its asymptotic cones are all simply connected.

Papasoglu’s proof shows that groups with quadratic Dehn function have linear IDiam and it then proceeds to the criterion in Theorem 5.3.1. In particular, by Theorem 5.3.3, one can deduce (without using asymptotic cones) that groups with quadratic Dehn functions have linear FL, and this seems a non-trivial result. More generally, it is shown in [52] that groups with Area($n$) $\leq n^\alpha$ for some $\alpha \geq 2$ have IDiam($n$) $\leq n^{\alpha-1}$, and in [54, Theorem 8.2] the additional conclusion GL($n$) $\leq n^{\alpha}$ is drawn.

Exercise 5.3.6. Give a direct proof that finite presentations with Area($n$) $\leq n^2$ have IDiam($n$) $\leq n$, and (harder) FL($n$) $\leq n$.

5.4 Higher dimensions

The notions of finite generability and finite presentability are $\mathcal{F}_1$ and $\mathcal{F}_2$ in a family of finiteness conditions $\mathcal{F}_n$ for a group $G$. We say $G$ is of type $\mathcal{F}_n$ if there is a $K(G,1)$ which is a complex with finite $n$-skeleton. A $K(G,1)$ is a space with fundamental group $G$ and all other homotopy groups trivial.) For example, if $G$ has finitely presentation $\mathcal{P}$ then one can attach 3-cells, then 4-cells, and so on, to the presentation 2-complex of $\mathcal{P}$ in such a way as to kill off $\pi_2$, then $\pi_1$, and so on, of its universal cover, producing a $K(G,1)$ with finite 1- and 2-skeleta.

Suppose a group $G$ is of type $\mathcal{F}_n$ for some $n \geq 2$. Then there are reasonable notions of higher dimensional filling inequalities concerning filling $(n-1)$-spheres with $n$-discs in some appropriate sense. The following approach of Bridson [20] is a geometric/combinatorial interpretation of the algebraic definitions of Alonso, Pride & Wang [3]. By way of warning we remark that there is no clear agreement between this definition and that of Epstein et al. [42], that concerns spheres in a Riemannian manifold on which $G$ acts properly, discontinuously and cocompactly.

Starting with a finite presentation for $G$ build a \textit{presentation $(n+1)$-complex} by attaching finitely many 3-cells to kill $\pi_2$ of the presentation 2-complex, and then finitely many 4-cells to kill $\pi_3$, and so on up to $(n+1)$-cells to kill $\pi_n$. The data involved in this construction is termed a \textit{finite $(n+1)$-presentation $\mathcal{P}_{n+1}$}. Let $\text{Cay}^{n+1}(\mathcal{P}_{n+1})$ denote the universal cover of the presentation $(n+1)$-complex of $\mathcal{P}_{n+1}$. An important technicality here is that the $(j+1)$-cells may be attached not by combinatorial maps from their boundary combinatorial $n$-spheres, but by \textit{singular combinatorial maps}. These can collapse $i$-cells, mapping them into the $(i-1)$-skeleton of their range – see [20] for more details.

Consider a singular combinatorial map $\gamma: S^n \to \text{Cay}^{n+1}(\mathcal{P}_{n+1})$, where $S^n$ is some combinatorial $n$-sphere. We can \textit{fill} $\gamma$ by giving a singular combinatorial extension $\tilde{\gamma}: D^{n+1} \to \text{Cay}^{n+1}(\mathcal{P}_{n+1})$ with respect to some combinatorial $(n+1)$-disk $D^{n+1}$ such that $S^n = \partial D^{n+1}$. We define $\text{Vol}_e(\gamma)$ and $\text{Vol}_{e+1}(\tilde{\gamma})$ to be the number of $n$-cells $e$ in $S^n$ and $D^{n+1}$, respectively, such that $\gamma|_e$ is a homeomorphism, and we define the \textit{filling volume} FVol$(\gamma)$ of $\gamma$ to be the minimum amongst all $\text{Vol}_{e+1}(\tilde{\gamma})$ such that $\tilde{\gamma}$ fills $\gamma$. This leads to the
Chapter 5. Asymptotic cones

definition of the \( n \)-th order Dehn function, which, up to \( \sim \) equivalence, is independent of the choice of finite \((n + 1)\)-presentation [3].

We will be concerned with inequalities concerning not just filling volume but also (intrinsic) diameter. So define \( \text{Diam}_n(\gamma) \) and \( \text{Diam}_{n+1}(\bar{\gamma}) \) to be the maxima of the distances between two vertices on \( S^n \) or \( D^{n+1} \), respectively, in the combinatorial metric on their 1-skeleta. And define \( \text{FDiam}(\gamma) \) to be the minimum of \( \text{Diam}_{n+1}(\bar{\gamma}) \) amongst all \( \bar{\gamma} \) filling \( \gamma \).

Recall from Section 2.3 how we coned off a loop (a 1-sphere) as in Figure 2.8 to get an upper bound on the Dehn function. In the same way it is possible to cone off a singular combinatorial \( n \)-sphere and then fill the rods (the cones over each of the \( \text{Vol}_n(\gamma) \) non-collapsing \( n \)-cells) to prove the following generalisation of Theorem 2.3.4.

**Theorem 5.4.1.** [42, 49] Suppose \( G \) is a finitely generated, asynchronously combable group with length function \( k \mapsto L(k) \). Then \( G \) is of type \( F_n \) for all \( n \). Further, given a finite \((n + 1)\)-presentation \( \mathcal{P}_{n+1} \) for \( G \), every singular combinatorial \( n \)-sphere \( \gamma : S^n \to \text{Cay}^{n+1}(\mathcal{P}_{n+1}) \) can be filled by some singular combinatorial \((n + 1)\)-disc \( \bar{\gamma} : D^{n+1} \to \text{Cay}^{n+1}(\mathcal{P}_{n+1}) \) with

\[
\text{FVol}_{n+1}(\gamma) \leq \text{Vol}_n(\gamma) L(\text{Diam}_n(\gamma)),
\]

\[
\text{FDiam}_{n+1}(\gamma) \leq L(\text{Diam}_n(\gamma)).
\]

Moreover, these bounds are realisable simultaneously.

Similarly, by coning off and then filling the rods it is possible to obtain filling inequalities for groups with \( n \)-connected asymptotic cones. (A space is \( n \)-connected if its homotopy groups \( \pi_0, \pi_1, \ldots, \pi_n \) are all trivial.) Each rod has filling volume that is at most polynomial in its diameter for reasons similar to those that explain the appearance of the polynomial area bound in Theorem 5.3.3.

**Theorem 5.4.2.** [94] Suppose \( G \) is a finitely generated group whose asymptotic cones are all \( n \)-connected \((n \geq 1)\). Then \( G \) is of type \( F_{n+1} \) and, given any finite \((n + 1)\)-presentation \( \mathcal{P}_{n+1} \) for \( G \), every singular combinatorial \( n \)-sphere \( \gamma : S^n \to \text{Cay}^{n+1}(\mathcal{P}_{n+1}) \) can be filled by some singular combinatorial \((n + 1)\)-disc \( \bar{\gamma} : D^{n+1} \to \text{Cay}^{n+1}(\mathcal{P}_{n+1}) \) with

\[
\text{FVol}_{n+1}(\gamma) \leq \text{Vol}_n(\gamma) (\text{Diam}_n(\gamma))^\alpha_n,
\]

\[
\text{FDiam}_{n+1}(\gamma) \leq \text{Diam}_n(\gamma),
\]

for some \( \alpha_n \) depending only on \( \mathcal{P}_{n+1} \). Moreover, these bounds are realisable simultaneously.

**Open Question 5.4.3.** Do the higher order Dehn functions of hyperbolic groups admit linear upper bounds?

Perhaps the characterisation of hyperbolic groups in terms of \( \mathbb{R} \)-trees in Theorem 5.2.2 can be used to resolve this question. Mineyev [71] gets linear upper bounds on the volumes of homological fillings.
**Open Question 5.4.4.** Is there a sequence $G_n$ of groups (discrete if possible) such that the asymptotic cones of $G_n$ are all $n$-connected but not all $(n+1)$-connected?

Gromov [62, §2.B1] makes the tantalising suggestion that $\text{SL}_n(\mathbb{Z})$ might provide a family of examples. Epstein & Thurston [42, Chapter 10] show, roughly speaking, that any $(n-2)$-st order isoperimetric function for $\text{SL}_n(\mathbb{Z})$ is at least exponential and hence, by Theorem 5.4.2, the asymptotic cones of $\text{SL}_n(\mathbb{Z})$ are not all $(n-2)$-connected. It remains to show that all the asymptotic cones of $\text{SL}_n(\mathbb{Z})$ are $(n-3)$-connected.

In [62, §2.B.(f)] Gromov outlines a strategy for showing that a certain sequence of solvable Lie groups $S_n$ will have every asymptotic cone $(n-3)$-connected but of nontrivial (uncountably generated, even) $\pi_{n-2}$. The reason one expects the non-triviality of $\pi_{n-2}$ is that similar arguments to those in [42] should show the $(n-2)$-nd order isoperimetric function of $S_n$ again to be at least exponential.

Stallings gave the first example of a finitely presented group that is not of type $F_3$ [101]. Bieri generalised this to a family of groups $\text{SB}_n$ of type $F_{n-1}$ but not $F_n$ [9]. These groups are discussed in [15]. Their asymptotic cones are ripe for investigation.
Bibliography


[98] P. Schupp. personal communication.


