

The Length of Conjugators in Solvable Groups and Lattices of Semisimple Lie Groups



Andrew W. Sale
Balliol College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2012

For my Father.

Acknowledgements

Foremost, I would like to thank my supervisor Cornelia Druțu for her dedicated support, guidance and encouragement over the last three years. A huge debt of gratitude is also owed to Tim Riley who, while at the University of Bristol, introduced me to geometric group theory and in particular the conjugacy length function. The Topology Group in Oxford has been a fantastic, dynamic place to study as a graduate student so I would like to thank everybody there, especially those with whom I shared office SGG9.

I would like to thank Pierre Pansu and Romain Tessera for their time while I visited Paris 11 and ENS Lyon. I would also like to thank Alexander Olshanskii for his helpful comments on a draft copy of [Sal12] and Jennifer Taback for a helpful discussion from which the geometric techniques for lamplighter groups were developed, replacing the algebraic methods of [Sal11].

The support of my family, above all, has been invaluable and irreplaceable.

The Length of Conjugators in Solvable Groups and Lattices of Semisimple Lie Groups

Andrew W. Sale
Balliol College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2012

Abstract

The conjugacy length function of a group Γ determines, for a given a pair of conjugate elements $u, v \in \Gamma$, an upper bound for the shortest $\gamma \in \Gamma$ such that $u\gamma = \gamma v$, relative to the lengths of u and v . This thesis focuses on estimating the conjugacy length function in certain finitely generated groups.

We first look at a collection of solvable groups. We see how the lamplighter groups have a linear conjugacy length function; we find a cubic upper bound for free solvable groups; for solvable Baumslag–Solitar groups it is linear, while for a larger family of abelian-by-cyclic groups we get either a linear or exponential upper bound; also we show that for certain polycyclic metabelian groups it is at most exponential. We also investigate how taking a wreath product effects conjugacy length, as well as other group extensions.

The Magnus embedding is an important tool in the study of free solvable groups. It embeds a free solvable group into a wreath product of a free abelian group and a free solvable group of shorter derived length. Within this thesis we show that the Magnus embedding is a quasi-isometric embedding. This result is not only used for obtaining an upper bound on the conjugacy length function of free solvable groups, but also for giving a lower bound for their L_p compression exponents.

Conjugacy length is also studied between certain types of elements in lattices of higher-rank semisimple real Lie groups. In particular we obtain linear upper bounds for the length of a conjugator from the ambient Lie group within certain families of real hyperbolic elements and unipotent elements. For the former we use the geometry of the associated symmetric space, while for the latter algebraic techniques are employed.

“You don’t have the talent,” said Jason Bourne. “You’re lacking. You can’t think geometrically.”

“What does that mean?”

“Ponder it.”

ROBERT LUDLUM – THE BOURNE SUPREMACY

“It always seems impossible until it’s done.”

NELSON MANDELA

Contents

1	Introduction	1
1.1	The Conjugacy Length Function	3
1.2	Graphs and Cayley Graphs	4
1.3	Subgroup Distortion	5
1.4	Wreath Products	5
1.5	Summary of Results	7
1.5.1	Solvable groups	7
1.5.2	Semisimple Lie Groups and Higher-rank lattices	11
2	Solvable Groups	13
2.1	Lamplighter Groups	14
2.1.1	Horocyclic products and Diestel–Leader graphs	15
2.1.2	Diestel–Leader groups	16
2.1.3	Conjugacy length for lamplighter groups	21
2.2	Wreath Products and Free Solvable Groups	26
2.2.1	Preliminaries	26
2.2.2	Conjugacy in wreath products	32
2.2.3	Subgroup distortion	49
2.2.4	Conjugacy in free solvable groups	51
2.2.5	Compression exponents	57
2.3	Group Extensions	59
2.3.1	Twisted and restricted conjugacy length functions	59
2.3.2	Conjugacy length in group extensions	62
2.3.3	Abelian-by-cyclic groups	71
2.3.4	Semidirect products $\mathbb{Z}^n \rtimes \mathbb{Z}^k$	79
2.3.5	Behaviour under finite extensions	86

3	Higher Rank Lattices	89
3.1	Symmetric Spaces	90
3.1.1	Symmetric spaces and Lie groups	90
3.1.2	Lattices	91
3.1.3	Classifications of elements: two terminologies	92
3.1.4	Structure of symmetric spaces	95
3.1.5	The asymptotic cone of a symmetric space	102
3.2	Real Hyperbolic Elements	104
3.2.1	Centralisers of real hyperbolic elements	104
3.2.2	Bounding the size of a conjugator from G	108
3.2.3	Dependence on slope	120
3.2.4	Finding a conjugator in Γ	122
3.3	Unipotent Elements	125
3.3.1	Preliminaries	126
3.3.2	Outline of the method	131
3.3.3	Relating the root system to conjugation	132
3.3.4	An ordering on the root system	135
3.3.5	Reduction of the simple case	139
3.3.6	The collateral damage of Algorithm A	142
3.3.7	The last step towards finding a short conjugator	144
A	Tables and Figures for Lemma 3.3.9	149
	Bibliography	158

List of Figures

1.1	An annular diagram	2
2.1	Identifying vertices of the tree with balls in the space of Laurent series	18
2.2	Visualising the action of $\mathbb{Z}_q \wr \mathbb{Z}$ on $DL_2(q)$	20
2.3	The climb and fall of g in the i -th tree.	21
2.4	Case 1 of Theorem 2.1.4	23
2.5	Case 2 of Theorem 2.1.4	24
2.6	A short path covering $\text{Supp}(h)$	38
2.7	Bounding D	39
2.8	The support of h_n	45
2.9	The shaded region indicates $\text{Supp}(h)$, while the dark shaded region is a triangle contained in $\text{Supp}(h)$ which contains $\frac{1}{2}n(n+1)$ elements.	49
2.10	A section of the Cayley graph of $BS(1, 2)$	67
2.11	A path from q to b^{96} in the Cayley graph of $BS(1, 3)$	68
3.1	The translation of a geodesic by a hyperbolic isometry	105
3.2	A quadrilateral demonstrating the conjugacy of hyperbolic elements	109
3.3	Using the distance to flats to bound the length of a conjugator	110
3.4	Bounding the distance to an apartment in a Euclidean building	112
3.5	The quadrilateral Q_n with side S	114
3.6	Defining $\hat{\psi}_n(t)$	115
3.7	Two asymptotic rays entering an apartment	116
3.8	Two intersecting flats	121
3.9	The intersection of a lattice with a flat when $\text{rank}_{\mathbb{R}}(G) = 2$	124
A.1	The root system E_8	155
A.2	The root system F_4	157
A.3	The root system G_2	157

Chapter 1

Introduction

In 1912, motivated by problems in low-dimensional manifolds, Max Dehn set out his celebrated list of decision problems in group theory. These are the word problem, the conjugacy problem and the isomorphism problem. These three problems are the most fundamental in combinatorial and geometric group theory and have received much attention over the last century. Dehn originally described these problems in group theory because of the significance he discovered they had in the geometry of 3-manifolds. He observed the interplay that occurs between the fundamental group of the manifold and its geometry. For example, the conjugacy problem in the fundamental group is equivalent to determining when two loops in the manifold are freely homotopic.

Let Γ be a recursively presented group with finite symmetric generating set A . The *word problem* on Γ asks whether there is an algorithm which determines when any given word on the generating set A represents the identity element of Γ . In the 1950's, Novikov and Boone showed independently that such an algorithm does not necessarily exist [Nov55], [Boo59]. Nevertheless, thanks to a century of hard work by many mathematicians we know that in a wide variety of groups such algorithms can be constructed. Examples include surface groups, shown by Dehn himself [Deh12], one-relator groups [Mag32], finitely presented residually finite groups [McK43], [Mos66] and hyperbolic groups [Gro87].

The Dehn function of a group Γ provides information on the geometric complexity of the word problem. It is a measure of the minimal area required to fill a loop in the Cayley 2-complex of Γ . Because of this geometric interpretation, determining the Dehn function of groups has been a fundamental question in geometric group theory over the last couple of decades.

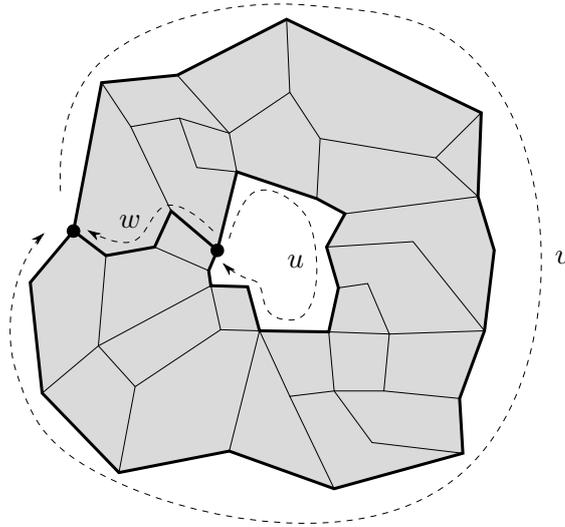


Figure 1.1: An annular diagram for $uw = vw$.

The conjugacy problem is of a similar flavour to the word problem. We say the conjugacy problem is solvable if we can write an algorithm which determines when any two given elements of the group are conjugate. By putting one of the two elements equal to the identity, we see that the word problem is a special case of the conjugacy problem. It is worth noting that Dehn solved the conjugacy problem for certain 2-manifold groups back in 1912 [Deh12], but a solution for 3-manifold groups has only very recently been found by Préaux, [Pré06] and [Pré12].

Suppose for the moment that $\Gamma = \langle \mathcal{A} \mid \mathcal{R} \rangle$ is finitely presented. If we take a word w on the (symmetric) generating set \mathcal{A} of Γ such that w represents the identity element in Γ then we can build a van Kampen diagram for w . That is, we have a finite, planar, combinatorial 2-complex \mathcal{D} with edges labelled by elements of \mathcal{A} such that around the perimeter of \mathcal{D} we read the word w and around each 2-cell we read an element of \mathcal{R} or \mathcal{R}^{-1} . The key result here is that a word w represents the identity in Γ if and only if it has a van Kampen diagram. The Dehn function measures the minimal number of 2-cells necessary to fill a loop labelled by a word representing the identity.

For the conjugacy problem there is an analogue to a van Kampen diagram: an *annular diagram*. These are discussed in Bridson and Haefliger [BH99, Ch.III.Γ Remark 2.13]. Given two words u and v on \mathcal{A} and a third word w such that $uw = vw$ we can build a van Kampen diagram for $w^{-1}uwv^{-1}$. Suppose that w is chosen so that the two boundary sections of the van Kampen diagram labelled by w and w^{-1} are

disjoint, simple paths. In this case we can identify these two segments and obtain a planar 2-complex which is homotopy equivalent to a circle (or to a point if one of u or v is trivial). Around the interior boundary we read u and around the exterior we read v . There is also an edge path from one boundary component to the other such that from the interior to the exterior we read the word w .

1.1 The Conjugacy Length Function

In geometric group theory there has often been a tendency to produce more effective results. For example, calculating the Dehn function of a group is an effective version of the word problem and it gives us a better understanding of its complexity. We can then use this extra information to determine more details of the group at hand. For instance, we can say that if a finitely presented group has a linear Dehn function then it is hyperbolic. Estimating the length of short conjugators in a group could be described as an *effective* version of the conjugacy problem, and finding a control on these lengths is the main motivation of this thesis.

Suppose we are given two words u and v which represent conjugate elements in $\Gamma = \langle \mathcal{A} \mid \mathcal{R} \rangle$. Then, by choosing a word w of minimal length which satisfies $uw = vw$, we can build an annular diagram as described above. The thickness of the annular diagram measures the length of a minimal conjugator between u and v , relative to the lengths of u and v . It is this measurement that we are interested in.

We relax the condition that Γ is finitely presented and just assume it is finitely generated. Suppose word-lengths in Γ , with respect to the given generating set \mathcal{A} , are denoted by $|\cdot|$. The *conjugacy length function* was introduced by T. Riley and is the minimal function $\text{CLF}_\Gamma : \mathbb{N} \rightarrow \mathbb{N}$ which satisfies the following: if u is conjugate to v in Γ and $|u| + |v| \leq n$ then there exists a conjugator $\gamma \in \Gamma$ such that $|\gamma| \leq \text{CLF}_\Gamma(n)$. One can define it more concretely to be the function which sends an integer n to

$$\max \left\{ \min \{ |w| : wu = vw \} : |u| + |v| \leq n \text{ and } u \text{ is conjugate to } v \text{ in } \Gamma \right\}.$$

The conjugacy length function is a geometric invariant of a group so, for example, has the potential to be used as a tool when classifying groups. It has been applied in the study of complexity of the conjugacy problem, for example in the recent work of Calvez and Wiest [CW12], and also in work on the stronger ℓ^1 -Bass conjecture by Ji, Ogle and Ramsey [JOR10].

We know various upper bounds for the conjugacy length function in certain classes of groups. For example, Gromov–hyperbolic groups have a linear upper bound; this is demonstrated by Bridson and Haefliger [BH99, Ch.III.Γ Lemma 2.9]. They also show that CAT(0) groups have an exponential upper bound for conjugacy length [BH99, Ch.III.Γ Theorem 1.12], though it is a significant open questions as to whether this upper bound is sharp. In their study of the stronger ℓ^1 –Bass conjecture, Ji, Ogle and Ramsey show that 2–step nilpotent groups have a quadratic conjugacy length function [JOR10]. In mapping class groups the conjugacy length function is linear. This was shown for all elements by J. Tao [Tao11], though Masur and Minsky [MM00] originally obtained the linear bound for pairs of conjugate pseudo-Anosov elements and Behrstock and Druţu obtained it for purely reducible elements [BD11]. The work of Crisp, Godelle and Wiest [CGW09] showing that the complexity of the conjugacy problem in right-angled Artin groups is linear also implies that these groups have a linear conjugacy length function.

1.2 Graphs and Cayley Graphs

A *graph* \mathcal{G} consists of a vertex set \mathcal{V} , an edge set \mathcal{E} and two maps $\iota, \tau : \mathcal{E} \rightarrow \mathcal{V}$ (the initial and terminal vertices). We will equip our graphs with an edge-inversion map, that is a map from \mathcal{E} to itself in which the image of $e \in \mathcal{E}$ is denoted by \bar{e} and has the property that $\iota(e) = \tau(\bar{e})$ and $\tau(e) = \iota(\bar{e})$. We will assume that all edges have a unique inverse in \mathcal{E} . The graph \mathcal{G} is said to be *simplicial* if for any pair of vertices $u, v \in \mathcal{V}$ there is at most one edge $e \in \mathcal{E}$ with the property that $\iota(e) = u$ and $\tau(e) = v$ and for every $e \in \mathcal{E}$ the two end-points $\iota(e)$ and $\tau(e)$ are distinct.

An *edge-path* is a sequence of edges $e_0, e_1, \dots, e_k \in \mathcal{E}$ such that $\iota(e_i) = \tau(e_{i-1})$ for each $i = 1, \dots, k$. An edge-path is a *loop* if $\iota(e_0) = \tau(e_k)$. A graph is said to be *connected* if for every pair of vertices $u, v \in \mathcal{V}$ there exists an edge-path e_0, e_1, \dots, e_k such that $\iota(e_0) = u$ and $\tau(e_k) = v$.

A graph \mathcal{G} is said to be a *simplicial tree* if it is a connected, simplicial graph which contains no loops.

An important example of a graph is that of a *Cayley graph* of a finitely generated group G with respect to a finite generating set X , which we denote by $\text{Cay}(G, X)$. The vertex set of $\text{Cay}(G, X)$ is identified with the elements of the group G and for each $g \in G$ and $x \in X$ there is an edge e with $\iota(e) = g$ and $\tau(e) = gx$.

1.3 Subgroup Distortion

A recurring theme when studying conjugacy length is the notion of subgroup distortion. Suppose we are looking at a finitely generated group G . We will often find ourselves with an element h whose word length we know with respect to a finite generating set for some subgroup H , but what we really want to know is its word length in G . By allowing ourselves to use elements in a finite generating set for G we may be able to reduce the number of letters needed to write a word representing h . The degree to which we can shorten these words is measured by the subgroup distortion function.

Suppose H and G have, with respect to some pair of finite generating sets, word lengths denoted by $|\cdot|_H$ and $|\cdot|_G$ respectively. The *subgroup distortion function* δ_H^G compares the size of an element in H with its size in G . It is defined as

$$\delta_H^G(n) = \max\{|h|_H \mid |h|_G \leq n\}.$$

Subgroup distortion is studied up to an equivalence relation of functions. For functions $f, g : \mathbb{N} \rightarrow [0, \infty)$ we write $f \preceq g$ if there exists an integer $C > 0$ such that $f(n) \leq Cg(Cn)$ for all $n \in \mathbb{N}$. The two functions are equivalent if both $f \preceq g$ and $g \preceq f$. In this case we write $f \asymp g$. Up to this equivalence we can talk about *the* distortion function for a group. If the distortion function of a subgroup H satisfies $\delta_H^G(n) \asymp n$ then we say H is *undistorted* in G , otherwise H is distorted.

As an example, consider the solvable Baumslag–Solitar group $BS(1, m)$ with presentation given by

$$\langle a, b \mid aba^{-1} = b^m \rangle.$$

The subgroup $\langle b \rangle$ is exponentially distorted in $BS(1, m)$. We can see the lower bound on the distortion function quite clearly because for any $k \in \mathbb{N}$ we can write the element b^{m^k} as $a^k b a^{-k}$.

1.4 Wreath Products

Let A, B be groups. Denote by $A^{(B)}$ the set of all functions from B to A with finite support and equip it with pointwise multiplication to make it a group. The (restricted) wreath product $A \wr B$ is the semidirect product $A^{(B)} \rtimes B$. To be more

precise, the elements of $A \wr B$ are pairs (f, b) where $f \in A^{(B)}$ and $b \in B$. Multiplication in $A \wr B$ is given by

$$(f, b)(g, c) = (fg^b, bc), \quad f, g \in A^{(B)}, \quad b, c \in B$$

where $g^b(x) = g(b^{-1}x)$ for each $x \in B$. The identity element in B will be denoted by e_B , while we use 1 to denote the trivial function from B to A .

The following Lemma deals with the word-length of elements in $\Gamma = A \wr B$ when A, B are finitely generated. It was given by de Cornulier (whose proof we follow here) in the Appendix of [dC06] in a slightly more general context, and also by Davis and Olshanskii [DO11, Theorem 3.4]. We fix a generating set X for B , let $S = X \cup X^{-1}$, and for each $b \in B$ denote the corresponding word-length as $|b|_B$. We consider the left-invariant word metric on B , given by $d_B(x, y) := |x^{-1}y|_B$. Similarly, fix a finite generating set T for A and let $|\cdot|_A$ denote the word-length. For $f \in A^{(B)}$, let

$$|f| = \sum_{x \in B} |f(x)|_A.$$

Let A_{e_B} be the subgroup of $A^{(B)}$ consisting of those elements whose support is contained in $\{e_B\}$. Then A_{e_B} is generated by $\{f_t \mid t \in T\}$ where $f_t(e_B) = t$ for each $t \in T$ and Γ is generated by $\{(1, s), (f_t, e_B) \mid s \in S, t \in T\}$. With respect to this generating set, we will let $|(f, b)|$ denote the corresponding word-length for $(f, b) \in \Gamma$. The notation for the following Lemma was lifted from the analogous result for more general wreath products in [dC06], whose proof we follow.

Lemma 1.4.1. *Let $(f, b) \in \Gamma = A \wr B$, where A, B are finitely generated groups. Then*

$$|(f, b)| = K(\text{Supp}(f), b) + |f|$$

where $K(\text{Supp}(f), b)$ is the length of the shortest path in the Cayley graph $\text{Cay}(B, S)$ of B from e_B to b , travelling via every point in $\text{Supp}(f)$.

Proof. Let $n = K(\text{Supp}(f), b)$, and suppose $e_B = c_0, c_1, \dots, c_n = b$ are the vertices of a path in the Cayley graph of B such that if x is in the support of f then $x = c_i$ for some $i \in \{0, 1, \dots, n\}$. For each i , let s_i be the element in S so that $c_{i+1} = c_i s_i$ and let $f_i \in A_{e_B}$ be the function such that $f_i(e_B) = f(c_i)$ unless $c_i = c_j$ for some $j < i$, in which case $f_i = 1$. Then

$$\begin{aligned} (f, b) &= (f_0 f_1^{c_1} \dots f_n^{c_n}, b) \\ &= (f_0, s_1)(f_1, s_2) \dots (f_{n-1}, s_n)(f_n, e_B) \\ &= (f_0, e_B)(1, s_1)(f_1, e_B)(1, s_2) \dots (1, s_n)(f_n, e_B). \end{aligned}$$

For each i , the element (f_i, e_B) has length equal to $|f_i(e_B)|$, hence we have the following upper bound:

$$\begin{aligned} |(f, b)| &\leq n + \sum_{i=1}^n |f_i(e_B)| \\ &= K(\text{Supp}(f), b) + |f|. \end{aligned}$$

Now suppose that $|(f, b)| = k$ and consider a geodesic word for (f, b) . After clustering together the adjacent generators of the form (f_i, e_B) we can write

$$(f, b) = (f_0, e_B)(1, s_1)(f_1, e_B)(1, s_2) \dots (f_{m-1}, e_B)(1, s_m)(f_m, e_B)$$

for some integer m . It follows from this expression that if $c_i = s_1 \dots s_i$ then $b = c_m$ and $f = f_0 f_1^{c_1} \dots f_m^{c_m}$. Since each f_i is in A_{e_B} the support of f is therefore contained in the set $\{e_B, c_1, c_2, \dots, c_m = b\}$. Thus $e_B, c_1, c_2, \dots, c_m$ describes a path in the Cayley graph of B of length m starting at e_B , passing through every point in $\text{Supp}(f)$ and finishing at b . Subsequently $m \geq K(\text{Supp}(f), b)$. Finally, note that each (f_i, e_B) has word-length equal to $|f_i(e_B)|$. Hence

$$k = m + \sum_{i=1}^m |f_i(e_B)| = m + |f| \geq K(\text{Supp}(f), b) + |f|$$

and the Lemma follows. □

1.5 Summary of Results

We will now outline the structure of this thesis, stating the main results that we obtain in each section.

1.5.1 Solvable groups

The content of this section is based on the author's results in the two papers, [Sal11] and [Sal12], plus also some new results on group extensions.

We begin in Section 2.1 by investigating the conjugacy length function of lamplighter groups $\mathbb{Z}_q \wr \mathbb{Z}$. It was shown by Bartholdi, Neuhauser and Woess [BNW08] that, when choosing an appropriate generating set, the Cayley graph of a lamplighter group is a horocyclic product of two $(q+1)$ -regular trees, an example of a Diestel–Leader graph. We use the geometry of Diestel–Leader graphs to show the following:

Theorem 2.1.4 *Let $\Gamma = \mathbb{Z}_q \wr \mathbb{Z}$. Then the conjugacy length function for Γ is linear. In particular, with respect to the generating set described above,*

$$\text{CLF}_\Gamma(n) \leq 3n.$$

Theorem 2.1.4 was initially given in [Sal11], but the proof here is more geometric and the potential for it to be applied to more general Diestel–Leader groups is more transparent.

The main motivation for Section 2.2 was to study conjugacy length in free solvable groups. When studying such groups, the Magnus embedding is a valuable tool. If N is a normal subgroup of a (non-abelian) free group F of rank r , whose derived subgroup is denoted N' , then the Magnus embedding expresses F/N' as a subgroup of the wreath product $\mathbb{Z}^r \wr F/N$. The embedding was introduced in 1939 by Wilhelm Magnus [Mag39], and in the 1950's Fox, with a series of papers [Fox53], [Fox54], [Fox56], [CFL58], [Fox60], developed a notion of calculus on free groups which enabled the Magnus embedding to be further exploited. The first result of [Sal12] and of Section 2.2 is the following:

Theorem 2.2.5. *The Magnus embedding $\varphi : F/N' \hookrightarrow \mathbb{Z}^r \wr F/N$ is 2-bi-Lipschitz for an appropriate choice of word metrics.*

The definition of a free solvable group is as follows: let $F' = [F, F]$ denote the derived subgroup of F , where F is the free group of rank r . Denote by $F^{(d)}$ the d -th derived subgroup, that is $F^{(d)} = [F^{(d-1)}, F^{(d-1)}]$. The *free solvable group* of rank r and derived length d is the quotient $S_{r,d} = F/F^{(d)}$. The conjugacy problem in free solvable groups was shown, using the Magnus embedding, to be solvable by Kargapolov and Remeslennikov [KR66] (see also [RS70]) extending the same result for free metabelian groups by Matthews [Mat66]. Recently, Vassileva [Vas11] has looked at the computational complexity of algorithms to solve the conjugacy problem and the conjugacy search problem in wreath products and free solvable groups. In particular Vassileva showed that the complexity of the conjugacy search problem in free solvable groups is at most polynomial. Using Theorem 2.2.5 we are able to improve our understanding of the length of short conjugators in free solvable groups:

Theorem 2.2.24. *Let $r, d > 1$. Then the conjugacy length function of the free solvable group $S_{r,d}$ is bounded above by a cubic polynomial.*

In order to use the Magnus embedding we must understand conjugacy in wreath products. For such groups the conjugacy problem was studied by Matthews [Mat66],

who showed that for two recursively presented groups A, B with solvable conjugacy problem, their wreath product $A \wr B$ has solvable conjugacy problem if and only if B has solvable power problem. In Section 2.2.2 we show that for two finitely generated groups A and B there is an upper bound for the conjugacy length function of $A \wr B$ which depends on the conjugacy length functions of A and B and on the subgroup distortion of infinite cyclic subgroups of B . In the case when the B -component of the conjugate elements are of infinite order, the conjugacy length does not depend on the conjugacy length in A . Furthermore, when we avoid a certain collection of conjugacy classes, the conjugacy length function of B will also not appear.

Theorems 2.2.10 & 2.2.12. *Suppose A and B are finitely generated groups. Let $u = (f, b), v = (g, c)$ be elements in $\Gamma = A \wr B$. Then u, v are conjugate if and only if there exists a conjugator $\gamma \in \Gamma$ such that*

$$\begin{aligned} d_\Gamma(1, \gamma) &\leq (n+1)P(2\delta_{\langle b \rangle}^B(P) + 1) && \text{if } b \text{ is of infinite order, or} \\ d_\Gamma(1, \gamma) &\leq P(N+1)(2n + \text{CLF}_A(n) + 1) && \text{if } b \text{ is of finite order } N, \end{aligned}$$

where $n = d_\Gamma(1, u) + d_\Gamma(1, v)$, δ_H^B is the subgroup distortion function of $H < B$ and $P = 2n$ if (f, b) is not conjugate to $(1, b)$ and $P = n + \text{CLF}_B(n)$ otherwise.

In the second half of Section 2.2.2 we look for lower bounds of $\text{CLF}_{A \wr B}$. We develop a technique for using subgroup distortion in B for this purpose, though it is not the only tool we use. In particular, when considering wreath products of the form $A \wr B$ when B contains a copy of \mathbb{Z}^2 we make use of the quadratic Dehn function of \mathbb{Z}^2 . This is used to give a quadratic lower bound on the conjugacy length function of these wreath products.

Theorems 2.2.14, 2.2.17 & Proposition 2.2.16 *Let A and B be finitely generated groups and let $\text{BS}(1, q)$ denote a solvable Baumslag–Solitar group. Then*

- for any $x \in B$ of infinite order, $\text{CLF}_{A \wr B}(n) \succeq \delta_{\langle x \rangle}^B(n)$;
- if B contains a copy of \mathbb{Z}^2 , $\text{CLF}_{A \wr B}(n) \succeq n^2$;
- $\text{CLF}_{A \wr \text{BS}(1, q)}(n) \asymp \exp(n)$.

The final result of Section 2.2 concerns the L_p compression exponent of free solvable groups. Using Theorem 2.2.5 and results of Li [Li10] and Naor and Peres [NP11] we can show that free solvable groups have non-zero L_p compression exponent:

Corollary 2.2.27. *For $r, d \in \mathbb{N}$, $r, d \neq 1$, the L_p compression exponent for $S_{r,d}$ satisfies*

$$\frac{1}{d-1} \max \left\{ \frac{1}{p}, \frac{1}{2} \right\} \leq \alpha_p^*(S_{r,d}).$$

Moving on to Section 2.3, we consider group extensions with a short exact sequence

$$1 \longrightarrow F \xrightarrow{\alpha} G \xrightarrow{\beta} H \longrightarrow 1$$

Bogopolski, Martino and Ventura [BMV10] investigate in what situations the solubility of the conjugacy problem for F can imply its solubility in G . Their work only applies to group extensions in which the quotient group H satisfies certain conditions on the structure of centralisers of elements in H . In particular, the centraliser in H of any $h \in H$ must be virtually cyclic. We can relax this condition slightly, replacing it with a geometric condition, asking for a function $\rho : G \rightarrow [0, \infty)$ which measures the diameter of the fundamental domain of $\beta(Z_G(u))$ inside $Z_H(\beta(u))$ for any $u \in G$, where $Z_G(u)$ denotes the centraliser in G of u . If we let

$$\rho_n = \max\{\rho(u) \mid u \in G, |u| \leq n\}$$

then we show the following:

Theorem 2.3.1 *The conjugacy length function of G satisfies*

$$\text{CLF}_G(n) \leq \text{CLF}_H(n) + \max \left\{ \text{RCL}_F^G(n), \rho_n + \mathcal{TCL}_F(2\delta_F^G(n + \rho_n); A_G^{(n)}) \right\}$$

where \mathcal{TCL}_F is the twisted conjugacy length function of F and RCL_F^G is the restricted conjugacy length function of F in G , which are defined in Section 2.3.1.

We use this Theorem to study conjugacy length in certain abelian-by-cyclic groups and abelian-by-abelian groups. Following the work of Bieri and Strebel [BS78], the finitely presented, torsion-free, abelian-by-cyclic groups are given by presentations of the form

$$\Gamma_M = \langle t, a_1, \dots, a_d \mid [a_i, a_j] = 1, ta_it^{-1} = \varphi_M(a_i); i, j = 1, \dots, d \rangle$$

where $M = (m_{ij})$ is a $d \times d$ matrix with integer entries and non-zero determinant and $\varphi_M(a_i) = a_1^{m_{1i}} \dots a_d^{m_{di}}$ for each $i = 1, \dots, d$.

Theorem 2.3.6 *Suppose M is a diagonalisable matrix, all of whose eigenvalues have absolute value greater than 1. Then there exists a constant C depending on M such that*

$$\text{CLF}_{\Gamma_M}(n) \leq C\lambda^{28n}$$

where λ is the largest absolute value of an eigenvalue of M .

The next result was originally shown in [Sal11], though the proof has been slightly modified here to fit with the framework of Theorem 2.3.1.

Theorem 2.3.15 *Let $\Gamma = \mathbb{Z}^d \rtimes_{\varphi} \mathbb{Z}^k$, where the image of $\varphi : \mathbb{Z}^k \hookrightarrow \mathrm{SL}_d(\mathbb{Z})$ is contained in an \mathbb{R} -split torus T . Then there exist positive constants A, B such that*

(1) *if $k = 1$ then $\mathrm{CLF}_{\Gamma}(n) \leq Bn$;*

(2) *if $k > 1$ then $\mathrm{CLF}_{\Gamma}(n) \leq A^n$.*

1.5.2 Semisimple Lie Groups and Higher-rank lattices

In the second half of this thesis we look at conjugacy length between elements in an irreducible lattice Γ of a higher-rank semisimple real Lie group G . We focus our attention on two types of elements. Firstly we look at *real hyperbolic elements*. To study these we use the geometry of the symmetric space X associated to G . The hyperbolic elements are those which translate a geodesic somewhere in X , while the *real hyperbolic elements* satisfy the property that if they translate one geodesic then they translate every other geodesic which is parallel to the first. We can assign a *slope* to each geodesic, by looking at where it lies in a Weyl chamber in X , and we can use this to assign a slope to hyperbolic elements of G .

Using the geometry of the translated geodesics we are able to prove the following:

Theorem 3.2.12 *Let p be any point in X . For each slope ξ there exists constants $\ell_{\xi}, d_{\xi} > 0$ such that two real hyperbolic elements a and b in G , satisfying $d_X(p, ap), d_X(p, bp) \geq d_{\xi}$, are conjugate if and only if there exists a conjugator $g \in G$ such that:*

$$d_X(p, gp) \leq 2\ell_{\xi}(d_X(p, ap) + d_X(p, bp)).$$

This naturally leads to the following result for lattices:

Corollary 3.2.13 *Let p be any point in X . For each slope ξ there exists a constant $L_{\xi} > 0$ such that two real hyperbolic elements a and b in Γ are conjugate if and only if there exists a conjugator $g \in G$ such that:*

$$d_{\Gamma}(1, g) \leq L_{\xi}(d_{\Gamma}(1, a) + d_{\Gamma}(1, b)).$$

Two questions arise from these results. Firstly, we would like to know whether we can replace ℓ_ξ and d_ξ with a constant which is independent of the slope. In Section 3.2.3 we show that when our conjugate elements are allowed to come from G we cannot relax this condition. Secondly, the conjugator obtained in Corollary 3.2.13 comes from the ambient Lie group. We can ask how far we have to go to find a conjugator from the lattice instead. In Section 3.2.4 we partially answer this question by showing that in certain situations we can find a lattice conjugator while maintaining the linear upper bound on its length.

The second type of element we look at are *unipotent* elements. These are elements which in some finite-dimensional, faithful, linear representation are conjugate to an upper triangular matrix with 1's on the diagonal. For these elements we apply a more algebraic approach. We look at pairs of conjugate unipotent elements for whom, in the matrix representing them, the superdiagonal entries are non-zero. We describe these elements as having non-zero simple entries. When the Lie algebra of G is split we can obtain the following:

Theorem 3.3.13 *Let $u, v \in \Gamma$ be conjugate unipotent elements in the same unipotent subgroup whose simple entries are all non-zero. We can construct $g \in G$ such that $gug^{-1} = v$ and which satisfies*

$$d_G(1, g) \leq L(d_G(1, u) + d_G(1, v))$$

where L is a positive constant depending on the root system associated to G and the unipotent subgroup which contains u and v .

Chapter 2

Solvable Groups

The groups of interest in this section are in the most part finitely generated, recursively presented and solvable. Kharlampovich [Har81] has shown that there exist finitely presented solvable groups of derived length 3 which have unsolvable word problem, and hence unsolvable conjugacy problem. However Noskov [Nos82] showed that all finitely presented metabelian groups have solvable conjugacy problem. This therefore includes the solvable Baumslag–Solitar groups, which we look at in Section 2.3.2.a, but excludes the lamplighter groups, which are the subject of Section 2.1, as they are not finitely presented.

The lamplighter groups, however, are conjugacy separable, but this still isn't enough to show they have solvable conjugacy problem. A group G is said to be conjugacy separable if for each pair of non-conjugate elements u, v in G there is a homomorphism of G onto a finite group H in such a way that the images of u, v in H are not conjugate. Mal'cev [Mal58] and Mostowski [Mos66] independently showed that a finitely presented conjugacy separable group has solvable conjugacy problem; but if the group in question is recursively presented and conjugacy separable, like the lamplighter group, then it is still open as to whether these conditions imply solubility of the conjugacy problem.

A different method though can be applied to the lamplighter group. Matthews [Mat66] showed that if two recursively presented groups A, B have solvable conjugacy problem then their wreath product $A \wr B$ has solvable conjugacy problem if and only if B has solvable power problem. The power problem is solvable if there is an algorithm which determines whether for two elements $x, y \in B$ there exists an integer n such that $y = x^n$. In particular, the lamplighter group $\mathbb{Z}_q \wr \mathbb{Z}$ satisfies these requirements and hence, by the theorem of Matthews, has solvable conjugacy problem.

In Section 2.1 we show that lamplighter groups enjoy a linear conjugacy length function. The method we use for estimating this takes advantage of the geometry of their Cayley graphs, which we may describe as the horocyclic product of two regular trees. When looking at general wreath products the ideas used for the lamplighter groups do not carry through and we need to use new techniques, which are based on the work of Matthews [Mat66]. We obtain an upper bound on the length of conjugators in a group $A \wr B$ which depends on the conjugacy length function of B and on the subgroup distortion of cyclic subgroups of B . The conjugacy length function of A may also appear, but only when dealing with torsion elements of B .

The conjugacy problem in free solvable groups was shown, using the Magnus embedding, to be solvable by Kargapolov and Remeslennikov [KR66] (see also [RS70]) extending the same result for free metabelian groups by Matthews [Mat66]. We show in Section 2.2 that free solvable groups have a cubic conjugacy length function, but in order to do this we show also that the image of the Magnus embedding is undistorted in the ambient wreath product.

A recent paper of Bogopolski, Martino and Ventura [BMV10] analyses the solvability of the conjugacy problem in certain group extensions. The group extensions to which their work applies must satisfy a strong condition on the nature of centralisers in the quotient, most notably the centralisers need to be virtually cyclic. In Section 2.3 we look at similar group extensions, though we replace their condition on the centralisers in the quotient by a geometric condition which allows us to look at more groups. We introduce the notions of a twisted conjugacy length function and a restricted conjugacy length function and see how they relate to the conjugacy length function in a group extension. We apply this work to show that the conjugacy length function in solvable Baumslag–Solitar groups is linear, in certain other abelian-by-cyclic groups it is at most exponential and in certain semidirect products $\mathbb{Z}^d \rtimes \mathbb{Z}^k$ it is linear if $k = 1$ or at most exponential otherwise.

2.1 Lamplighter Groups

In [Sal11] we gave an algebraic method for controlling conjugacy length in lamplighter groups $\mathbb{Z}_q \wr \mathbb{Z}$, using their geometry to understand the word lengths. Here we modify these methods and give a more geometric proof of the same result.

The lamplighter groups are both wreath products and cyclic extensions of abelian groups. Hence we could apply techniques from either Section 2.2 or Section 2.3 to obtain an upper bound for the conjugacy length function of $\mathbb{Z}_q \wr \mathbb{Z}$. There is however a bigger picture, as the lamplighter groups sit inside a larger class of groups, the Diestel–Leader groups. The techniques introduced in this section for obtaining a linear conjugacy length function for lamplighter groups are designed so that they can be generalised to all Diestel–Leader groups. However, the nature of centralisers of elements plays an important role in determining the conjugacy length function of $\mathbb{Z}_q \wr \mathbb{Z}$, and we do not yet know enough about centralisers in a general Diestel–Leader group to apply these techniques and find a bound on their conjugacy length functions as well.

2.1.1 Horocyclic products and Diestel–Leader graphs

We give here a brief introduction to horocyclic products and Diestel–Leader graphs. For a more complete description see [BNW08].

Let T be a simplicial tree and ω a boundary point of T . For any vertex x in T there is a unique geodesic ray emerging from x that is asymptotic to ω . Given a pair of vertices, x, y , the corresponding rays will coincide from some vertex $x \wedge y$ onwards. Using the terminology of [BNW08], $x \wedge y$ is called the *greatest common ancestor* of x and y . After fixing a basepoint x_0 in the vertex set of T we can define a Busemann function \mathfrak{h} on the vertices of T as

$$\mathfrak{h}(y) = d_T(y, x_0 \wedge y) - d_T(x_0, x_0 \wedge y).$$

The k -th horocycle of T based at ω is $H_k = \{y \in T \mid \mathfrak{h}(y) = k\}$.

Given a collection T_1, \dots, T_n of simplicial trees together with a chosen collection of respective Busemann functions $\mathfrak{h}_1, \dots, \mathfrak{h}_n$, we define the *horocyclic product* to be

$$\prod_{i=1}^n \mathfrak{h} T_i = \left\{ (y_1, \dots, y_n) \in T_1 \times \dots \times T_n \mid \sum_{i=1}^n \mathfrak{h}_i(y_i) = 0 \right\}. \quad (2.1)$$

The *Diestel–Leader graph* $\text{DL}(q_1, \dots, q_d)$ is the horocyclic product of trees \mathbb{T}_{q_i} , where \mathbb{T}_q is the $q + 1$ regular tree. When $q_1 = q_2 = \dots = q_d = q$, the corresponding Diestel–Leader graph is also denoted by $\text{DL}_d(q)$.

On a terminological note, the Busemann function described here coincides with the standard definition in a non-positively curved space for a Busemann function at

ω , with the zero level-set passing through x_0 . The level-sets of Busemann functions on manifolds or CAT(0)–spaces are usually called *horospheres*. When dealing with symmetric spaces, the term “horocycle” usually refers to the orbits of a maximal unipotent subgroup of the group of isometries of the symmetric space. However, when we look at the hyperbolic plane, the horocycles and horospheres are in fact equal. And indeed, when we consider the action of the lamplighter group on $\text{DL}_2(q)$, as described below, the orbit of all elements in $\mathbb{Z}_q \wr \mathbb{Z}$ of a unipotent type, that is all elements of the form

$$\gamma = \begin{pmatrix} t^0 & f \\ 0 & 1 \end{pmatrix},$$

are indeed pairs of horocycles as defined above, one horocycle in each tree.

The horocyclic product (2.1) can be recognised as a horosphere in the CAT(0)–space $T_1 \times \dots \times T_n$. That is, the horocyclic product is the level set of a Busemann function defined on $T_1 \times \dots \times T_n$ (see [BGS85]). The ray defining the horosphere is the unit-speed reparametrisation of the ray $(\rho_1(t), \dots, \rho_n(t))$, where ρ_i is a ray in the tree T_i which determines the Busemann function \mathfrak{h}_i .

2.1.2 Diestel–Leader groups

The fact that the Diestel–Leader graph $\text{DL}_2(q)$ is a Cayley graph for the lamplighter group $\mathbb{Z}_q \wr \mathbb{Z}$ was explained in [Woe05]. This is a special case of the following result of Bartholdi, Neuhauser and Woess:

Theorem 2.1.1 (Bartholdi–Neuhauser–Woess [BNW08, (3.14)]). *The Diestel–Leader graph $\text{DL}_d(q)$ is a Cayley graph of a group, denoted $\Gamma_d(\mathfrak{L}_q)$.*

The groups $\Gamma_d(\mathfrak{L}_q)$ are called *Diestel–Leader groups*. We will not discuss them any further, but just note that when $d = 2$ we obtain the lamplighter groups and when $d = 3$ we obtain groups previously considered by Baumslag [Bau72], [Bau74]. The descriptions given below for the lamplighter groups can be extended to the Diestel–Leader groups of more than 2 trees. The word length in Diestel–Leader groups is studied in a recent paper of Stein and Taback [ST12] where they give a formula for its calculation.

Let $\Gamma = \mathbb{Z}_q \wr \mathbb{Z}$ be the lamplighter group. We can represent Γ as the group $\Gamma_2(\mathbb{Z}_q)$ of affine matrices

$$\left\{ \begin{pmatrix} t^k & P \\ 0 & 1 \end{pmatrix} : k \in \mathbb{Z}, P \in \mathbb{Z}_q[t^{-1}, t] \right\},$$

which has a symmetric generating set

$$\left\{ \begin{pmatrix} t & b \\ 0 & 1 \end{pmatrix} : b \in \mathbb{Z}_q \right\}.$$

With respect to this generating set, the Cayley graph of Γ is the Diestel–Leader graph $\text{DL}_2(q)$.

The relationship between $\text{DL}_2(q)$ and these affine matrices comes about by identifying the vertices of the tree \mathbb{T}_q with closed balls in the ring $\mathbb{Z}_q((t))$ of Laurent series

$$f = \sum_{k=-\infty}^{\infty} a_k t^k$$

where $a_k \in \mathbb{Z}_q$ and there exists some $n \in \mathbb{Z}$ such that $a_k = 0$ for all $k < n$. The valuation $v_0(f)$ is defined to be the maximal such n (note that we define $v_0(0) = \infty$), and the absolute value of f is defined to be $q^{-v_0(f)}$. This absolute value determines an ultrametric d_q on $\mathbb{Z}_q((t))$:

$$d_q(f, g) = q^{-v_0(f-g)}, \quad \text{for } f, g \in \mathbb{Z}_q((t)).$$

See [CKW94, §4] for a more complete picture.

Fix a boundary point $\omega \in \partial_\infty \mathbb{T}_q$. For each vertex x in the tree there is a unique geodesic ray in the equivalence class ω which starts at x . Say that the first edge in this ray is *above* x , and all others are *below* it. There are q edges below x , label these with the elements of \mathbb{Z}_q . Apply this process to every vertex in \mathbb{T}_q . We can number the edges in such a way that the geodesic ray from any vertex in \mathbb{T}_q asymptotic to ω passes through finitely many edges which are labelled with non-zero elements from \mathbb{Z}_q . To do this, first fix a basepoint $o \in \mathbb{T}_q$ and label with 0 each edge in the geodesic ray emerging from o which is asymptotic to ω . The remaining edges can be labelled arbitrarily, provided the labelling agrees with the description above. Every geodesic ray asymptotic to ω will eventually merge with the geodesic ray emitted from o . Hence, after a finite distance, it will travel only along edges labelled by 0.

Suppose the vertex x is in the k -th horocycle with respect to ω . Then we can read off an element of $\mathbb{Z}_q((t))$ from the ray emerging from x : along each edge is a label, and each edge connects two distinct, adjacent horocycles. The ray consists of exactly one edge between the j -th horocycle and the $(j-1)$ -th horocycle for $j \leq k$, suppose it is labelled by $a_j \in \mathbb{Z}_q$. We assign to x the Laurent series

$$f_x = \sum_{j=-\infty}^k a_j t^j. \tag{2.2}$$

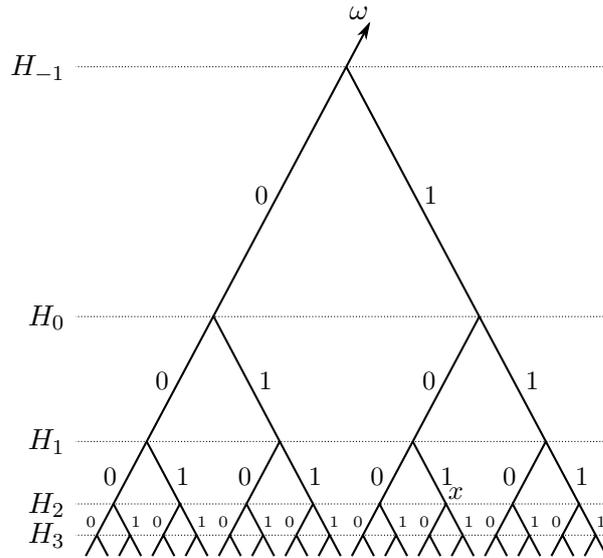


Figure 2.1: The vertex labelled x is associated with the Laurent series $f_x = t^2 + 1$, and identified with the ball $B(f_x, q^{-2})$.

Suppose we are given a closed ball

$$B(f, q^{-n}) = \{P \in \mathbb{Z}_q((t)) \mid v_0(f - P) \geq n\}$$

in $\mathbb{Z}_q((t))$ with respect to the ultrametric determined by the absolute value defined above. We identify this with a vertex in the tree as follows: the radius q^{-n} of the ball tells us which horocycle the vertex is in, while the Laurent series f tells us precisely which vertex to take. To be more precise, given a radius q^{-n} , this tells us we should be in the n -th horocycle. Hence we only care about the coefficients in f for the terms t^k where $k \leq n$. So we obtain f' from f by setting the coefficient of t^m to be zero for each $m > n$. That is, if

$$f = \sum_{j=-\infty}^{\infty} a_j t^j$$

then

$$f' = \sum_{j=-\infty}^n a_j t^j.$$

Then we find the vertex x in \mathbb{T}_q for which $f_x = f'$, where f_x is as in (2.2), and identify this vertex with $B(f, q^{-n})$.

The above identification describes a map

$$\mathcal{V}(\mathbb{T}_q) \rightarrow \mathcal{B} = \{B(f, q^{-n}) : f \in \mathbb{Z}_q((t)), n \in \mathbb{Z}\}$$

where $\mathcal{V}(\mathbb{T}_q)$ is the vertex set of \mathbb{T}_q . Note that $P \in B(f, q^{-n})$ is equivalent to saying that the coefficients of q^r in P and f agree for all $r < n$, and hence it is also equivalent to $B(f, q^{-n}) = B(P, q^{-n})$. It follows from this that the above map is a bijection.

The group $\Gamma_2(\mathbb{Z}_q)$ has underlying set

$$\Gamma_2(\mathbb{Z}_q) = \left\{ \begin{pmatrix} t^s & P \\ 0 & 1 \end{pmatrix} \mid P \in \mathbb{Z}_q[t^{-1}, t] \text{ and } n \in \mathbb{Z} \right\}.$$

An element of $\Gamma_2(\mathbb{Z}_q)$ acts on \mathcal{B} as

$$\begin{pmatrix} t^s & P \\ 0 & 1 \end{pmatrix} \cdot B(f, q^{-n}) = B(P + t^s f, q^{-n-s}).$$

Before proceeding further, we make the following observation concerning the height of a common ancestor under the action of $\Gamma_2(\mathbb{Z}_q)$.

Lemma 2.1.2. *Let x, y be vertices in \mathbb{T}_q and suppose $\gamma = \begin{pmatrix} t^s & P \\ 0 & 1 \end{pmatrix} \in \Gamma_2(\mathbb{Z}_q)$. Then*

$$\mathfrak{h}(\gamma x \wedge \gamma y) = \mathfrak{h}(x \wedge y) + s.$$

Proof. Let x be identified with $B(X, q^{-\alpha})$ and y with $B(Y, q^{-\beta})$. The common ancestor of x and y must be above both x and y , and will correspond to the smallest ball in $\mathbb{Z}_q((t))$ containing both X and Y . This tells us that

$$\mathfrak{h}(x \wedge y) = \min\{v_0(X - Y) - 1, \mathfrak{h}(x), \mathfrak{h}(y)\}.$$

Consider $\gamma x \wedge \gamma y$. The vertex γx will be identified with the ball $B(P + t^s X, q^{-\alpha-s})$ and γy with $B(P + t^s Y, q^{-\beta-s})$. We obtain the analogous expression to the above equation for $\mathfrak{h}(\gamma x \wedge \gamma y)$, noting that the terms in P both cancel, leaving

$$\mathfrak{h}(\gamma x \wedge \gamma y) = \min\{v_0(t^s X - t^s Y) - 1, \mathfrak{h}(\gamma x), \mathfrak{h}(\gamma y)\}.$$

Since $v_0(t^s X - t^s Y) = v_0(X - Y) + s$, $\mathfrak{h}(\gamma x) = \mathfrak{h}(x) + s$ and similarly for γy , the Lemma holds. \square

We have so far described the action of the lamplighter group on only one tree. We must ask how it acts on the second tree in the horocyclic product. To answer this, instead of considering f as an element of $\mathbb{Z}_q((t))$ we see it as an element in $\mathbb{Z}_q((t^{-1}))$; we can do this because $f \in \mathbb{Z}_q[t^{-1}, t]$. The (negative of the) valuation of $f \in \mathbb{Z}_q((t^{-1}))$ will be denoted by $v_0^-(f)$ and will be equal to the largest integer k such that the coefficient of t^k is non-zero. The absolute value of f will be $q^{v_0^-(f)}$. We can

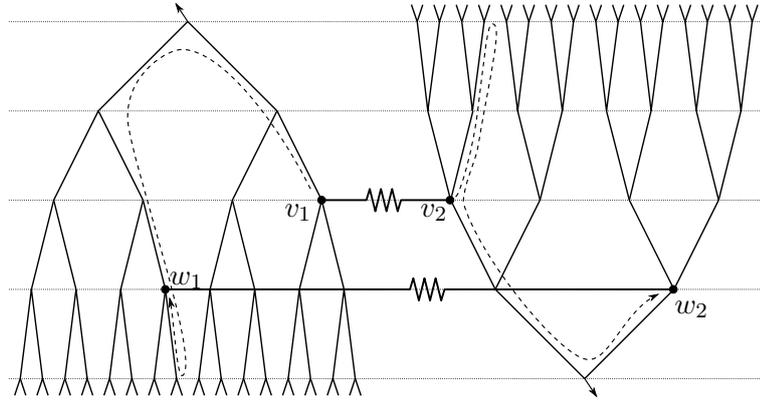


Figure 2.2: Moving in $DL_2(q)$ from (v_1, v_2) to (w_1, w_2) : the dotted lines demonstrate the routes each bead needs to take. When we move the left-hand bead we need to keep the elastic horizontal, so the right hand bead will have to move upwards as well. Similarly, once the left-hand bead has reached w_1 , we need to keep moving it while we move the right-hand bead to get it to w_2 .

identify vertices of the second tree with closed balls $B^-(f, q^{-n})$ in $\mathbb{Z}_q((t^{-1}))$ in much the same way as we did before, but this time round the n -th horocycle will instead consist of the balls of radius q^{-n-1} . The reason for this slight adjustment is so that we consider each coefficient in f exactly once. The action on the horocyclic product is

$$\begin{pmatrix} t^s & P \\ 0 & 1 \end{pmatrix} \cdot (B(0, q^0), B^-(0, q^{-1})) = (B(P, q^{-s}), B^-(P, q^{s-1})).$$

This takes into account each coefficient exactly once because the ball $B(P, q^{-s})$ is determined by the coefficients in P of t^k for each $k \leq s$. Meanwhile, $B^-(P, q^{s-1})$ is determined by coefficients in P of $(t^{-1})^k$ for $k \leq s-1$, or equivalently t^j for each $j > s$.

One can visualise the action of the lamplighter group on $DL_2(q)$ with two beads and a piece of elastic. Woess has previously described this method, see for example [Woe05, §2]. The idea is as follows: take your two trees and, as in Figure 2.2, draw the first tree on the left-hand side with the chosen boundary point at the top and the second tree on the right with the boundary point below. Place the two trees so that the two 0-th horocycles appear on the same horizontal line, and in general the n -th horocycle in the left-hand tree lines up with the $(-n)$ -th horocycle on the right. In each tree fix a basepoint in the 0-th horocycle. Place one bead on each tree, one at each of the basepoints. Connect the two beads by a piece of elastic (we may assume the elastic can stretch infinitely long). By construction, the elastic will be horizontal.

Any point in $DL_2(q)$ can now be obtained by moving the two beads in their respective trees in any way we like, provided the piece of elastic is always kept horizontal.

2.1.3 Conjugacy length for lamplighter groups

In order to estimate the conjugacy length function we need to be able to understand the word lengths of elements in $\mathbb{Z}_q \wr \mathbb{Z}$.

2.1.3.a Word length

The word length of elements in a general Diestel–Leader group has been studied by Stein and Taback [ST12]. They give a formula for the word length of an element by looking at the climb and fall of a geodesic path in each tree from the basepoint to its image under the action of the element. Let o_i denote the basepoint of the i -th tree. For $g \in \Gamma_d(\mathfrak{L}_q)$, denote by $m_i(g)$ the length of the climb of the geodesic from o_i to go_i and $l_i(g)$ the length of the fall. More concretely we mean:

$$m_i(g) = d(o_i, o_i \wedge go_i), \quad l_i(g) = d(go_i, o_i \wedge go_i).$$

A consequence of Stein and Taback’s formula are the following upper and lower bounds for the word length of g :

$$\sum_{i=1}^d m_i(g) \leq |g| \leq 3 \sum_{i=1}^d m_i(g)$$

where $|g|$ is the word length of u with respect to the generating set described above. Each generator of $\Gamma_d(\mathfrak{L}_q)$ corresponds to an instruction: go up in tree i and down in tree j . When given a word, these instructions are read from left to right.

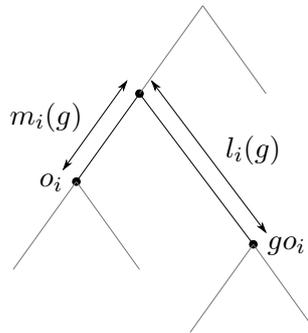


Figure 2.3: The climb and fall of g in the i -th tree.

The notion of climbing and falling can be extended to paths describing the concatenation of words. In particular, for $g, h \in \Gamma_d(\mathfrak{L}_q)$ let

$$m_{h,i}(g) = d(ho_i, ho_i \wedge go_i), \quad l_{h,i} = d(go_i, ho_i \wedge go_i).$$

Stein and Taback show that $|g^{-1}h|$ behaves in the same way but instead with respect to the climb and fall functions $m_{h,i}(g)$ and $l_{h,i}(g)$. Hence

$$\sum_{i=1}^d m_{h,i}(g) \leq |g^{-1}h| \leq 3 \sum_{i=1}^d m_{h,i}(g). \quad (2.3)$$

Please note that in these bounds for $|g^{-1}h|$ we may replace $m_{h,i}$ with $l_{h,i}$ since their sums are equal.

A consequence of Lemma 2.1.2 is that we can measure the length of the word by looking at its action on different points in each tree:

Lemma 2.1.3. *For every $g, h \in \Gamma_d(\mathfrak{L}_q)$ and each $i = 1, \dots, d$ we have the following:*

$$m_i(g) = m_{h,i}(hg), \quad l_i(g) = l_{h,i}(hg).$$

Proof. By Lemma 2.1.2, in each tree, the height of the common ancestor $o_i \wedge go_i$ never strays too far from the height of $ho_i \wedge hgo_i$. To be precise:

$$\mathfrak{h}_i(o_i \wedge go_i) = \mathfrak{h}_i(ho_i \wedge hgo_i) + r_i$$

where $r_i = \mathfrak{h}_i(ho_i)$. Hence

$$\begin{aligned} m_i(g) &= d(o_i, o_i \wedge go_i) \\ &= -\mathfrak{h}_i(o_i \wedge go_i) \\ &= -\mathfrak{h}_i(ho_i \wedge hgo_i) - r_i \end{aligned}$$

But $d(ho_i, ho_i \wedge hgo_i) = \mathfrak{h}_i(ho_i) - \mathfrak{h}_i(ho_i \wedge hgo_i)$, so we get $m_i(g) = m_{h,i}(hg)$. The result for $l_i(g)$ follows from the result for $m_i(g)$, the relationships:

$$m_i(g) - l_i(g) = \mathfrak{h}_i(go_i) \quad \text{and} \quad m_{h,i}(hg) - l_{h,i}(hg) = \mathfrak{h}_i(hgo_i) - \mathfrak{h}_i(ho_i)$$

and the fact that $\mathfrak{h}_i(go_i) = \mathfrak{h}_i(hgo_i) - \mathfrak{h}_i(ho_i)$. □

2.1.3.b The conjugacy length function

Theorem 2.1.4. *Let $\Gamma = \mathbb{Z}_q \wr \mathbb{Z}$. Then the conjugacy length function for Γ is linear. In particular, with respect to the generating set described above,*

$$\text{CLF}_\Gamma(n) \leq 3n.$$

Proof. We use the representation of elements of Γ as matrices. Suppose

$$u = \begin{pmatrix} t^s & P \\ 0 & 1 \end{pmatrix}, \quad v = \begin{pmatrix} t^r & Q \\ 0 & 1 \end{pmatrix}, \quad \gamma = \begin{pmatrix} t^k & f \\ 0 & 1 \end{pmatrix}.$$

Then, by direct calculation, $u\gamma = \gamma v$ if and only if the following equations hold:

$$s + k = k + r \tag{2.4}$$

$$P + t^s f = f + t^k Q \tag{2.5}$$

We therefore split into the two cases according to whether r is zero or not.

CASE 1: $r = 0$.

This corresponds to the case when u and v map each basepoint to a vertex of the same height. Equation (2.5) becomes $P = t^k Q$, so we may set $f = 0$. This means that γ will act on each tree by either a sequence of consecutive up movements or a sequence of consecutive down movements, but never a mixture of both. Suppose γ acts in a purely downwards motion on tree i . Figure 2.4 demonstrates what is happening in this tree.

As long as v is non-trivial, we may assume that $vo_i \neq o_i$ because otherwise we can just replace o_i with some other basepoint corresponding to a ball $B(0, q^{-\alpha})$, in which

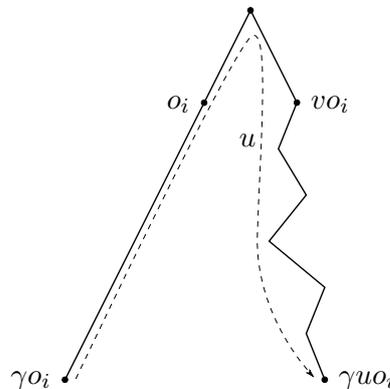


Figure 2.4: The action of u on γo_i in Case 1.

α is large enough to ensure that the action of v on this ball is non-trivial. Since the action of γ on the i -th tree corresponds to a consecutive list of downward movements, the path read out by u from γo_i to $\gamma u o_i$ must pass through $v o_i$. Hence $m_{\gamma,i}(\gamma u) \geq |k|$. Using equation 2.3 we then obtain

$$|u| = |(\gamma u)^{-1} \gamma| \geq m_{\gamma,i}(\gamma u) \geq |k| = |\gamma|$$

giving a linear upper bound in this case.

CASE 2: $r \neq 0$.

By exchanging u and v with their inverses if necessary, we may assume that $r > 0$. The important step here is to pick the right conjugator. Take any conjugator γ' , satisfying $u\gamma = \gamma v$, and premultiply it by a suitable power of u so that we obtain an element $\gamma = u^m \gamma'$, written as above, with $0 \leq k < r$. We claim that for each i

$$m_i(\gamma) \leq \max\{m_i(u), m_i(v)\} + r.$$

To prove this claim we will show that if this were not true then $\gamma u o_i$ and $v \gamma o_i$ would have to be on different branches of their respective trees.

If we were to assume that $m_i(u) < m_i(\gamma) - r$ then we would have ensured that $o_i \wedge \gamma u o_i = o_i \wedge \gamma o_i$ — since, by Lemma 2.1.3, $d(\gamma o_i, \gamma o_i \wedge o_i) \leq d(\gamma o_i, \gamma o_i \wedge o_i)$. On the other hand, the common ancestor $v o_i \wedge v \gamma o_i$ will lie in a different horocycle to $o_i \wedge \gamma o_i$. Indeed, considering the values of the Busemann functions:

$$\mathfrak{h}_i(o_i \wedge \gamma o_i) = -m_i(\gamma) \neq -m_i(\gamma) \pm r = \mathfrak{h}_i(v o_i \wedge v \gamma o_i).$$

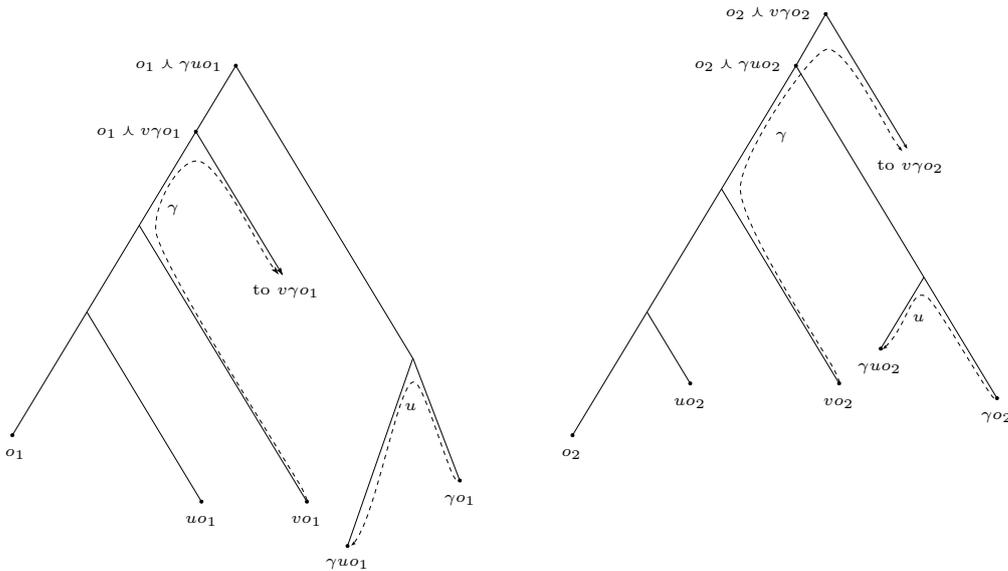


Figure 2.5: The common ancestor $o_i \wedge \gamma u o_i$ lies in a different horocycle to $o_i \wedge v \gamma o_i$.

If $m_i(\gamma) > m_i(v) + r$ then this guarantees that $vo_i \wedge v\gamma o_i = o_i \wedge v\gamma o_i$. Hence, $o_i \wedge \gamma u o_i \neq o_i \wedge v\gamma o_i$ since they lie in different horocycles, contradicting the fact that $\gamma u = v\gamma$.

With the claim justified, we see that

$$|\gamma| \leq 3m_1(\gamma) + 3m_2(\gamma) \leq 3 \max\{m_1(u), m_1(v)\} + 3 \max\{m_2(u), m_2(v)\} + 3r.$$

The last term is bounded above by $3n$, where $n = |u| + |v|$, thus proving the Theorem. \square

The nature of this proof gives an idea of how it may be transferred to prove a similar statement for more general Diestel-Leader groups $\Gamma_d(\mathfrak{L}_q)$, groups for which the Cayley graph is the horocyclic product of more than two trees. The main obstacle in doing this seems to be in understanding the nature of the centralisers in $\Gamma_d(\mathfrak{L}_q)$. In case 2 of the above proof, the first thing we did was to premultiply an arbitrary conjugator by a suitable power of u so that we obtained a conjugator γ for which $r = \mathfrak{h}_1(\gamma o_1) \in [0, \mathfrak{h}_1(u o_1))$. This was enough to ensure also that $\mathfrak{h}_2(\gamma o_2) \in (\mathfrak{h}_2(u o_2), 0]$, then allowing us to proceed and show γ is of bounded length.

However, if we have more than two trees, naively forcing γo_i into the correct horocycle in one tree is not enough to control which horocycles the other vertices γo_j will lie in. Suppose $r_i = \mathfrak{h}_i(u o_i)$ for $i = 1, \dots, d$. In the case when $d = 2$ we have $r_1 = r = -r_2$. Given any $\gamma \in \Gamma_2(\mathfrak{L}_q)$ it is clear that we can premultiply it by a suitable power of u and get that $|\mathfrak{h}_i(\gamma o_i)| < |r|$ for $i = 1, 2$. However, if $d > 2$ then using $\langle u \rangle$ on its own is not enough to bound each $\mathfrak{h}_i(\gamma o_i)$. In order to do this we would have to extend our reach into the wider depths of the centraliser $Z_\Gamma(u)$ of u inside $\Gamma = \Gamma_d(\mathfrak{L}_q)$.

We will observe a similar problem in Section 2.3.4 when studying the groups $\mathbb{Z}^d \rtimes \mathbb{Z}^k$. When $k = 1$ we can deal with this problem in the same way as we did here. But when $k > 1$ we need to look at the projection of the centraliser of u into the \mathbb{Z}^k component. In Theorem 2.3.15 we will show that, essentially, the fundamental domain of this projection in \mathbb{R}^k is compact and we can control the diameter by an exponential function. In the case of $\Gamma_d(\mathfrak{L}_q)$, we consider the projection onto \mathbb{Z}^{d-1} , which we see as

$$\{(\mathfrak{h}_1(g o_1), \dots, \mathfrak{h}_d(g o_d)) \in \mathbb{Z}^d \mid g \in \Gamma_d(\mathfrak{L}_q)\}$$

and ask if we can find a similar control on the fundamental domain of this projection inside \mathbb{R}^{d-1} .

2.2 Wreath Products and Free Solvable Groups

In this section we look at the effect that taking wreath products has on conjugacy length. The lamplighter group that we saw in the previous section is a special case of this, but we used very different techniques for it than we will use for general wreath products. The main motivation for this section is to find a control on the conjugacy length function of free solvable groups. The majority of the content of this section is taken from [Sal12].

The structure is as follows: we begin in Section 2.2.1 with the appropriate preliminary definitions, including a brief account of Fox calculus. This section builds up to Theorem 2.2.5 which asserts that the image of the Magnus embedding is undistorted in the ambient wreath product. In Section 2.2.2 we study conjugacy in wreath products, first obtaining an upper bound and then a lower bound for the length of short conjugators. The subgroup distortion of cyclic subgroups of $S_{r,d}$ is discussed in Section 2.2.3, before we move onto conjugacy length in free solvable groups in the penultimate section.

Finally, in Section 2.2.5, we apply Theorem 2.2.5 to study the L_p compression exponents for free solvable groups. Compression exponents were first introduced by Guentner and Kaminker [GK04], building on the idea of uniform embeddings introduced by Gromov [Gro93] and Yu [Yu00]. In particular we show that free solvable groups have non-zero L_p compression exponent.

2.2.1 Preliminaries

2.2.1.a Fox calculus

In Section 2.2.1.b we will introduce the Magnus embedding. This is a classical tool that plays an important role in the study of free solvable groups $S_{r,d}$. In order to make effective use of the Magnus embedding we need to understand Fox derivatives. These were introduced by Fox in the 1950's in a series of papers [Fox53], [Fox54], [Fox56], [CFL58], [Fox60].

Recall that a derivation on a group ring $\mathbb{Z}(G)$ is a mapping $\mathcal{D} : \mathbb{Z}(G) \rightarrow \mathbb{Z}(G)$ which satisfies the following two conditions for every $a, b \in \mathbb{Z}(G)$:

$$\begin{aligned}\mathcal{D}(a + b) &= \mathcal{D}(a) + \mathcal{D}(b) \\ \mathcal{D}(ab) &= \mathcal{D}(a)\varepsilon(b) + a\mathcal{D}(b)\end{aligned}$$

where $\varepsilon : \mathbb{Z}(G) \rightarrow \mathbb{Z}$ sends each element of G to 1. That is, for $g_1, \dots, g_n \in G$ and $\kappa_1, \dots, \kappa_n$ integers, $\varepsilon(\kappa_1 g_1 + \dots + \kappa_n g_n) = \kappa_1 + \dots + \kappa_n$.

Suppose $G = F$, the free group on generators $X = \{x_1, \dots, x_r\}$. For each generator we can define a unique derivation $\frac{\partial}{\partial x_i}$ which satisfies

$$\frac{\partial x_j}{\partial x_i} = \delta_{ij}$$

where δ_{ij} is the Kronecker delta. Any derivation \mathcal{D} can be expressed as a linear combination of these: for each \mathcal{D} there exists some elements $k_i \in \mathbb{Z}(F)$ such that

$$\mathcal{D}(a) = \sum_{i=1}^n k_i \frac{\partial a}{\partial x_i}$$

for each $a \in \mathbb{Z}(F)$.

Fox describes the following Lemma as the “fundamental formula” and it can be found in [Fox53, (2.3)].

Lemma 2.2.1 (Fundamental formula of Fox calculus). *Let $a \in \mathbb{Z}(F)$. Then*

$$a - \varepsilon(a)1 = \sum_{i=1}^r \frac{\partial a}{\partial x_i} (x_i - 1).$$

Fox derivatives also accept a form of integration, see [CF63, Ch.VII (2.10)]. In particular, given $\beta_1, \dots, \beta_r \in \mathbb{Z}(F)$ one can find $c \in \mathbb{Z}(F)$ such that $\frac{\partial c}{\partial x_i} = \beta_i$ for each i . The element c is unique up to addition of scalar multiples of the identity.

Given a normal subgroup N in F and a derivation \mathcal{D} of $\mathbb{Z}(F)$ we can define a derivation $\mathcal{D}^* : \mathbb{Z}(F) \rightarrow \mathbb{Z}(F/N)$ through the composition of maps

$$\begin{array}{ccc} \mathbb{Z}(F) & \xrightarrow{\mathcal{D}} & \mathbb{Z}(F) \\ & \searrow \mathcal{D}^* & \downarrow \alpha^* \\ & & \mathbb{Z}(F/N) \end{array}$$

where α^* is the extension of the quotient homomorphism $\alpha : F \rightarrow F/N$.

The following Lemma can be deduced from the Magnus embedding, but it also follows from [Fox53, (4.9)].

Lemma 2.2.2. *Let $g \in F$. Then $\mathcal{D}^*(g) = 0$ for every derivation \mathcal{D} if and only if $g \in N' = [N, N]$.*

Strictly speaking, α^* is a map from $\mathbb{Z}(F)$ to $\mathbb{Z}(F/N)$, but if instead we consider the canonical map $\bar{\alpha} : F/N' \rightarrow F/N$ then $\bar{\alpha}^*$ becomes a map from $\mathbb{Z}(F/N')$ to $\mathbb{Z}(F/N)$. Understanding the kernel of $\bar{\alpha}^*$ will be helpful in Section 2.2.4.

Lemma 2.2.3 (See also Gruenberg [Gru67, §3.1 Theorem 1]). *An element of the kernel of $\bar{\alpha}^* : \mathbb{Z}(F/N') \rightarrow \mathbb{Z}(F/N)$ can be written in the form*

$$\sum_{j=1}^m r_j (h_j - 1)$$

for some integer m , where $r_j \in F/N'$ and $h_j \in N/N'$ for each $j = 1, \dots, m$.

Proof. Take an arbitrary element a in the kernel of $\bar{\alpha}^*$. Suppose we can write

$$a = \sum_{g \in F/N'} \beta_g g$$

where $\beta_g \in \mathbb{Z}$ for each $g \in F/N'$. Fix a coset xN . Then

$$\sum_{\bar{\alpha}(g)=xN} \beta_g = 0$$

since this is the coefficient of xN in $\bar{\alpha}^*(a)$. Notice that $\bar{\alpha}(g) = xN$ if and only if there is some $h \in N$ such that $g = xh$. Thus the sum can be rewritten as

$$\sum_{h \in N \setminus \{1\}} \beta_{xh} = -\beta_x.$$

This leads us to

$$\sum_{\bar{\alpha}(g)=xN} \beta_g g = \sum_{h \in N \setminus \{1\}} \beta_{xh} x(h - 1)$$

which implies the Lemma after summing over all left-cosets. \square

2.2.1.b The Magnus embedding

Let F be the free group of rank r with generators $X = \{x_1, \dots, x_r\}$ and let N be a normal subgroup of F . The Magnus embedding gives a way of recognising F/N' , where N' is the derived subgroup of N , as a subgroup of the wreath product $M(F/N) = \mathbb{Z}^r \wr F/N$.

Consider the group ring $\mathbb{Z}(F/N)$ and let \mathcal{R} be the free $\mathbb{Z}(F/N)$ -module with generators t_1, \dots, t_r . We define a homomorphism

$$\varphi : F \longrightarrow M(F/N) = \left(\begin{array}{c|c} F/N & \mathcal{R} \\ \hline 0 & 1 \end{array} \right) = \left\{ \left(\begin{array}{cc} g & a \\ 0 & 1 \end{array} \right) \mid g \in F/N, a \in \mathcal{R} \right\}$$

by

$$\varphi(w) = \begin{pmatrix} \alpha(w) & \frac{\partial w}{\partial x_1} t_1 + \dots + \frac{\partial w}{\partial x_r} t_r \\ 0 & 1 \end{pmatrix}$$

where α is the quotient homomorphism $\alpha : F \rightarrow F/N$. Magnus [Mag39] recognised that the kernel of φ is equal to N' and hence φ induces an injective homomorphism from F/N' to $M(F/N)$ which is known as the *Magnus embedding*. In the rest of this paper we will use φ to denote both the homomorphism defined above and the Magnus embedding it induces.

Given $w \in F$, its image under the Magnus embedding can be identified with $(f, b) \in \mathbb{Z}^r \wr F/N$ in the following way: we take $b = \alpha(w)$ and f will be the function $f^w = (f_1^w, \dots, f_r^w)$, where for each i the function $f_i^w : F/N \rightarrow \mathbb{Z}$ satisfies the equation

$$\sum_{g \in F/N} f_i^w(g)g = \frac{\partial^* w}{\partial x_i} \in \mathbb{Z}(F/N).$$

Let $d_{F/N'}$ denote the word metric in F/N' with respect to the generators determined by the elements of X and let d_M denote the word metric on $M(F/N)$ with respect to the generating set

$$\left\{ \begin{pmatrix} \alpha(x_1) & 0 \\ 0 & 1 \end{pmatrix}, \dots, \begin{pmatrix} \alpha(x_r) & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & t_1 \\ 0 & 1 \end{pmatrix}, \dots, \begin{pmatrix} 1 & t_r \\ 0 & 1 \end{pmatrix} \right\}.$$

Note that this generating set is the same as that used for Lemma 1.4.1. The aim is to compare the metrics $d_{F/N'}$ and d_M . We first give a result of Droms, Lewin and Servatius [DLS93, Theorem 2] on the word-length of elements in F/N' . In order to do so we need to set up some notation.

Let E be the edge set of the Cayley graph $\text{Cay}(F/N, X)$ of F/N with respect to the generating set determined by X . Given a word w on X we obtain a path ρ_w in $\text{Cay}(F/N, X)$ labelled by w . Define a function $\pi_w : E \rightarrow \mathbb{Z}$ such that for each edge $(g, gx) \in E$ the value of $\pi_w(g, gx)$ is equal to the net number of times the path ρ_w traverses this edge — for each time the path travels from g to gx count $+1$; for each time the path goes from gx to g count -1 . Since the path is finite, π_w has finite support.

Let $\text{Supp}(\pi_w)$ denote the subgraph of $\text{Cay}(F/N, X)$ containing all edges e such that $\pi_w(e) \neq 0$. Consider a new path $\sigma(\pi_w)$ which is a path travelling through every point in $\text{Supp}(\pi_w) \cup \{1\}$ so that it minimises the number of edges not contained in $\text{Supp}(\pi_w)$. Let $W(\pi_w)$ denote this number.

Lemma 2.2.4 (Droms–Lewin–Servatius [DLS93]). *Let w be a word on generators X which determines the element $g \in F/N'$. Then*

$$d_{F/N'}(1, g) = \sum_{e \in E} |\pi_w(e)| + 2W(\pi_w).$$

The next Theorem tells us how word lengths behave under the Magnus embedding. We see in particular that with respect to the chosen generating sets described above for F/N' and for the wreath product $M(F/N)$ the Magnus embedding is 2-bi-Lipschitz.

Geometrically, we take a geodesic word w for $g \in F/N'$ and construct a path ρ_w from this word in the Cayley graph of F/N . The length of ρ_w is equal to $d_{F/N'}(1, g)$ so we need to compare its length with the size of $\varphi(g)$ in the wreath product. We will see how, if $\varphi(g) = (f^w, \alpha(w))$, then the function f^w describes the route which ρ_w takes, telling us the net number of times ρ_w transverses each edge. From this we can deduce a relationship between the size of g and the size of $\varphi(g)$.

Theorem 2.2.5. *The subgroup $\varphi(F/N')$ is undistorted in $M(F/N)$. To be precise, for each $g \in F/N'$*

$$\frac{1}{2}d_{F/N'}(1, g) \leq d_M(1, \varphi(g)) \leq 2d_{F/N'}(1, g).$$

Proof. The aim is to compare the word-lengths given by Lemma 2.2.4 and Lemma 1.4.1. Let w be a word on X representing $g \in F/N$. The image of g under the Magnus embedding is $\varphi(w) = (f^w, \alpha(w))$, with $f^w = (f_1^w, \dots, f_r^w)$ satisfying

$$\sum_{g \in F/N} f_i^w(g)g = \frac{\partial^* w}{\partial x_i} \in \mathbb{Z}(F/N).$$

We claim that $f_i^w(g) = \pi_w(g, gx_i)$, and will prove this by induction on the word-length of w . If $w = x_j$ then $\frac{\partial^* w}{\partial x_i} = \delta_{ij}$. The path ρ_w consists of just one edge: $(1, x_j)$. Hence $\pi_w(g, gx_i)$ is zero everywhere except when $g = 1$ and $i = j$, where it takes the value 1. Thus, in this case, the claim holds. If $w = x_j^{-1}$ then $\frac{\partial^* w}{\partial x_i} = -\delta_{ij}x_j^{-1}$. The path ρ_w this time consists of the edge $(x_j^{-1}, 1)$ and one can check that the claim holds here too.

Now suppose w has length at least 2 and that the claim holds for all words shorter than w . Suppose also that w is of the form $w = w'x_j^\varepsilon$ where w' is a non-trivial word and $\varepsilon = \pm 1$. Then

$$\frac{\partial^*(w'x_j^\varepsilon)}{\partial x_i} = \frac{\partial^* w'}{\partial x_i} + \alpha(w') \frac{\partial^* x_j^\varepsilon}{\partial x_i}$$

and it follows that $f_i^w(g) = f_i^{w'}(g)$ whenever $i \neq j$. When $i = j$ we get

$$f_i^w(g) = \begin{cases} f_i^{w'}(g) & \text{if } \varepsilon = 1 \text{ and } g \neq \alpha(w'), \text{ or } \varepsilon = -1 \text{ and } g \neq \alpha(w), \\ f_i^{w'}(g) + 1 & \text{if } \varepsilon = 1 \text{ and } g = \alpha(w'), \\ f_i^{w'}(g) - 1 & \text{if } \varepsilon = -1 \text{ and } g = \alpha(w). \end{cases}$$

Meanwhile, ρ_w is obtained from $\rho_{w'}$ by attaching one extra edge on to its final vertex, namely the edge $(w', w'x_i)$ if $\varepsilon = 1$ or (w, wx_i) is $\varepsilon = -1$. Hence $\pi_w(g, gx_i) = \pi_{w'}(g, gx_i)$ whenever $i \neq j$ and when $i = j$ we get

$$\pi_w(g, gx_i) = \begin{cases} \pi_{w'}(g, gx_i) & \text{if } \varepsilon = 1 \text{ and } g \neq \alpha(w'), \text{ or } \varepsilon = -1 \text{ and } g \neq \alpha(w) \\ \pi_{w'}(g, gx_i) + 1 & \text{if } \varepsilon = 1 \text{ and } g = \alpha(w'), \\ \pi_{w'}(g, gx_i) - 1 & \text{if } \varepsilon = -1 \text{ and } g = \alpha(w). \end{cases}$$

Thus, applying the inductive hypothesis gives $f_i^w(g) = \pi_w(g, gx_i)$ and the claim therefore holds for all words w .

From Lemma 1.4.1 the word-length in $M(F/N)$ is given by

$$d_M(1, (f^w, \alpha(w))) = K(\text{Supp}(f^w), \alpha(w)) + \sum_{y \in F/N} \|f^w(y)\|$$

where $\|\cdot\|$ is the ℓ_1 -norm on \mathbb{Z}^r . The above expression of f^w in terms of π_w leads us to the equation

$$\sum_{y \in F/N} \|f^w(y)\| = \sum_{e \in E} |\pi_w(e)|.$$

Since an edge e is in $\text{Supp}(\pi_w)$ only if one of its ends is in $\text{Supp}(f^w)$, we see that $\text{Supp}(f^w)$ is contained in the subgraph $\text{Supp}(\pi_w)$. Take a path q starting at 1 and travelling through every point in $\text{Supp}(f^w)$, in particular we may take q to be a path realising $K(\text{Supp}(f^w), \alpha(w))$. Any edge in $\text{Supp}(\pi_w)$ which is not in this path must have one vertex lying in the path q . Adding these edges to q (along with the corresponding backtracking) gives a new path q' passing through every point of $\text{Supp}(\pi_w) \cup \{1\}$. Note that every edge in q' that is not in $\text{Supp}(\pi_w)$ was already in q . Hence the length of q is bounded below by the size of $W(\pi_w)$. In particular $W(\pi_w) \leq K(\text{Supp}(f^w), \alpha(w))$ and hence

$$\frac{1}{2}d_{F/N'}(1, g) \leq d_M(1, \varphi(g)).$$

On the other hand, suppose $w = x_{i_1}^{\varepsilon_1} \dots x_{i_m}^{\varepsilon_m}$ is a minimal word representing g , that is $d_{F/N'}(1, g) = m$. Then $\varphi(g) = \varphi(x_{i_1})^{\varepsilon_1} \dots \varphi(x_{i_m})^{\varepsilon_m}$ gives an expression for $\varphi(g)$ in

terms of $2m$ generators, since for each i_j

$$\begin{aligned}\varphi(x_{i_j}) &= \begin{pmatrix} x_{i_j} & t_{i_j} \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & t_{i_j} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{i_j} & 0 \\ 0 & 1 \end{pmatrix}.\end{aligned}$$

Hence $d_M(1, \varphi(g)) \leq 2m$ and the result follows. \square

2.2.2 Conjugacy in wreath products

Let A and B be finitely generated groups. By a result of Matthews [Mat66], when A and B are recursively presented with solvable conjugacy problem and when B also has solvable power problem, the group $\Gamma = A \wr B$ has solvable conjugacy problem. In what follows we will not need these assumptions, we will only assume that A and B are finitely generated.

Fix $b \in B$ and let $\{t_i \mid i \in I\}$ be a set of right-coset representatives for $\langle b \rangle$ in B . We associate to this a family of maps $\pi_{t_i}^{(z)} : A^{(B)} \rightarrow A$ for each z in B as follows:

$$\pi_{t_i}^{(z)}(f) = \begin{cases} \prod_{j=0}^{N-1} f(z^{-1}b^j t_i) & \text{for } b \text{ of finite order } N \\ \prod_{j=-\infty}^{\infty} f(z^{-1}b^j t_i) & \text{for } b \text{ of infinite order.} \end{cases}$$

The products above are taken so that the order of multiplication is such that $f(t_i b^j z^{-1})$ is to the left of $f(t_i b^{j-1} z^{-1})$ for each j . When $z = e_B$ we denote $\pi_{t_i}^{(z)}$ by π_{t_i} .

Proposition 2.2.6 (Matthews [Mat66]). *Fix a family $\{t_i \mid i \in I\}$ of right-coset representatives for $\langle b \rangle$ in B . Two elements (f, b) and (g, c) are conjugate in $A \wr B$ if and only if there exists an element z in B such that $bz = zc$ and for all $i \in I$ either*

- $\pi_{t_i}^{(z)}(g) = \pi_{t_i}(f)$ if b is of infinite order; or
- $\pi_{t_i}^{(z)}(g)$ is conjugate to $\pi_{t_i}(f)$ if b is of finite order.

For such z in B , a corresponding function h such that $(f, b)(h, z) = (h, z)(g, c)$ is defined as follows: if b is of infinite order then for each $i \in I$ and each $k \in \mathbb{Z}$ we set

$$h(b^k t_i) = \left(\prod_{j \leq k} f(b^j t_i) \right) \left(\prod_{j \leq k} g(z^{-1} b^j t_i) \right)^{-1}$$

or if b is of finite order N , then for each $i \in I$ and each $k = 0, \dots, N - 1$ we set

$$h(b^k t_i) = \left(\prod_{j=0}^k f(b^j t_i) \right) \alpha_{t_i} \left(\prod_{j=0}^k g(z^{-1} b^j t_i) \right)^{-1}$$

where α_{t_i} is any element satisfying $\pi_{t_i}(f)\alpha_{t_i} = \alpha_{t_i}\pi_{t_i}^{(z)}(g)$.

2.2.2.a Upper bounds for lengths of short conjugators

Proposition 2.2.6 gives us an explicit description of a particular conjugator for two elements in $A \wr B$. The following Lemma tells us that any conjugator between two elements has a concrete description similar to that given by Matthews in the preceding Proposition. With this description at our disposal we will be able to determine their size and thus find a short conjugator.

Lemma 2.2.7. *Let $(h, z), (f, b), (g, c) \in A \wr B$ be such that $(f, b)(h, z) = (h, z)(g, c)$. Then there is a set of right-coset representatives $\{t_i \mid i \in I\}$ of $\langle b \rangle$ in B such that, if b is of infinite order then*

$$h(b^k t_i) = \left(\prod_{j \leq k} f(b^j t_i) \right) \left(\prod_{j \leq k} g(z^{-1} b^j t_i) \right)^{-1}$$

for every $i \in I$ and $k \in \mathbb{Z}$; if b is of finite order N then

$$h(b^k t_i) = \left(\prod_{j=0}^k f(b^j t_i) \right) \alpha_{t_i} \left(\prod_{j=0}^k g(z^{-1} b^j t_i) \right)^{-1}$$

for every $i \in I$ and $k = 0, \dots, N - 1$ and where α_{t_i} satisfies $\pi_{t_i}(f)\alpha_{t_i} = \alpha_{t_i}\pi_{t_i}^{(z)}(g)$. Furthermore, for any element α_{t_i} satisfying this relationship there exists some conjugator (h, z) with h of the above form.

Proof. Fix a set of coset representatives $\{s_i \mid i \in I\}$. By Matthews' argument there exists a conjugator $(h_1, z_1) \in A \wr B$ for (f, b) and (g, c) as described in Proposition 2.2.6, with respect to the coset representatives $\{s_i \mid i \in I\}$. Since (h, z) and (h_1, z_1) are both conjugators, it follows that there exists some (ψ, y) in $Z_\Gamma(f, b)$ such that $(h, z) = (\psi, y)(h_1, z_1)$. This tells us that $z = yz_1$ and also that $h(x) = \psi(x)h_1(y^{-1}x)$ for each $x \in B$. Since (ψ, y) is in the centraliser of (f, b) , we obtain two identities:

$$yb = by \tag{2.6}$$

$$\psi(x)f(y^{-1}x) = f(x)\psi(b^{-1}x) \quad \forall x \in B. \tag{2.7}$$

For each $i \in I$ we set $t_i = ys_i$. First suppose that b is of infinite order. Then

$$\begin{aligned}
h(b^k t_i) &= \psi(b^k t_i) h_1(y^{-1} b^k t_i) \\
&= \psi(b^k t_i) h_1(b^k s_i) \\
&= \psi(b^k t_i) \left(\prod_{j \leq k} f(b^j s_i) \right) \left(\prod_{j \leq k} g(z_1^{-1} b^j s_i) \right)^{-1} \\
&= \psi(b^k t_i) \left(\prod_{j \leq k} f(y^{-1} b^j t_i) \right) \left(\prod_{j \leq k} g(z^{-1} b^j t_i) \right)^{-1}.
\end{aligned}$$

We can apply equation (2.7) once, and then repeat this process to shuffle the ψ term past all the terms involving f . This process terminates and the ψ term vanishes because of the finiteness of support of both ψ and of f . Hence, as required, we obtain:

$$\begin{aligned}
h(b^k t_i) &= f(b^k t_i) \psi(b^{k-1} t_i) \left(\prod_{j \leq k-1} f(y^{-1} b^j t_i) \right) \left(\prod_{j \leq k} g(z^{-1} b^j t_i) \right)^{-1} \\
&\vdots \\
&= \left(\prod_{j \leq k} f(b^j t_i) \right) \left(\prod_{j \leq k} g(z^{-1} b^j t_i) \right)^{-1}.
\end{aligned}$$

If instead b is of finite order, N say, then for $0 \leq k \leq N-1$ we obtain

$$h(b^k t_i) = \psi(b^k t_i) \left(\prod_{j=0}^k f(y^{-1} b^j t_i) \right) \alpha_{s_i} \left(\prod_{j=0}^k g(z^{-1} b^j t_i) \right)^{-1}$$

where α_{s_i} is some element satisfying $\pi_{s_i}(f)\alpha_{s_i} = \alpha_{s_i}\pi_{s_i}^{(z_1)}(g)$. With equation (2.7) the $\psi(b^k t_i)$ term can be shuffled past the terms involving f . Unlike in the infinite order case, however, the ψ term will not vanish:

$$h(b^k t_i) = \left(\prod_{j=0}^k f(b^j t_i) \right) \psi(b^{-1} t_i) \alpha_{s_i} \left(\prod_{j=0}^k g(z^{-1} b^j t_i) \right)^{-1}.$$

To confirm that h is of the required form, all that is left to do is to verify that if we set $\alpha_{t_i} = \psi(b^{-1} t_i) \alpha_{s_i}$ then it will satisfy $\pi_{t_i}(f)\alpha_{t_i} = \alpha_{t_i}\pi_{t_i}^{(z)}(g)$. We will prove this while proving the final statement of the Lemma: that any element α_{t_i} satisfying $\pi_{t_i}(f)\alpha_{t_i} = \alpha_{t_i}\pi_{t_i}^{(z)}(g)$ will appear in this expression for some conjugator between (f, b) and (g, c) . Set

$$\mathcal{C}_{t_i} = \{\alpha \mid \pi_{t_i}(f)\alpha = \alpha\pi_{t_i}^{(z)}(g)\} \quad \text{and} \quad \mathcal{C}_{s_i} = \{\alpha \mid \pi_{s_i}(f)\alpha = \alpha\pi_{s_i}^{(z_1)}(g)\}.$$

By Proposition 2.2.6, we can choose h_1 above so that any element of \mathcal{C}_{s_i} appears above in the place of α_{s_i} . We need to check that $\mathcal{C}_{t_i} = \psi(b^{-1}t_i)\mathcal{C}_{s_i}$. Observe that we have two equalities:

$$\psi(b^{-1}t_i)\pi_{s_i}(f) = \pi_{t_i}(f)\psi(b^{-1}t_i) \quad (2.8)$$

$$\pi_{t_i}^{(z)}(g) = \pi_{s_i}^{(z_1)}(g) \quad (2.9)$$

Equation (2.9) is straight-forward to show and was used above in the infinite order argument, while equation (2.8) follows by applying equation (2.7) N times:

$$\psi(b^{-1}t_i) \prod_{j=0}^{N-1} f(b^j s_i) = \prod_{j=0}^{N-1} f(y^{-1}b^j s_i)\psi(b^{-(N+1)}t_i)$$

and then using the facts that b has order N and y is in the centraliser of b .

Suppose that $\alpha_{s_i} \in \mathcal{C}_{s_i}$. Then, using equations (2.8) and (2.9):

$$\begin{aligned} e_A &= \alpha_{s_i}^{-1}\pi_{s_i}(f)^{-1}\alpha_{s_i}\pi_{s_i}^{(z_1)}(g) \\ &= \alpha_{s_i}^{-1}\psi(b^{-1}t_i)^{-1}\pi_{t_i}(f)^{-1}\psi(b^{-1}t_i)\alpha_{s_i}\pi_{t_i}^{(z)}(g). \end{aligned}$$

This confirms that $\psi(b^{-1}t_i)\mathcal{C}_{s_i} \subseteq \mathcal{C}_{t_i}$. On the other hand, suppose instead that $\alpha_{t_i} \in \mathcal{C}_{t_i}$. Then

$$\begin{aligned} e_A &= \alpha_{t_i}^{-1}\pi_{t_i}(f)^{-1}\alpha_{t_i}\pi_{t_i}^{(z)}(g) \\ &= \alpha_{t_i}^{-1}\psi(b^{-1}t_i)\pi_{s_i}(f)^{-1}\psi(b^{-1}t_i)^{-1}\alpha_{t_i}\pi_{s_i}^{(z_1)}(g). \end{aligned}$$

Hence $\psi(b^{-1}t_i)^{-1}\alpha_{t_i} \in \mathcal{C}_{s_i}$. In particular we get $\mathcal{C}_{t_i} = \psi(t^{-1})\mathcal{C}_{s_i}$ as required. \square

Obtaining a short conjugator will require two steps. Lemma 2.2.8 is the first of these steps. Here we actually find the short conjugator, while in Lemma 2.2.9 we show that the size of a conjugator (h, z) can be bounded by a function involving the size of z but independent of h altogether.

Recall that the conjugacy length function of B is the minimal function

$$\text{CLF}_B : \mathbb{N} \rightarrow \mathbb{N}$$

such that if b is conjugate to c in B and $d_B(e_B, b) + d_B(e_B, c) \leq n$ then there exists a conjugator $z \in B$ such that $d_B(e_B, z) \leq \text{CLF}_B(n)$.

Lemma 2.2.8. *Suppose $u = (f, b), v = (g, c)$ are conjugate elements in $\Gamma = A \wr B$ and let $n = d_\Gamma(1, u) + d_\Gamma(1, v)$. Then there exists $\gamma = (h, z) \in \Gamma$ such that $u\gamma = \gamma v$ and either:*

- (1) $d_B(e_B, z) \leq \text{CLF}_B(n)$ if (f, b) is conjugate to $(1, b)$; or
- (2) $d_B(e_B, z) \leq n$ if (f, b) is not conjugate to $(1, b)$.

Proof. Without loss of generality we may assume that $d_\Gamma(1, u) \leq d_\Gamma(1, v)$. By Lemma 2.2.7, if (h_0, z_0) is a conjugator for u and v then there exists a family of right-coset representatives $\{t_i \mid i \in I\}$ for $\langle b \rangle$ in B such that

$$\pi_{t_i}^{(z_0)}(g) = \pi_{t_i}(f) \quad \text{or} \quad \pi_{t_i}^{(z_0)}(g) \text{ is conjugate to } \pi_{t_i}(f)$$

for every $i \in I$ according to whether b is of infinite or finite order respectively (the former follows from the finiteness of the support of the function h given by Lemma 2.2.7).

By Proposition 2.2.6, (f, b) is conjugate to $(1, b)$ if and only if $\pi_{t_i}(f) = e_A$ for every $i \in I$. In this case we take

$$h(b^k t_i) = \begin{cases} \prod_{j \leq k} f(b^j t_i) & \text{if } b \text{ is of infinite order;} \\ \prod_{j=0}^k f(b^j t_i) & \text{if } b \text{ is of finite order } N \text{ and } 0 \leq k < N. \end{cases}$$

One can then verify that $(f, b)(h, e_B) = (h, e_B)(1, b)$. Thus we have reduced (1) to the case when $u = (1, b)$ and $v = (1, c)$. For this we observe that any conjugator z for b, c in B will give a conjugator $(1, z)$ for u, v in $A \wr B$. Thus (1) follows.

If on the other hand (f, b) is not conjugate to $(1, b)$ then by Proposition 2.2.6, $\pi_{t_i}(f) \neq e_A$ for some $i \in I$. Fix some such i , observe that there exists $k \in \mathbb{Z}$ satisfying $b^k t_i \in \text{Supp}(f)$ and there must also exist some $j \in \mathbb{Z}$ so that $z_0^{-1} b^j t_i \in \text{Supp}(g)$. Premultiply (h_0, z_0) by $(f, b)^{k-j}$ to get $\gamma = (h, z)$, where $z = b^{k-j} z_0$ and γ is a conjugator for u and v since $(f, b)^{k-j}$ belongs to the centraliser of u in Γ . By construction, $z^{-1} b^k t_i = z_0^{-1} b^j t_i$ and hence is contained in the support of g . We finish by applying the triangle inequality and using the left-invariance of the word metric d_B as follows:

$$\begin{aligned} d_B(e_B, z^{-1}) &\leq d_B(e_B, z^{-1} b^k t_i) + d_B(z^{-1} b^k t_i, z^{-1}) \\ &\leq d_B(e_B, b^k t_i) + d_B(e_B, z^{-1} b^k t_i) \\ &\leq K(\text{Supp}(f), b) + K(\text{Supp}(g), c) \\ &\leq d_\Gamma(1, u) + d_\Gamma(1, v). \end{aligned}$$

This completes the proof. □

Soon we will give Theorem 2.2.10, which will describe the length of short conjugators in wreath products $A \wr B$ where B is torsion-free. Before we dive into this however, it will prove useful in Section 2.2.4, when we look at conjugacy in free solvable groups, to understand how the conjugators are constructed. In particular, it is important to understand that the size of a conjugator $(h, z) \in A \wr B$ can be expressed in terms of the size of z in B with no need to refer to the function h at all. This is what we explain in Lemma 2.2.9.

For $b \in B$, let $\delta_{\langle b \rangle}^B(n) = \max\{m \in \mathbb{Z} \mid d_B(e_B, b^m) \leq n\}$ be the subgroup distortion of $\langle b \rangle$ in B . Fix a finite generating set X for B and let $\text{Cay}(B, X)$ be the corresponding Cayley graph.

Lemma 2.2.9. *Suppose $u = (f, b), v = (g, c)$ are conjugate elements in $\Gamma = A \wr B$ and let $n = d_\Gamma(1, u) + d_\Gamma(1, v)$. Suppose also that b and c are of infinite order in B . If $\gamma = (h, z)$ is a conjugator for u and v in Γ then*

$$d_\Gamma(1, \gamma) \leq (n + 1)P(2\delta_{\langle b \rangle}^B(P) + 1)$$

where $P = d_B(1, z) + n$.

Proof. Without loss of generality we may assume $d_\Gamma(1, u) \leq d_\Gamma(1, v)$. From Lemma 2.2.7 we have an explicit expression for h . We use this expression to give an upper bound for the size of (h, z) , making use of Lemma 1.4.1 which tells us

$$d_\Gamma(1, \gamma) = K(\text{Supp}(h), z) + |h|$$

where $K(\text{Supp}(h), z)$ is the length of the shortest path in $\text{Cay}(B, X)$ from e_B to z travelling via every point in $\text{Supp}(h)$ and $|h|$ is the sum of terms $d_A(e_A, f(x))$ over all $x \in B$.

We begin by obtaining an upper bound on the size of $K(\text{Supp}(h), z)$. To do this we build a path from e_B to z , zig-zagging along cosets of $\langle b \rangle$, see Figure 2.6. Lemma 2.2.7 tells us that there is a family of right-coset representatives $\{t_i\}_{i \in I}$ such that

$$h(b^k t_i) = \left(\prod_{j \leq k} f(b^j t_i) \right) \left(\prod_{j \leq k} g(z^{-1} b^j t_i) \right)^{-1}$$

for every $i \in I$ and $k \in \mathbb{Z}$. This expression for h tells us where in each coset the support of h will lie. In particular, note that if we set $C = \text{Supp}(f) \cup z\text{Supp}(g)$, then

$$\text{Supp}(h) \cap \langle b \rangle t_i \neq \emptyset \implies C \cap \langle b \rangle t_i \neq \emptyset.$$

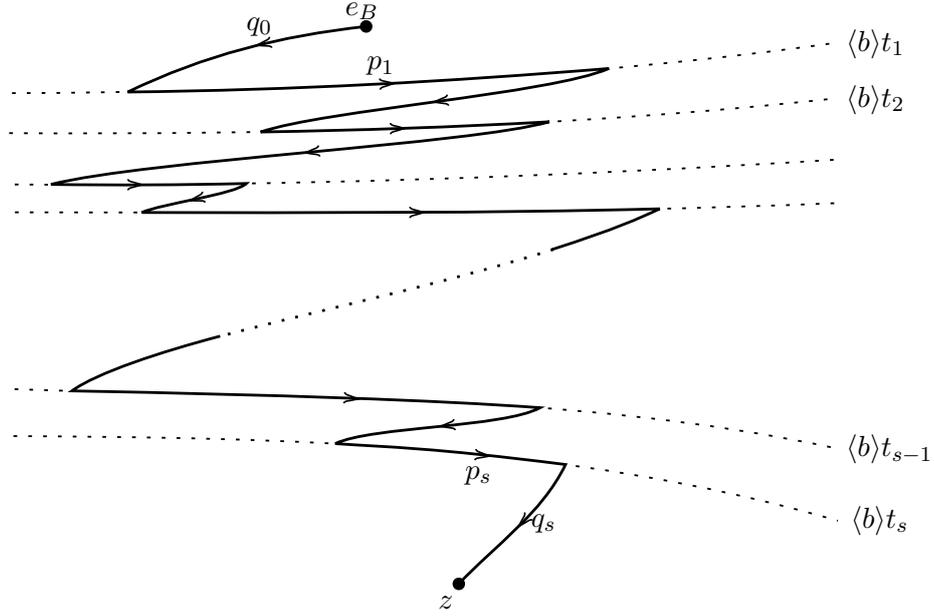


Figure 2.6: We build a path from e_B to b by piecing together paths q_i and p_i , where the paths p_i run through the intersection of $\text{Supp}(h)$ with a coset $\langle b \rangle t_i$ and the paths q_i connect these cosets.

Furthermore, in each coset the support of h must lie between some pair of elements in C . Let t_1, \dots, t_s be all the coset representatives for which $\text{Supp}(h)$ intersects the coset $\langle b \rangle t_i$. The number s of such cosets is bounded above by the size of the set $\text{Supp}(f) \cup \text{Supp}(g)$, which is bounded above by $d_\Gamma(1, u) + d_\Gamma(1, v) = n$.

If we restrict our attention to one of these cosets, $\langle b \rangle t_i$, then there exist integers $m_1 < m_2$ such that $b^j t_i \in \text{Supp}(h)$ implies $m_1 \leq j \leq m_2$. We can choose m_1 and m_2 so that $b^m t_i \in C$ for $m \in \{m_1, m_2\}$. Let p_i be a piecewise geodesic in the Cayley graph of B which connects $b^{m_1} t_i$ to $b^{m_2} t_i$ via $b^j t_i$ for every $m_1 < j < m_2$. The length of p_i will be at most

$$d_B(b^j t_i, b^{j+1} t_i) \delta_{\langle b \rangle}^B(\text{diam}(C))$$

for any $j \in \mathbb{Z}$. Choose $j \in \mathbb{Z}$ such that $b^j t_i \in C$. In that case that $b^j t_i \in \text{Supp}(f)$ we get that

$$\begin{aligned} d_B(b^j t_i, b^{j+1} t_i) &\leq d_B(b^j t_i, b) + d_B(b, b^{j+1} t_i) \\ &= d_B(b^j t_i, b) + d_B(e_B, b^j t_i) \\ &\leq K(\text{Supp}(f), b) \leq n \end{aligned}$$

where the last line follows because any path from e_B to b via all points in $\text{Supp}(f)$ will have to be at least as long as the path from e_B to b via the point $b^j t_i$. Similarly,

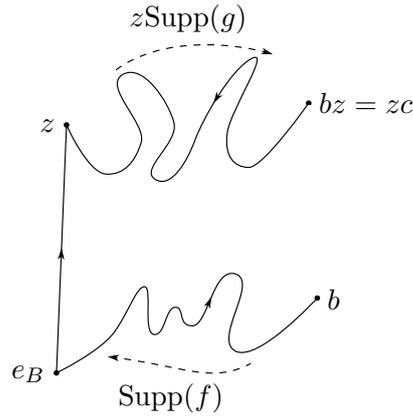


Figure 2.7: The concatenation of three paths: a path from b to e_B through $\text{Supp}(f)$, followed by a path from e_B to z , then finish off with a path from z to zc through each point in $z\text{Supp}(g)$.

in the case when $z^{-1}b^j t_i \in \text{Supp}(g)$, we get

$$\begin{aligned} d_B(b^j t_i, b^{j+1} t_i) &\leq d_B(b^j t_i, zc) + d_B(zc, b^{j+1} t_i) \\ &= d_B(b^j t_i, zc) + d_B(z, b^j t_i) \\ &\leq K(\text{Supp}(g), c) \leq n \end{aligned}$$

where we obtain the last line because a shortest path from z to zc via $z\text{Supp}(g)$ will have length precisely $K(\text{Supp}(g), c)$. Hence, in either case we get that the path p_i has length bounded above by $n\delta_{(b)}^B(\text{diam}(C))$.

We will now show that $\text{diam}(C \cup \{e_B, z\}) \leq n + d_B(1, z) = P$. This diameter will be given by the length of a path connecting some pair of points in this set. We take a path through e_B , z and all points in the set C , a path such as that in Figure 2.7. The length of this path will certainly be bigger than the diameter. Hence we have

$$\begin{aligned} \text{diam}(C) &\leq K(\text{Supp}(f), b) + d_B(1, z) + K(\text{Supp}(g), c) \\ &\leq n + d_B(1, z) = P. \end{aligned}$$

For $i = 1, \dots, s-1$ let q_i be a geodesic path which connects the end of p_i with the start of p_{i+1} . Let q_0 connect e_B with the start of p_1 and q_s connect the end of p_s with z . Then the concatenation of paths $q_0, p_1, q_1, \dots, q_{s-1}, p_s, q_s$ is a path from e_B to z via every point in $\text{Supp}(h)$.

For each i , the path q_i will be a geodesic connecting two points of $C \cup \{e_B, z\}$. The above upper bound for the diameter of this set tells us that each q_i will have length at most $n + d_B(1, z) = P$.

Hence our path $q_0, p_1, q_1, \dots, q_{s-1}, p_s, q_s$ has length bounded above by

$$(n+1)P + n^2\delta_{\langle b \rangle}^B(P) \leq (n+1)P(\delta_{\langle b \rangle}^B(P) + 1)$$

thus giving an upper bound for $K(\text{Supp}(h), z)$.

Now we need to turn our attention to an upper bound for $|h|$. By the value of $h(b^k t_i)$ given to us by Lemma 2.2.7 we see that

$$\begin{aligned} d_A(e_A, h(b^k t_i)) &\leq \sum_{j \leq k} d_A(e_A, g(z^{-1} b^j t_i)) + \sum_{j \leq k} d_A(e_A, f(b^j t_i)) \\ &\leq |g| + |f| \leq n \end{aligned}$$

The number of elements $b^k t_i$ in the support of h can be counted in the following way. Firstly, the number of $i \in I$ for which $\langle b \rangle t_i \cap \text{Supp}(h) \neq \emptyset$ is equal to s , which we showed above to be bounded by n . Secondly, for each such i , recall that there exists $m_1 \leq m_2$ such that $b^j t_i \in \text{Supp}(h)$ implies $m_1 \leq j \leq m_2$. Hence for each i the number of $k \in \mathbb{Z}$ for which $b^k t_i \in \text{Supp}(h)$ is bounded above by $m_2 - m_1 \leq \delta_{\langle b \rangle}^B(P)$. So in conclusion we have

$$\begin{aligned} d_\Gamma(1, \gamma) &= K(\text{Supp}(h), z) + |h| \\ &\leq (n+1)P(\delta_{\langle b \rangle}^B(P) + 1) + n^2\delta_{\langle b \rangle}^B(P) \\ &\leq (n+1)P(2\delta_{\langle b \rangle}^B(P) + 1) \end{aligned}$$

where $n = d_\Gamma(1, u) + d_\Gamma(1, v)$ and $P = d_B(e_B, z) + n$. □

Theorem 2.2.10. *Suppose A and B are finitely generated and B is also torsion-free. Let $u = (f, b), v = (g, c) \in \Gamma = A \wr B$, with $b, c \neq e_B$, and set $n = d_\Gamma(1, u) + d_\Gamma(1, v)$. Then u, v are conjugate if and only if there exists a conjugator $\gamma \in \Gamma$ such that*

$$d_\Gamma(1, \gamma) \leq (n+1)P(2\delta_{\langle b \rangle}^B(P) + 1)$$

where $P = 2n$ if (f, b) is not conjugate to $(1, b)$ and $P = n + \text{CLF}_B(n)$ otherwise.

Proof. By Lemma 2.2.8 we can find a conjugator $\gamma = (h, z)$ which satisfies the inequality $d_B(e_B, z) \leq \text{CLF}_B(n)$ if (f, b) is conjugate to $(1, b)$ or $d_B(e_B, z) \leq n$ otherwise. Therefore if we set $P = n + \text{CLF}_B(n)$ if (f, b) is conjugate to $(1, b)$ and $P = 2n$ otherwise then the result follows immediately by applying Lemma 2.2.9 to the conjugator γ obtained from Lemma 2.2.8. □

When we look at elements whose B -components are non-trivial, or in general when they may be of finite order, we can still obtain some information on the conjugator length. Theorem 2.2.10 does the work when we look at elements in $A \wr B$ such that the B -components are of infinite order. However, if they have finite order we need to understand the size of the conjugators α_i as in Proposition 2.2.6 and Lemma 2.2.7. When the order of b is finite, the construction of the function h by Matthews in Proposition 2.2.6 will work for any conjugator α_{t_i} between $\pi_{t_i}^{(z)}(g)$ and $\pi_{t_i}(f)$. Then, since

$$|\pi_{t_i}^{(z)}(g)| + |\pi_{t_i}(f)| \leq |g| + |f| \leq n$$

where $n = d_\Gamma(1, u) + d_\Gamma(1, v)$, for each coset representative t_i and each $b^k \in \langle b \rangle$ we have

$$d_B(e_B, h(b^k t_i)) \leq |f| + |g| + \text{CLF}_A(n) \leq n + \text{CLF}_A(n).$$

With the aid of the conjugacy length function for A we can therefore give the following:

Lemma 2.2.11. *Suppose $u = (f, b), v = (g, c)$ are conjugate elements in $\Gamma = A \wr B$ and let $n = d_\Gamma(1, u) + d_\Gamma(1, v)$. Suppose also that b and c are of finite order N . If $\gamma = (h, z)$ is a conjugator for u and v in Γ then*

$$d_\Gamma(1, \gamma) \leq P(N + 1)(2n + \text{CLF}_A(n) + 1)$$

where $P = d_B(e_B, z) + n$.

Proof. For the most part this proof is the same as for Lemma 2.2.9. It will differ in two places. As mentioned above, we obtain

$$d_B(e_B, h(b^k t_i)) \leq |f| + |g| + \text{CLF}_A(n)$$

for each coset representative t_i and $b^k \in \langle b \rangle$. By a similar process as that in Lemma 2.2.9 we deduce the upper bound

$$|h| \leq nN(n + \text{CLF}_A(n)).$$

The second place where we need to modify the proof is in the calculation of an upper bound for the length of each path p_i . Since b is of finite order, each coset will give a loop in $\text{Cay}(B, X)$. We will let p_i run around this loop, so its length will be bounded above by $Nd_B(t_i, bt_i)$. As before we get $d_B(t_i, bt_i) \leq n$, so in the upper

bound obtained for $K(\text{Supp}(h), z)$ we need only replace the distortion function $\delta_{\langle b \rangle}^B$ by the order N of b in B . Thus

$$K(\text{Supp}(h), z) \leq P(N + 1)(n + 1)$$

where $P = d_B(e_B, z) + n$. Combining this with the upper bound above for $|h|$ we get

$$\begin{aligned} d_\Gamma(1, \gamma) &\leq P(N + 1)(n + 1) + nN(n + \text{CLF}_A(n)) \\ &\leq P(N + 1)(2n + \text{CLF}_A(n) + 1) \end{aligned}$$

proving the Lemma. □

We finish this section by applying Lemma 2.2.8 and Lemma 2.2.11 to give the complete picture for the length of short conjugators in the case when B may contain torsion.

Theorem 2.2.12. *Suppose A and B are finitely generated groups. Let $u = (f, b), v = (g, c) \in \Gamma$ where the order of b and c is $N \in \mathbb{N} \cup \{\infty\}$. Then u, v are conjugate if and only if there exists a conjugator $\gamma \in \Gamma$ such that either*

$$\begin{aligned} d_\Gamma(1, \gamma) &\leq P(N + 1)(2n + \text{CLF}_A(n) + 1) && \text{if } N \text{ is finite; or} \\ d_\Gamma(1, \gamma) &\leq (n + 1)P(2\delta_{\langle b \rangle}^B(P) + 1) && \text{if } N = \infty, \end{aligned}$$

where $n = d_\Gamma(1, u) + d_\Gamma(1, v)$ and $P = 2n$ if (f, b) is conjugate to $(1, b)$ or $P = n + \text{CLF}_B(n)$ otherwise.

2.2.2.b Lower bounds for lengths of short conjugators

We saw in Section 2.2.2.a that the distortion of cyclic subgroups plays an important role in the upper bound we obtained for the conjugacy length function. We will make use of the distortion to determine lower bounds as well. Firstly, however, we give a straightforward lower bound. In the following, let $|\cdot|_B$ denote word length in the finitely generated group B and $|\cdot|$ without the subscript denote word length in $A \wr B$.

Proposition 2.2.13. *Let A and B be finitely generated groups. Then*

$$\text{CLF}_{A \wr B}(n) \geq \text{CLF}_B(n).$$

Proof. Let $n \in \mathbb{N}$. The value $\text{CLF}_B(n)$ is defined to be the smallest integer such that whenever b, c are conjugate elements in B and satisfy $|b|_B + |c|_B \leq n$ then there is a conjugator $z \in B$ such that $|z|_B \leq \text{CLF}_B(n)$. Let b_n, c_n be elements which realise this minimum. That is:

1. $|b_n|_B + |c_n|_B \leq n$; and
2. a minimal length conjugator $z_n \in B$ satisfies $|z_n|_B = \text{CLF}_B(n)$.

Consider the elements $u_n = (1, b_n), v_n = (1, c_n)$ in $A \wr B$, where 1 represents the trivial function. Then by Lemma 1.4.1

$$|u_n| + |v_n| \leq n.$$

Any conjugator (h, x) must satisfy $h^{b_n} = h$ and $b_n x = x c_n$. We may take $h = 1$ since any non-trivial function h (only possible when b_n is of finite order) will lead to a larger conjugator. Thus a minimal length conjugator for u_n and v_n will have the form $(1, x)$ where x can be chosen to be any conjugator for b_n and c_n . In particular, this shows that the minimal length conjugator for u_n and v_n has length $\text{CLF}_B(n)$. \square

Distorted elements. Let B be a finitely generated group containing an element x of infinite order. If the centraliser of x in B , denoted $Z_B(x)$, is sufficiently large (see Lemma 2.2.15), then we can use the distortion of $\langle x \rangle$ in B to construct two sequences of functions from B to A that allow us to demonstrate a lower bound on the conjugacy length function of $A \wr B$ in terms of this distortion. Given any element b of infinite order in B , by taking $x = b^3$ we ensure that x has sufficiently large centraliser in order to apply Lemma 2.2.15. Since the distortion of $\langle b \rangle$ in B is (roughly) a third of that of $\langle x \rangle$, we can conclude that the distortion function of any cyclic subgroup in B provides a lower bound for the conjugacy length function of $A \wr B$.

Theorem 2.2.14. *Let A and B be finitely generated groups and let $b \in B$ be any element of infinite order. Then*

$$\text{CLF}_{A \wr B}(4n + 4 + 10|b|_B) \geq \frac{4}{3} \delta_{\langle b \rangle}^B(n) - 4.$$

In order to prove Theorem 2.2.14 we will use the following closely-related Lemma:

Lemma 2.2.15. *If the set $\{y \in Z_B(x) : y^2 \notin \langle x \rangle\}$ is non-empty then*

$$\text{CLF}_{AB}(4(n + L_x + 1) + 2|x|_B) \geq 4\delta_{\langle x \rangle}^B(n)$$

where $L_x = \min\{|y|_B : y \in Z_B(x), y^2 \notin \langle x \rangle\}$.

Proof of Theorem 2.2.14. To obtain the Theorem we need to apply Lemma 2.2.15, taking $x = b^3$. Then $b \in Z_B(x)$ and $b^2 \notin \langle x \rangle$, hence $L_x \leq |b|$. The distortion function for b satisfies $\delta_{\langle b \rangle}^B(n) \geq \frac{1}{3}\delta_{\langle x \rangle}^B(n) - 1$. Hence the Theorem follows by application of Lemma 2.2.15. \square

Proof of Lemma 2.2.15. Take an element y in $Z_B(x)$ which realises this minimum. Let a be any element in the chosen generating set of A and consider two functions $f_n, g_n : B \rightarrow A$ which take values of either e_A or a and which have the following supports:

$$\begin{aligned} \text{Supp}(f_n) &= \{e_B, y\} \\ \text{Supp}(g_n) &= \{x^{-\delta(n)}, x^{\delta(n)}y\} \end{aligned}$$

where $\delta(n) = \delta_{\langle x \rangle}^B(n)$. We use these functions to define a pair of conjugate elements: $u_n = (f_n, x), v_n = (g_n, x)$. First we will show that the sum of the sizes of these elements grows with n . Observe that

$$|u_n| + |v_n| = K(\text{Supp}(f_n), x) + 2 + K(\text{Supp}(g_n), x) + 2$$

where the notation is as in Lemma 1.4.1. For every n ,

$$\begin{aligned} K(\text{Supp}(f_n), x) &\leq 2|y|_B + |x|_B \\ n \leq K(\text{Supp}(g_n), x) &\leq 2|y|_B + 4n + |x|_B. \end{aligned}$$

Hence $n \leq |u_n| + |v_n| \leq 4n + 4L_x + 2|x|_B + 4$. As an example of a conjugator we may take $\gamma_n = (h_n, e_B)$, where h_n is given by

$$\begin{aligned} h_n(x^i y) &= a \quad \text{if } 0 \leq i \leq \delta(n) - 1; \\ h_n(x^{-i}) &= a^{-1} \quad \text{if } 1 \leq i \leq \delta(n); \\ h_n(b) &= e_A \quad \text{otherwise.} \end{aligned}$$

We will now verify that this is indeed a conjugator. To do so, we need to verify that $f_n h_n^x = h_n g_n$. We need only check this holds for elements of the form $x^i y$ or x^{-i} for

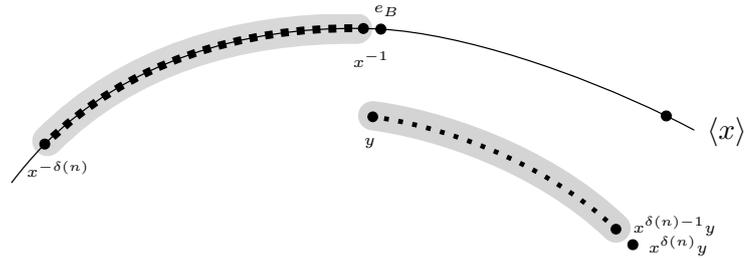


Figure 2.8: The support of h_n is the union of the two shaded regions.

$0 \leq i \leq \delta(n)$ since otherwise both sides evaluate to the identity. The reader can verify that

$$f_n(x^i y) h_n(x^{i-1} y) = a = h_n(x^i y) g_n(x^i y)$$

whenever $0 \leq i \leq \delta(n)$. Provided $1 \leq i \leq \delta(n) - 1$ we get

$$f_n(x^{-i}) h_n(x^{-i-1}) = a^{-1} = h_n(x^{-i}) g_n(x^{-i})$$

and in the last two cases, that is for $i \in \{0, \delta(n)\}$, both sides equal the identity.

Now we will show that any conjugator for u_n and v_n will have to have size bounded below by $4\delta(n)$. This is done by showing the support of the associated function will contain at least $2\delta(n)$ elements. We will first give the family of elements $z \in B$ for which there exists a conjugator for u_n and v_n of the form (h, z) for some function h . Suppose z is some such element. Lemma 2.2.7 then tells us what the corresponding function h will look like. In particular, in order for the support of h to be finite, we must have that

$$\text{Supp}(f_n) \cap \langle x \rangle t \neq \emptyset \quad \text{if and only if} \quad z \text{Supp}(g_n) \cap \langle x \rangle t \neq \emptyset$$

for any $t \in B$ (note that this does not apply in general, but it does here because the functions have been designed so their supports intersect each coset with at most one element). Hence $z \text{Supp}(g_n)$ will intersect the cosets $\langle x \rangle y^i$ exactly once for each $i = 0, 1$ and it will not intersect any other coset. Let σ be the permutation of $\{0, 1\}$ such that $z x^{\delta(n)} y \in \langle x \rangle y^{\sigma(1)}$ and $z x^{-\delta(n)} \in \langle x \rangle y^{\sigma(0)}$. Since y is in the centraliser of x it follows that $z \in \langle x \rangle y^{\sigma(i)-i}$ for each i . If $\sigma(i) \neq i$ then this implies that $z \in \langle x \rangle y^{-1} \cap \langle x \rangle y$, so $y^2 \in \langle x \rangle$, contradicting our choice of y . Hence $\sigma(i) = i$ for $i \in \{0, 1\}$, and thus $z \in \langle x \rangle$.

The support of g_n was chosen in such a way that it is sufficiently spread out in the two cosets of $\langle x \rangle$. It means that shifting $\text{Supp}(g_n)$ by any power of x does not prevent the support of h needing at least $2\delta(n)$ elements. In particular, if $z = x^k$

for some $-\delta(n) < k < \delta(n)$ then the support of h will consist of elements $x^i y$ for $0 \leq i < \delta(n) + k$ and x^{-i} for $1 \leq i \leq \delta(n) - k$. Here the support has precisely $2\delta(n)$ elements. If k lies outside this range then the support will contain at least as many as $2\delta(n)$ elements, for example if $k \geq \delta(n)$ then the support will consist of elements $x^i y$ for $0 \leq i < \delta(n) + k$ as well as x^i for any i satisfying $0 \leq i < -\delta(n) + k$. This implies that, by Lemma 1.4.1, any conjugator for u_n and v_n will have to have size at least $4\delta(n)$, providing the required lower bound for the conjugacy length function. \square

Osin [Osi01] has described the distortion functions of subgroups of finitely generated nilpotent groups. In particular, for a c -step nilpotent group N his result implies that the maximal distortion of a cyclic subgroup of N will be n^c , and this occurs when the subgroup is contained in the centre of N . A consequence of Theorem 2.2.10, Theorem 2.2.14 and Osin's work is that when restricting to elements in $A \wr N$ not conjugate to an element of the form $(1, b)$, the (restricted) conjugacy length function will be n^α , where $\alpha \in [c, c + 2]$. However, since we do not yet know the conjugacy length function of a general c -step nilpotent group, apart from this lower bound we cannot estimate the conjugacy length function of $A \wr N$. In the particular case when N is 2-step nilpotent we know its conjugacy length function is quadratic by Ji, Ogle and Ramsey [JOR10]. Hence the conjugacy length function of $A \wr N$, when N is torsion-free, is n^α for some $\alpha \in [2, 7]$.

Baumslag–Solitar groups. Using the exponential distortion of elements in the solvable Baumslag–Solitar groups $BS(1, q)$ we are able to show that the conjugacy length function for $A \wr BS(1, q)$ is exponential for any finitely generated group A . We will write $BS(1, q)$ with presentation

$$\langle a, b \mid aba^{-1} = b^q \rangle.$$

The subgroup generated by b is exponentially distorted since $b^{q^n} = a^n b a^{-n}$. Hence by Theorem 2.2.14 the conjugacy length function of $A \wr BS(1, q)$ will be at least exponential. We must show that it also has an exponential upper bound. To do this we will use a normal form which is admitted by the solvable Baumslag–Solitar groups: every element $g \in BS(1, q)$ can be written uniquely in the form $g = a^{-r} b^t a^s$, where $r, s \geq 0$ and if $r, s > 0$ then q does not divide t .

Proposition 2.2.16. *The conjugacy length function for $A \wr BS(1, q)$ is exponential.*

Proof. The exponential lower bound is a consequence of Theorem 2.2.14. The conjugacy length function for $\text{BS}(1, q)$ was shown to be linear in [Sal11], as well as in Theorem 2.3.5 below. In order to deduce from the result for wreath products (Theorem 2.2.10) that we have an exponential upper bound on conjugacy length, we need to show that the distortion of cyclic subgroups generated by elements of $\text{BS}(1, q)$ of length at most n is bounded by a function which is exponential in n .

Let $g = a^{-r}b^t a^s$ be an element of $\text{BS}(1, q)$ given in normal form and contained in the ball of radius n about the identity. We will consider the distortion of the subgroup $\langle g \rangle$. First suppose that $r = s$. Then by writing $g^{q^k} = a^{-r+k} b a^{r-k}$ we get an upper bound on $\delta_{\langle g \rangle}^{\text{BS}(1, q)}(m)$ of $q^{\frac{m-1}{2}+r}$. But $r \leq n$ since g has length at most n . Hence

$$\delta_{\langle g \rangle}^{\text{BS}(1, q)}(m) \leq q^{\frac{m-1}{2}+n}$$

among all elements g with normal form $a^{-r}b^t a^r$ with length at most n .

Now suppose $s \neq r$. Consider the action of $\text{BS}(1, q)$ on its Bass–Serre tree T , which will be a $(q+1)$ -regular tree. Fix a basepoint x of T , let ω be the boundary point of T determined by the ray consisting of vertices $a^k x$ for $k \in \mathbb{N}$ and let \mathfrak{h} be the associated Busemann function, as defined in Section 2.1.1. The element $g = a^{-r}b^t a^s$ will act on T by moving a point in the p -th horocycle of \mathfrak{h} to a point in the $(p+r-s)$ -th horocycle. Hence $|g^k| \geq |k(r-s)|$. This implies that

$$\delta_{\langle g \rangle}^{\text{BS}(1, q)}(m) \leq \frac{m}{n}$$

for any g of the form $a^{-r}b^t a^s$ with $r \neq s$ and of length at most n . \square

Using the area of triangles. We will show that the conjugacy length function of wreath products $A \wr B$, where A is finitely generated and B contains a copy of \mathbb{Z}^2 , are non-linear, and in particular are at least quadratic. Combined with Theorem 2.2.10 we learn, for example, that $\text{CLF}_{A \wr \mathbb{Z}^r}(n) \asymp n^\alpha$ for some $\alpha \in [2, 3]$. The methods used here differ to those used above in that we do not use subgroup distortion. Instead we rely on the area of triangles in \mathbb{Z}^2 being quadratic with respect to the perimeter length.

Suppose x and y generate a copy of \mathbb{Z}^2 in B . For each $n \in \mathbb{N}$ let $f_n, g_n : B \rightarrow A$ be functions which take values of either e_A or a , where a is an element of a generating set for A , and have supports given by

$$\text{Supp}(f_n) = \{x^{-n}, \dots, x^{-1}, e_B, x, \dots, x^n\},$$

$$\text{Supp}(g_n) = \{x^{-n}y^{-n}, \dots, x^{-1}y^{-1}, e_B, xy, \dots, x^n y^n\}.$$

Consider the two elements $u_n = (f_n, y)$ and $v_n = (g_n, y)$. These are conjugate via the element (h_n, e_B) , where h_n is defined by

$$\begin{aligned} h_n(x^i y^j) &= a & \text{if } 0 < i \leq n \text{ and } 0 \leq j < i; \\ h_n(x^i y^j) &= a^{-1} & \text{if } -n \leq i < 0 \text{ and } i \leq j < i; \\ h_n(z) &= e_A & \text{otherwise.} \end{aligned}$$

It is clear that the sizes of u_n and v_n grows linearly with n . In particular, using Lemma 1.4.1, one can verify that

$$4n + 2 \leq |u_n| \leq 2n(2|x| + 1) + |y| + 1,$$

$$4n + 2 \leq |v_n| \leq 2n(2|xy| + 1) + |y| + 1.$$

Suppose that (h, z) is any conjugator for u_n and v_n . First we claim that z must be a power of y . This follows from a similar argument as in the proof of Theorem 2.2.14. By Lemma 2.2.7, given z we can construct the function h . The support of h will be finite only if for any $t \in B$

$$\text{Supp}(f_n) \cap \langle y \rangle t \neq \emptyset \quad \text{if and only if} \quad z \text{Supp}(g_n) \cap \langle y \rangle t \neq \emptyset. \quad (2.10)$$

First observe that $z \in \langle x, y \rangle = \mathbb{Z}^2$ since otherwise we would have $\text{Supp}(f_n) \cap \langle y \rangle z = \emptyset$ while on the other hand $z \in z \text{Supp}(g_n) \cap \langle y \rangle z$. By the nature of \mathbb{Z}^2 we can see that if z acts on it by translation in any direction except those parallel to y then (2.10) cannot hold, implying h will have infinite support. Thus any conjugator for u_n and v_n must be of the form (h, y^k) for some integer k .

A simple geometric argument now gives us a quadratic lower bound on the size of (h, y^k) relative to n . The support of h will be contained in the cosets $\langle y \rangle x^i$ for $-n \leq i \leq n$. Within each coset it will include precisely one of x^i or $x^r y^{i+k}$, as well as all elements in between. Specifically, $x^i y^j \in \text{Supp}(h)$ if and only if either $0 \leq j < i+k$ or $i+k \leq j < 0$. A triangle of elements is therefore contained in the support of h . Regardless of what value of k is chosen this triangle can be chosen to have at least n elements along the base and side, giving a minimum of $\frac{1}{2}n(n+1)$ elements in $\text{Supp}(h)$. Thus any conjugator must have size bounded below by $n(n+1)$.

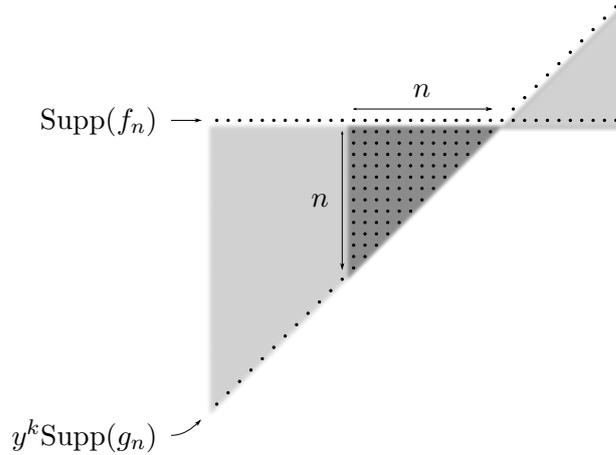


Figure 2.9: The shaded region indicates $\text{Supp}(h)$, while the dark shaded region is a triangle contained in $\text{Supp}(h)$ which contains $\frac{1}{2}n(n + 1)$ elements.

Theorem 2.2.17. *Let A and B be finitely generated groups and suppose that B contains a copy of \mathbb{Z}^2 generated by elements x and y . The conjugacy length function of $A \wr B$ satisfies*

$$\text{CLF}_{A \wr B}(4n(|x| + |xy| + 1) + 2|y| + 2) \geq n^2 + n.$$

When $B = \mathbb{Z}^r$ for some $r \geq 2$ we get the following upper and lower bounds:

$$\frac{(n + 14)(n - 2)}{256} \leq \text{CLF}_{A \wr \mathbb{Z}^r}(n) \leq 7n(n + 1)(14n + 1).$$

Proof. The lower bounds follow from the argument preceding the Proposition. The upper bound for the second expression is an immediate consequence of Theorem 2.2.10, using the facts that $\text{CLF}_{\mathbb{Z}^r}$ is identically zero and the distortion function for cyclic subgroups in \mathbb{Z}^r is the identity function. \square

The use of area in this way to provide a lower bound on the conjugacy length function raises the question of whether, in general, one can use the Dehn function of a group B to provide a lower bound for $\text{CLF}_{A \wr B}$.

QUESTION: Let $\text{Area}_B : \mathbb{N} \rightarrow \mathbb{N}$ be the Dehn function of a finitely presented group B . Is it true that $\text{CLF}_{A \wr B}(n) \succeq \text{Area}_B(n)$ for any finitely generated group A ?

2.2.3 Subgroup distortion

We saw in Theorem 2.2.10 that in order to understand the conjugacy length function of a wreath product $A \wr B$ we need to understand the distortion function for infinite cyclic subgroups in B .

Recall subgroup distortion is studied up to an equivalence relation of functions. For functions $f, g : \mathbb{N} \rightarrow [0, \infty)$ we write $f \preceq g$ if there exists an integer $C > 0$ such that $f(n) \leq Cg(Cn)$ for all $n \in \mathbb{N}$. The two functions are equivalent if both $f \preceq g$ and $g \preceq f$. In this case we write $f \asymp g$.

We will see that all cyclic subgroups of free solvable groups $S_{r,d}$ are undistorted. This is not always the case in finitely generated solvable groups. For example, in the solvable Baumslag-Solitar groups $BS(1, q) = \langle a, b \mid aba^{-1} = b^q \rangle$ the subgroup generated by b is at least exponentially distorted since $b^{q^n} = a^n b a^{-n}$. Because this type of construction doesn't work in $\mathbb{Z} \wr \mathbb{Z}$ or free metabelian groups it leads to a question of whether all subgroups of these groups are undistorted (see [DO11, §2.1]). Davis and Olshanskii answered this question in the negative, giving, for any positive integer t , 2-generated subgroups of these groups with distortion function bounded below by a polynomial of degree t .

The following Lemma is given in [DO11, Lemma 2.3].

Lemma 2.2.18. *Let A, B be finitely generated abelian groups. Then every finitely generated abelian subgroup of $A \wr B$ is undistorted.*

Davis and Olshanskii prove this by showing that such subgroups are retracts of a finite index subgroup of $A \wr B$. A similar process can be applied to finitely generated abelian subgroups of free solvable groups to show that they are undistorted [Ols]. Below we give an alternative proof for cyclic subgroups which provides an effective estimate for the constant¹ and which uses only results given in this paper.

Proposition 2.2.19. *Every cyclic subgroup of a free solvable group is undistorted. In particular, suppose $d \geq 1$ and let x be a non-trivial element of $S_{r,d}$. Then*

$$\delta_{\langle x \rangle}^{S_{r,d}}(n) \leq 2n.$$

Proof. Let w be a non-trivial element of the free group F . There exists an integer c such that $w \in F^{(c)} \setminus F^{(c+1)}$, where we include the case $F^{(0)} = F$. First we suppose that $d = c + 1$.

If $c = 0$ then we have $x \in F/F' = \mathbb{Z}^r$ and we apply linear distortion in \mathbb{Z}^r . If $c > 0$ then we take a Magnus embedding $\varphi : S_{r,d} \hookrightarrow \mathbb{Z}^r \wr S_{r,c}$ and observe that since $w \in F^{(c)} \setminus F^{(d)}$ the image of x in φ is $(f, 1)$ for some non-trivial function $f : S_{r,c} \rightarrow \mathbb{Z}^r$.

¹We thank Olshanskii for improving the constant from 2^d to 2.

If f^k denotes the function such that $f^k(b) = kf(b)$ for $b \in S_{r,c}$, then for any $k \in \mathbb{Z}$, since the Magnus embedding is 2-bi-Lipschitz (Theorem 2.2.5),

$$d_{S_{r,d}}(1, x^k) \geq \frac{1}{2}d_M(1, (f, 1)^k) = \frac{1}{2}d_M(1, (f^k, 1)).$$

We can apply Lemma 1.4.1 to get

$$\frac{1}{2}d_M(1, (f^k, 1)) = \frac{1}{2} \left(K(\text{Supp}(f), 1) + \sum_{b \in S_{r,c}} |kf(b)| \right)$$

and since the image of f lies in \mathbb{Z}^r and f is non-trivial

$$\sum_{b \in S_{r,c}} |kf(b)| = |k| \sum_{b \in S_{r,c}} |f(b)| \geq |k|.$$

Hence

$$d_{S_{r,d}}(1, x^k) \geq \frac{1}{2}d_M(1, (f^k, 1)) \geq \frac{1}{2}|k|.$$

This implies $\delta_{(x)}^{S_{r,d}}(n) \leq 2n$.

Now suppose that $d > c + 1$. Then we define a homomorphism

$$\psi : S_{r,d} \rightarrow S_{r,c+1}$$

by sending the free generators of $S_{r,d}$ to the corresponding free generator of $S_{r,c+1}$. Then, as before, set $x = wF^{(d)}$ and define y to be the image of x under ψ . By the construction of ψ we have that $y = wF^{(c+1)}$. Note that y is non-trivial and ψ does not increase the word length. Hence, using the result above for $S_{r,c+1}$, we observe that

$$d_{S_{r,d}}(1, x^k) \geq d_{S_{r,c+1}}(1, y^k) \geq \frac{1}{2}|k|.$$

This suffices to show that the distortion function is always bounded above by $2n$. \square

2.2.4 Conjugacy in free solvable groups

The fact that the conjugacy problem is solvable in free solvable groups was shown by Kargapolov and Remeslennikov [KR66]. The following Theorem was given by Remeslennikov and Sokolov [RS70]. They use it, alongside Matthews' result for wreath products, to show the decidability of the conjugacy problem in $S_{r,d}$.

Theorem 2.2.20 (Remeslennikov–Sokolov [RS70]). *Suppose F/N is torsion-free and let $u, v \in F/N'$. Then $\varphi(u)$ is conjugate to $\varphi(v)$ in $M(F/N)$ if and only if u is conjugate to v in F/N' .*

The proof of Theorem 2.2.21 more-or-less follows Remeslennikov and Sokolov's proof of the preceding theorem. With it one can better understand the nature of conjugators and how they relate to the Magnus embedding. In particular it tells us that once we have found a conjugator in the wreath product $\mathbb{Z}^r \wr S_{r,d-1}$ for the image of two elements in $S_{r,d}$, we need only modify the function component of the element to make it lie in the image of the Magnus embedding.

Theorem 2.2.21. *Let u, v be two elements in F/N' such that $\varphi(u)$ is conjugate to $\varphi(v)$ in $M(F/N)$. Let $g \in M(F/N)$ be identified with $(f, \gamma) \in \mathbb{Z}^r \wr F/N$. Suppose that $\varphi(u)g = g\varphi(v)$. Then there exists $w \in F/N'$ such that $\varphi(w) = (f_w, \gamma)$ is a conjugator.*

Proof. Let $g \in M(F/N)$ be such that $\varphi(u)g = g\varphi(v)$. Suppose

$$g = \begin{pmatrix} \gamma & a \\ 0 & 1 \end{pmatrix} = (f, \gamma)$$

for $\gamma \in F/N$ and $a \in \mathcal{R}$. Recall that $\bar{\alpha} : F/N' \rightarrow F/N$ is the canonical homomorphism and by Lemma 2.2.2 we may consider the derivations $\frac{\partial^*}{\partial x_i}$ to be maps from $\mathbb{Z}(F/N')$ rather than $\mathbb{Z}(F)$. By direct calculation we obtain the two equations

$$\bar{\alpha}(u)\gamma = \gamma\bar{\alpha}(v) \tag{2.11}$$

$$\sum_{i=1}^r \frac{\partial^* u}{\partial x_i} t_i + \bar{\alpha}(u)a = \gamma \sum_{i=1}^r \frac{\partial^* v}{\partial x_i} t_i + a \tag{2.12}$$

We now split the proof into two cases, depending on whether or not $\bar{\alpha}(u)$ is the identity element.

CASE 1: $\bar{\alpha}(u)$ is trivial.

In this case equation (2.12) reduces to

$$\sum_{i=1}^r \frac{\partial^* u}{\partial x_i} t_i = \gamma \sum_{i=1}^r \frac{\partial^* v}{\partial x_i} t_i$$

and it follows that $\varphi(\gamma_0) = (h, \gamma)$ will be a conjugator for $\varphi(u)$ and $\varphi(v)$, where γ_0 is any lift of γ in F/N' .

CASE 2: $\bar{\alpha}(u)$ is non-trivial.

Note that in this case we actually show a stronger result, that any conjugator for $\varphi(u)$ and $\varphi(v)$ must lie in the subgroup $\varphi(F/N')$. This is clearly not necessarily true in the first case.

First conjugate $\varphi(u)$ by $\varphi(\gamma_0)$, where γ_0 is a lift of γ in F/N' . This gives us two elements which are conjugate by a unipotent matrix in $M(F/N)$, in particular there exist b_1, \dots, b_r in $\mathbb{Z}(F/N)$ such that the conjugator is of the form

$$\varphi(\gamma_0)^{-1}g = \gamma' = \begin{pmatrix} 1 & b_1 t_1 + \dots + b_r t_r \\ 0 & 1 \end{pmatrix}.$$

Hence the aim now is to show that there is some element y in N such that $\gamma' = \varphi(y)$, in particular $\frac{\partial^* y}{\partial x_i} = b_i$ for each i . Therefore, without loss of generality, we assume that $\varphi(u)$ and $\varphi(v)$ are conjugate by such a unipotent matrix.

Assume that $\varphi(u)\gamma' = \gamma'\varphi(v)$. Then equation (2.11) tells us that $\bar{\alpha}(u) = \bar{\alpha}(v)$. Hence $uv^{-1} = z \in N$. Observe that $\frac{\partial^* z}{\partial x_i} = \frac{\partial^* u}{\partial x_i} - \frac{\partial^* v}{\partial x_i}$, hence from equation (2.12) we get

$$(1 - \bar{\alpha}(u))b_i = \frac{\partial^* z}{\partial x_i}$$

for each $i = 1, \dots, r$.

Let c be an element of $\mathbb{Z}(F/N')$ such that $\frac{\partial^* c}{\partial x_i} = b_i$ for each i . We therefore have the following:

$$(1 - \bar{\alpha}(u)) \sum_{i=1}^r \frac{\partial^* c}{\partial x_i} (\alpha(x_i) - 1) = \sum_{i=1}^r \frac{\partial^* z}{\partial x_i} (\alpha(x_i) - 1).$$

We can choose c so that $\varepsilon(c) = 0$, and then apply the fundamental formula of Fox calculus, Lemma 2.2.1, to both sides of this equation to get

$$(1 - \bar{\alpha}(u))c = z - 1$$

since $z \in N$ implies $\varepsilon(z) = 1$. In $\mathbb{Z}(F/N)$ the right-hand side is 0. Furthermore, since F/N is torsion-free, $(1 - \bar{\alpha}(u))$ is not a zero divisor, so c lies in the kernel of the homomorphism $\bar{\alpha}^* : \mathbb{Z}(F/N') \rightarrow \mathbb{Z}(F/N)$. Hence there is an expression of c in the following way (see Lemma 2.2.3):

$$c = \sum_{j=1}^m r_j (h_j - 1)$$

where m is a positive integer and for each $j = 1, \dots, m$ we have $r_j \in F/N'$ and $h_j \in N/N'$. Differentiating this expression therefore gives

$$\begin{aligned} b_i = \frac{\partial^* c}{\partial x_i} &= \sum_{j=1}^m \left(\frac{\partial^* r_j}{\partial x_i} \varepsilon(h_j - 1) + \bar{\alpha}(r_j) \frac{\partial^* (h_j - 1)}{\partial x_i} \right) \\ &= \sum_{j=1}^m \bar{\alpha}(r_j) \frac{\partial^* h_j}{\partial x_i} \end{aligned}$$

We set

$$y = \prod_{j=1}^m r_j h_j r_j^{-1} \in N.$$

Since $h_j \in N/N'$ for each j , we have the following equations:

$$\frac{\partial^*(h_1 h_2)}{\partial x_i} = \frac{\partial^* h_1}{\partial x_i} + \frac{\partial^* h_2}{\partial x_i}, \quad \frac{\partial^*(r_j h_j r_j^{-1})}{\partial x_i} = r_j \frac{\partial^* h_j}{\partial x_i}.$$

Using these, the condition $\frac{\partial^* y}{\partial x_i} = b_i$ can be verified. Hence, $g = \varphi(\gamma_0)\varphi(y)$, so taking $w = \gamma_0 y$ gives us a conjugator $\varphi(\gamma_0 y) = (f_w, \gamma)$ of the required form. \square

Theorem 2.2.21 tells us that when considering two conjugate elements in F/N' we may use the wreath product result, Theorem 2.2.12, and the fact that the Magnus embedding does not distort word lengths, Theorem 2.2.5, to obtain a control on the length of conjugators in F/N' in terms of the conjugacy length function of F/N . Recall that in $A \wr B$ the conjugacy length function of B plays a role only when we consider elements conjugate to something of the form $(1, b)$. Hence, with the use of the following Lemma, we can obtain an upper bound on the conjugacy length function of F/N' which is independent of the conjugacy length function of F/N and depends only on the distortion of cyclic subgroups in F/N .

Lemma 2.2.22. *Let (f, b) be a non-trivial element in the image $\varphi(F/N')$, where b is of infinite order in F/N . Then (f, b) is not conjugate to any element of the form $(1, c)$ in $\mathbb{Z}^r \wr F/N$.*

Proof. First note that we may assume $b \neq e$, since otherwise for (f, e) to be conjugate to $(1, c)$, c would have to be e and $f = 1$. Suppose, for contradiction, that $(1, c) = (h, z)^{-1}(f, b)(h, z)$. Then

$$0 = -h(zg) + f(zg) + h(b^{-1}zg) \tag{2.13}$$

for every $g \in F/N$. We will show that for this to be true, the support of h must be infinite. Write f and h in component form, that is

$$f(g) = (f_1(g), \dots, f_r(g)) \quad \text{and} \quad h(g) = (h_1(g), \dots, h_r(g))$$

for $g \in F/N$ and where $f_i, h_i : F/N \rightarrow \mathbb{Z}$. Recall that the generators of F are $X = \{x_1, \dots, x_r\}$. We will abuse notation, letting this set denote a generating set for F/N as well. Consider the function defined by

$$\Sigma_f(g) = \sum_{i=1}^r f_i(g) - \sum_{i=1}^r f_i(gx_i^{-1})$$

for $g \in F/N$ and similarly Σ_h for h in place of f . Since (f, b) is in the image of the Magnus embedding it gives a path ρ in the Cayley graph of F/N , as discussed in Section 2.2.1.b. We also showed in the proof of Theorem 2.2.5 that $f_i(g)$ counts the net number of times ρ passes along the edge (g, gx_i) , counting $+1$ when it travels from g to gx_i and -1 each time it goes from gx_i to g . Therefore $\Sigma_f(g)$ counts the net number of times this path leaves the vertex labelled g and in particular we deduce that

$$\Sigma_f(e) = 1, \quad \Sigma_f(b) = -1 \quad \text{and} \quad \Sigma_f(g) = 0 \quad \text{for} \quad g \neq e, b.$$

From equation (2.13) we get $\Sigma_f(g) = \Sigma_h(g) - \Sigma_h(b^{-1}g)$ for every $g \in F/N$. Note that if the support of h is to be finite, the support of Σ_h must also be finite.

Consider any coset $\langle b \rangle t$ in F/N . Suppose first that t is not a power of b . Then Σ_f is identically zero on $\langle b \rangle t$. Since $0 = \Sigma_f(b^k t) = \Sigma_h(b^k t) - \Sigma_h(b^{k-1} t)$ implies that Σ_h is constant on $\langle b \rangle t$, if h is to have finite support then this must always be zero since b is of infinite order. Looking now at $\langle b \rangle$, we similarly get $\Sigma_h(b^k) = \Sigma_h(b^{k-1})$ for all $k \neq 0, 1$. Again, the finiteness of the support of h implies that $\Sigma_h(b^k) = 0$ for all $k \neq 0$. However, $1 = \Sigma_f(e) = \Sigma_h(e) - \Sigma_h(b^{-1})$, so $\Sigma_h(e) = 1$. To summarise:

$$\Sigma_h(g) \neq 0 \quad \text{if and only if} \quad g = e.$$

We can use this to construct an infinite path in $\text{Cay}(F/N, X)$, which contains infinitely many points in $\text{Supp}(h)$ and thus obtain a contradiction. Start by setting g_0 to be e . Since $\Sigma_h(e) = 1$ there is some $x_{i_1} \in X$ such that either $h_{i_1}(e) \neq 0$ or $h_{i_1}(x_{i_1}^{-1}) \neq 0$. If the former is true then let $g_1 = x_{i_1}$; in the latter case take $g_1 = x_{i_1}^{-1}$. Since $\Sigma_h(g_1) = 0$ there must be some adjacent vertex g_2 such that either $g_2 = g_1 x_{i_2}$ and $h_{i_2}(g_1) \neq 0$ or $g_2 = g_1 x_{i_2}^{-1}$ and $h_{i_2}(g_2) \neq 0$. We can extend this construction endlessly, building an infinite sequence (g_m) for $m \in \mathbb{Z} \cup \{0\}$. Furthermore, for each edge in the induced path in $\text{Cay}(F/N, X)$, at least one of its end points will be in the support of h . Since $h_i(g_m)$ is finite for every i and every m , this path must have infinitely many edges. Hence $\text{Supp}(h)$ must be infinite. \square

We now use Theorem 2.2.21, Lemma 2.2.22 and work from Section 2.2.2 to give an estimate of the conjugacy length function of a group F/N' with respect to the subgroup distortion of its cyclic subgroups. Recall that $\bar{\alpha} : F/N' \rightarrow F/N$ denotes the canonical homomorphism and $\delta_{\langle \bar{\alpha}(u) \rangle}^{F/N}$ is the distortion function for the subgroup of F/N generated by $\bar{\alpha}(u)$.

Theorem 2.2.23. *Let N be a normal subgroup of F such that F/N is torsion-free. Let u, v be elements in F/N' . Then u, v are conjugate in F/N' if and only if there exists a conjugator $\gamma \in F/N'$ such that*

$$d_{F/N'}(1, \gamma) \leq (16n^2 + 8n)(2\delta_{\langle \bar{\alpha}(u) \rangle}^{F/N}(4n) + 1).$$

In particular,

$$\text{CLF}_{F/N'}(n) \leq (16n^2 + 8n)(2\Delta_{\text{cyc}}^{F/N}(4n) + 1)$$

where $\Delta_{\text{cyc}}^{F/N}(m) = \sup \left\{ \delta_{\langle x \rangle}^{F/N}(m) \mid x \in F/N \right\}$.

Proof. We begin by choosing a conjugator $(h, \gamma) \in \mathbb{Z}^r \wr F/N$ for which γ is short, as in Lemma 2.2.8. Then Theorem 2.2.21 tells us that there exists some lift γ_0 of γ in F/N' such that $\varphi(\gamma_0) = (h_0, \gamma)$ is a conjugator. First suppose that $u \notin N$. Then $\bar{\alpha}(u)$ has infinite order in F/N and we may apply Lemma 2.2.9 to give us

$$d_M(1, \varphi(\gamma_0)) \leq (n' + 1)P'(2\delta_{\langle \bar{\alpha}(u) \rangle}^{F/N}(P') + 1)$$

where $n' = d_M(1, \varphi(u)) + d_M(1, \varphi(v))$ and $P' = 2n'$ if $\varphi(u)$ is not conjugate to $(1, \bar{\alpha}(u))$ or $P' = n' + \text{CLF}_{F/N}(n')$ otherwise. Lemma 2.2.22 tells us that we can discount the latter situation and set $P' = 2n'$. By Theorem 2.2.5 we see that $d_{F/N'}(1, \gamma) \leq 2d_M(1, \gamma_0)$ and $n' \leq 2n$. Then, provided $u \notin N$, the result follows.

If, on the other hand, $u \in N$, then we must apply the torsion version, Lemma 2.2.11. The upper bound we get on $d_M(1, \varphi(\gamma_0))$ this time will be $P'(2n' + 1)$, where $p' = 2n'$ and $n' \leq 2n$ as in the previous case. This leads to

$$d_{F/N'}(1, \gamma_0) \leq 16n^2 + 4n.$$

Clearly this upper bound suffices to give the stated result. \square

In the special case where $N' = F^{(d)}$ the quotient F/N' is the free solvable group $S_{r,d}$ of rank r and derived length d . Plugging Proposition 2.2.19 into Theorem 2.2.23 gives us an upper bound for the length of short conjugators between two elements in free solvable groups.

Theorem 2.2.24. *Let $r, d > 1$. Then the conjugacy length function of $S_{r,d}$ is bounded above by a cubic polynomial.*

Proof. In light of Proposition 2.2.19, applying Theorem 2.2.23 gives us a conjugator $\gamma \in S_{r,d}$ such that

$$d_{S_{r,d}}(1, \gamma) \leq (16n^2 + 8n)(16n + 1).$$

□

We may ask whether this upper bound is sharp. Indeed, Theorem 2.2.17 tells us it is possible to find elements in $\mathbb{Z}^r \wr \mathbb{Z}^r$ which observe a quadratic conjugacy length relationship. However it seems that this will not necessarily carry through to the free metabelian groups $S_{r,2}$ as the elements u_n and v_n considered in Theorem 2.2.17 cannot be recognised in the image of the Magnus embedding for $S_{r,2}$. Restricting to elements in this image seems to place too many restrictions on the nature of the support of the corresponding functions of the conjugate elements. It therefore seems plausible that the conjugacy length function for $S_{r,2}$ should be subquadratic.

2.2.5 Compression exponents

We can use the fact that the Magnus embedding is a quasi-isometric embedding to obtain a lower bound for the L_p compression exponent of free solvable groups. The L_p compression exponent is a way of measuring how a group embeds into L_p . A non-zero L_p compression exponent implies the existence of a uniform embedding of G into L_p , a notion which was introduced by Gromov [Gro93]. Gromov claimed that if a group admits a uniform embedding into a Hilbert space then the Novikov conjecture holds in this group. The claim was later proved by Yu [Yu00], where he also showed that amenable groups admit such embeddings. Kasparov and Yu [KY06] extended this result to groups which admit a uniform embedding into any uniformly convex Banach space.

Let G be a finitely generated group with word metric denoted by d_G and let Y be a metric space with metric d_Y . A map $f : G \rightarrow Y$ is called a *uniform embedding* if there are two functions $\rho_{\pm} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ such that $\rho_-(r) \rightarrow \infty$ as $r \rightarrow \infty$ and

$$\rho_-(d_G(g_1, g_2)) \leq d_Y(f(g_1), f(g_2)) \leq \rho_+(d_G(g_1, g_2))$$

for $g_1, g_2 \in G$.

One can define the L_p compression exponent for a finitely generated group G , denoted by $\alpha_p^*(G)$, to be the supremum over all $\alpha \geq 0$ such that there exists a

Lipschitz map $f : G \rightarrow L_p$ satisfying

$$Cd_G(g_1, g_2)^\alpha \leq \|f(g_1) - f(g_2)\|$$

for any positive constant C . For $p = 2$, the Hilbert compression exponent, which is denoted by $\alpha^*(G)$, was introduced by Guentner and Kaminker [GK04].

Of particular interest to us is what happens to compression under taking a wreath product. The first estimate for compression exponents in wreath products was given by Arzhantseva, Guba and Sapir [AGS06] where they show that the Hilbert compression exponent of $\mathbb{Z} \wr H$, where H has super-polynomial growth, is bounded above by $1/2$. More recently Naor and Peres have given a lower bound for the compression of $A \wr B$ when B is of polynomial growth [NP11, Theorem 3.1].

Theorem 2.2.25 (Naor–Peres [NP11]). *Let A, B be finitely generated groups such that B has polynomial growth. Then, for $p \in [1, 2]$,*

$$\alpha_p^*(A \wr B) \geq \min \left\{ \frac{1}{p}, \alpha_p^*(A) \right\}.$$

Li showed in particular that a positive compression exponent is preserved by taking wreath products [Li10].

Theorem 2.2.26 (Li [Li10]). *Let A, B be finitely generated groups. For $p \geq 1$ we have*

$$\alpha_p^*(A \wr B) \geq \max \left\{ \frac{1}{p}, \frac{1}{2} \right\} \min \left\{ \alpha_1^*(A), \frac{\alpha_1^*(B)}{1 + \alpha_1^*(B)} \right\}.$$

We can deduce from the result of Naor and Peres that the L_1 compression exponent for $\mathbb{Z}^r \wr \mathbb{Z}^r$ is equal to 1. Hence the L_1 compression exponent for free metabelian groups, using Theorem 2.2.5, is also equal to 1. Then, with the result of Li, induction on the derived length gives us that $\alpha_1^*(S_{r,d}) \geq \frac{1}{d-1}$. Finally, another application of Li's result gives us the following:

Corollary 2.2.27. *Let $r, d \in \mathbb{N}$. Then*

$$\alpha_1^*(S_{r,d}) \geq \frac{1}{d-1}$$

and for $p > 1$

$$\alpha_p^*(S_{r,d}) \geq \frac{1}{d-1} \max \left\{ \frac{1}{p}, \frac{1}{2} \right\}.$$

It would be interesting to determine an upper bound on $\alpha_p^*(S_{r,d})$, in particular to check if it is ever strictly less than $\frac{1}{2}$ since no example of a solvable group with non-zero Hilbert compression exponent strictly less than $\frac{1}{2}$ is known. Austin [Aus11] has constructed solvable groups with L_p compression exponent equal to zero. His examples are modified versions of double wreath products of abelian groups. Therefore to find such an example it seems natural to look in the class of iterated wreath products of solvable groups and special families of their subgroups, such as the free solvable groups.

2.3 Group Extensions

In 1977 Collins and Miller showed that the solubility of the conjugacy problem does not pass to finite index subgroups or to finite extensions [CM77]. Recent work of Bogopolski, Martino and Ventura investigate certain group extensions and what circumstances are necessary for the solubility of the conjugacy problem to carry through to the extension [BMV10]. The extensions they study require a strong assumption to be placed on the structure of centralisers in the quotient group, limiting the application of their work. However, their result applies in cases where the quotient is, for example, cyclic (or indeed finite), enabling them to study such groups as abelian-by-cyclic groups or free-by-cyclic groups.

In this section we look at conjugacy length in extensions similar to those considered in [BMV10], but instead of their assumption on centralisers in the quotient we put a geometric condition on them. We use this to then study conjugacy length in certain abelian-by-cyclic groups, polycyclic abelian-by-abelian groups and in finite extensions.

2.3.1 Twisted and restricted conjugacy length functions

In the following, suppose that Γ is a group which admits a left-invariant metric d_Γ . For $\gamma \in \Gamma$, denote by $|\gamma|_\Gamma$ the distance $d_\Gamma(1, \gamma)$. We will usually omit the subscript in $|\cdot|_\Gamma$ when we discuss lengths in Γ , favouring the subscript notation when dealing with subgroups of Γ .

Twisted conjugacy length

We first recall the *twisted conjugacy problem* in a group Γ . For an automorphism φ of Γ we say two elements $u, v \in \Gamma$ are φ -twisted conjugate if there exists $\gamma \in \Gamma$

such that $u = \gamma v \varphi(\gamma)^{-1}$. In such cases we denote this relationship by $u \sim_\varphi v$. The twisted conjugacy problem asks whether there is an algorithm which, on input an automorphism φ and two elements u and v , determines whether $u \sim_\varphi v$.

Suppose we are given two elements u and v that are φ -twisted conjugate. We can ask what can be said about the length of the shortest γ such that $u = \gamma v \varphi(\gamma)^{-1}$. In particular, we can look for a function $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that whenever $|u| + |v| \leq x$, for $x \in \mathbb{R}_+$, then $u \sim_\varphi v$ if and only if there exists γ such that $|\gamma| \leq f(x)$ which satisfies $u = \gamma v \varphi(\gamma)^{-1}$. We call the minimal such function the φ -twisted conjugacy length function and denote it by $\mathcal{TCL}_\Gamma(n; \varphi)$. Observe that $\text{CLF}_\Gamma(n) = \mathcal{TCL}_\Gamma(n; \text{Id})$. We can extend this notation to subsets $A \subseteq \text{Aut}(\Gamma)$, by defining $\mathcal{TCL}_\Gamma(n; A) = \sup\{\mathcal{TCL}_\Gamma(n, \varphi) : \varphi \in A\}$. The twisted conjugacy length function of Γ is $\mathcal{TCL}_\Gamma(n) = \mathcal{TCL}_\Gamma(n; \text{Aut}(\Gamma))$.

Restricted conjugacy length

Given a subgroup B of a group Γ , the *restricted conjugacy problem* of Γ to B asks if there is an algorithm which determines when two elements $a, b \in B$ are conjugate in Γ (see [BMV10]).

We can associate to the restricted conjugacy length problem a corresponding function, $\text{RCL}_B^\Gamma : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, called the *restricted conjugacy length function* of B from Γ . It is defined to be the minimal function satisfying the property that whenever $|a| + |b| \leq x$, for $a, b \in B$ and $x \in \mathbb{R}_+$, then a is conjugate to b in Γ if and only if there exists a conjugator $\gamma \in \Gamma$ for which $|\gamma| \leq \text{RCL}_B^\Gamma(x)$.

Note that in the definition of the restricted conjugacy length function we always consider the length of the involved players as elements of Γ , rather than using a metric d_B on B . This naturally leads us to a lower bound for the conjugacy length function of Γ :

$$\text{RCL}_B^\Gamma \leq \text{CLF}_\Gamma.$$

In fact we need not even assume that B is a subgroup to define the restricted conjugacy problem of B from Γ and hence RCL_B^Γ . In order for the lower bound above to be useful though, we would need B to be unbounded in d_Γ .

2.3.1.a Example: twisted conjugacy length in free abelian groups

Let $\Gamma = \mathbb{Z}^r$ for some positive integer r . Let $u, v \in \mathbb{Z}^r$ and $\varphi \in \text{SL}_r(\mathbb{Z})$ be diagonalisable with all eigenvalues real and positive. We wish to find some control on the size of the

shortest element $\gamma \in \mathbb{Z}^r$ satisfying

$$u + \varphi(\gamma) = \gamma + v. \quad (2.14)$$

Suppose φ has an eigenvalue equal to 1 with corresponding eigenspace E_1 . Let V be the sum of the remaining eigenspaces, so $\mathbb{R}^n = E_1 \oplus V$. With respect to this decomposition, write

$$\gamma = \gamma_1 + \gamma_2, \quad u = u_1 + u_2, \quad v = v_1 + v_2$$

where $\gamma_1, u_1, v_1 \in E_1$ and $\gamma_2, u_2, v_2 \in V$. Equation (2.14) tells us that $u_1 = v_1$ and $u_2 + \varphi'(\gamma_2) = \gamma_2 + v_2$, where φ' is a matrix which corresponds to the action of φ on V and hence has no eigenvalues equal to 1. We may therefore take $\gamma_1 = 0$ and hence assume that φ has no eigenvalues equal to 1.

Rewrite equation (2.14) as $(\text{Id} - \varphi)\gamma = u - v$. Since 1 is not an eigenvalue of φ , we notice that $\gamma = (\text{Id} - \varphi)^{-1}(u - v)$. Hence

$$\|\gamma\| \leq (1 + \|\varphi\|)(\|u\| + \|v\|).$$

Therefore, if λ is the largest absolute value of an eigenvalue of φ , then

$$\mathcal{TCL}_{\mathbb{Z}^r}(n; \varphi) \leq (1 + \lambda)n.$$

2.3.1.b Example: restricted conjugacy length inside solvable Baumslag–Solitar groups

Consider $\text{BS}(1, m) = \langle a, b \mid aba^{-1} = b^m \rangle$, a solvable Baumslag–Solitar group. This is the semidirect product $\mathbb{Z}[\frac{1}{m}] \rtimes_m \mathbb{Z}$ where the action of \mathbb{Z} on $\mathbb{Z}[\frac{1}{m}]$ is by multiplication by m . The subgroup $\mathbb{Z}[\frac{1}{m}]$ corresponds to the subgroup generated by elements $a^{-r}ba^r$, for non-negative integers r . We will consider here just the subgroup generated by b . One can see that this is exponentially distorted since $b^{m^n} = a^n b a^{-n}$ for any $n \in \mathbb{N}$. In fact, Lemma 2.3.3 in the next Section gives us a more accurate picture:

$$\frac{1}{2} \log_m |r| \leq |b^r| \leq (m + 2) \log_m |r| + \frac{m}{2} + 1. \quad (2.15)$$

Suppose b^r is conjugate to b^s in $\text{BS}(1, m)$, where r and s may be taken to be non-zero, and assume that $|b^r| \leq |b^s|$. Every element in $\text{BS}(1, m)$ can be written uniquely in the normal form $a^{-j}b^l a^k$, for some $j, k, l \in \mathbb{Z}$ with $j, k \geq 0$ and, if j, k are both non-zero, then l is not divisible by m . Write a conjugator for b^r and b^s in this way. Then $b^r a^{-j} b^l a^k = a^{-j} b^l a^k b^s$, which leads to $a^{-j} b^{rm^j+l} a^k = a^{-j} b^{l+sm^k} a^k$. Note

that both sides of this equation are in normal form, since $rm^j + l$ and $l + sm^k$ are divisible by m if and only if l is as well. So $a^{-j}b^l a^k$ is a conjugator if and only if j, k and l satisfy:

$$rm^j = sm^k.$$

Then we may take $l = 0$ and we also have $k - j = \log_m |r| - \log_m |s|$. Since s is a non-zero integer, $\log_m |s| \geq 0$. We noted above that $\log_m |r| \leq 2|b^r|$, so

$$|a^{k-j}| \leq |k - j| \leq 2|b^r| \leq |b^r| + |b^s|.$$

This leads to the restricted conjugacy length function

$$\frac{n-2}{2} \leq \text{RCL}_{\langle b \rangle}^{\text{BS}(1,m)}(n) \leq n$$

where the lower bound follows from looking at the conjugate elements b^{m^r} and b and noting that the shortest conjugator for them is a^r .

2.3.2 Conjugacy length in group extensions

A solution to the conjugacy problem in certain group extensions is given by Bogopolski, Martino and Ventura [BMV10]. Given a short exact sequence

$$1 \longrightarrow F \xrightarrow{\alpha} G \xrightarrow{\beta} H \longrightarrow 1 \quad (2.16)$$

they show that, under certain conditions, the solubility of the conjugacy problem in G is equivalent to the subgroup $A_G = \{\varphi_g \mid \varphi_g(x) = g^{-1}\alpha(x)g, x \in F, g \in G\}$ of $\text{Aut}(F)$ having solvable orbit problem (that is to say, there is an algorithm which decides whether for any element $u \in F$ there is some $\varphi \in A_G$ such that $u = \varphi(u)$). The conditions that must apply to the short exact sequence are the following:

- (a) H has solvable conjugacy problem;
- (b) F has solvable twisted conjugacy problem;
- (c) for every non-trivial $h \in H$, the subgroup $\langle h \rangle$ has finite index in the centraliser $Z_H(h)$, and one can algorithmically produce a set of coset representatives.

Condition (c) is rather restrictive. In particular it implies that centralisers in H need to be virtually cyclic. The types of groups which this includes are typically extensions where H is a finitely generated hyperbolic group.

To study conjugacy length in a group extension, it seems natural therefore that we should require an understanding of the conjugacy length in H and the twisted conjugacy length in F . We should also expect the restricted conjugacy length function of G to F to make an appearance and there should be some condition based upon the centralisers of elements in H .

We will identify F with its image under α . Suppose that G is finitely generated with $|\cdot|$ denoting the word length in G with respect to some finite generating set. It will not always be the case that F is finitely generated. Suppose that d_F is any left-invariant metric on F . For example, we may take d_F to be equal to the word metric on G , or, if F is finitely generated, we may take it to be the word metric on F with respect to some finite generating set for it. We will denote by $|x|_F$ the distance $d_F(e_F, x)$. Let $\delta_F^G : \mathbb{N} \rightarrow \mathbb{N}$ be the subgroup distortion function for F in G , defined by

$$\delta_F^G(n) = \max\{|f|_F : f \in F, |f| \leq n\}.$$

Let $|gh|_H = \min\{|gh| : h \in H\}$ be the quotient metric on H . In the following, the twisted conjugacy length function for F is taken with respect to the metric d_F chosen above.

Theorem 2.3.1. *Let G be given by the short exact sequence (2.16). Suppose that it satisfies the following condition:*

- (c) *there exists a function $\rho : G \rightarrow [0, \infty)$ such that for each $u \in G$ the fundamental domain of $\beta(Z_G(u))$ in $Z_H(\beta(u))$ has diameter bounded above by $\rho(u)$.*

Then $u, v \in G$ are conjugate in G if and only if there exists a conjugator $g \in G$ such that either:

1. $|g| \leq \text{RCL}_F^G(n)$; or
2. $g = ha$ where $|\beta(h)|_H \leq \text{CLF}_H(n) + \rho(u)$ and $|a|_F \leq \mathcal{TCL}_F(2\delta_F^G(n + \rho(u)); \varphi_u)$.

where $n = |u| + |v|$. This leads to an upper bound on the conjugacy length function of

$$\text{CLF}_G(n) \leq \max \left\{ \text{RCL}_F^G(n), \text{CLF}_H(n) + \rho_n + \mathcal{TCL}_F \left(2\delta_F^G(n + \rho_n); A_G^{(n)} \right) \right\}$$

where $\rho_n = \max\{\rho(u) \mid u \in G, |u| \leq n\}$ and $A_G^{(n)} = \{\varphi_u \in A_G \mid u \in G, |u| \leq n\}$.

Proof. We split the proof into various cases, according to the relationship between $\beta(u)$ and $\beta(v)$, beginning with the easiest case. Throughout we will make the assumption that $|u| \leq |v|$.

CASE 1: $\beta(u) = \beta(v) = e_H$.

In this case u and v lie in the image of α . We therefore find a conjugator $x \in G$ such that $v = x^{-1}ux$ and $|x| \leq \text{RCL}_F^G(|u| + |v|)$.

CASE 2: $\beta(u) = \beta(v) \neq e_H$.

We need to reduce this case to the twisted conjugacy problem in F . Let \mathcal{H} be a set of left-coset representatives of F in G satisfying $|h| = |\beta(h)|_H$ for each $h \in \mathcal{H}$. Let g be any conjugator for u and v . We can write this as a product hf for some $h \in \mathcal{H}$ and $f \in F$. Consider the set

$$\mathcal{H}_g = \{h \in \mathcal{H} \mid \exists f \in F \text{ such that } hf \in Z_G(u)g\}.$$

Note that the image under β of \mathcal{H}_g will be precisely the image of $Z_G(u)g$: to say h is in \mathcal{H}_g is equivalent to saying there exists some $f \in F$ such that $hf \in Z_G(u)g$ and since $\beta(h) = \beta(hf)$ for all $f \in F$ we see that $\beta(\mathcal{H}_g) = \beta(Z_G(u)g)$.

Choose $h \in \mathcal{H}_g$ with $\beta(h)$ of minimal size. Since $\beta(u) = \beta(v)$ we deduce that $\beta(Z_G(u)g) \subseteq Z_H(\beta(u))$. Hence we may apply condition (c') and assume $|h| \leq \frac{1}{2}\rho(u)$.

Since $\beta(h) \in Z_H(\beta(u))$, it follows that $h^{-1}uh = uf_h$ for some $f_h \in F$. Also $\beta(u) = \beta(v)$ implies $u^{-1}v = f \in F$. Let $a \in F$ satisfy the twisted conjugacy relation

$$f = \varphi_u(a)^{-1}f_h a. \quad (2.17)$$

We will first show that ha is a conjugator for u and v and then show we have a control on its size. By unscrambling equation (2.17) we obtain the following:

$$\begin{aligned} u^{-1}v = f &= u^{-1}a^{-1}uf_h a \\ &= u^{-1}a^{-1}h^{-1}uha. \end{aligned}$$

Hence $v = (ha)^{-1}u(ha)$ as required.

The size of a is controlled by the twisted conjugacy length function of F :

$$|a|_F \leq \mathcal{TCL}_F(|f|_F + |f_h|_F; \varphi_u).$$

Applying the distortion function gives us $|f|_F \leq \delta_F^G(n)$. Meanwhile $f_h = [u, h]$, so $|f_h|_F \leq \delta_F^G(2|u| + 2|h|) \leq \delta_F^G(n + \rho(u))$, since we have $|u| \leq |v|$. In summary, we have found a conjugator ha satisfying

$$|ha| \leq \rho(u) + \mathcal{TCL}_F(2\delta_F^G(n + \rho(u)); \varphi_u).$$

CASE 3: $\beta(u) \neq \beta(v)$.

Let u, v be conjugate elements in G . Then in particular $\beta(u)$ is conjugate to $\beta(v)$ in H . Apply the conjugacy length function of H and we get that there exists $h_0 \in H$ such that $\beta(u) = h_0^{-1}\beta(v)h_0$ and

$$|h_0|_H \leq \text{CLF}_H(|\beta(u)|_H + |\beta(v)|_H).$$

Let g_0 be a minimal length element in the pre-image $\beta^{-1}(h_0)$. Set $v_0 = g_0^{-1}vg_0$. Then $\beta(v_0) = \beta(u)$ and v_0 is conjugate to u via an element g_0 satisfying

$$|g_0| = |h_0|_H \leq \text{CLF}_H(n).$$

Now we apply Case 2, above, to find a bounded conjugator ha for u and v_0 . Then all we need to do is to pre-multiply it by g_0 to obtain a conjugator for u and v . In other words, we have a conjugator g_0ha for u and v such that

$$|g_0ha| \leq \text{CLF}_H(n) + \rho(u) + \mathcal{TCL}_F(2\delta_F^G(2n + 2\rho(u)); \varphi_u).$$

This is enough to complete the proof. □

By taking a group extension with cyclic quotient we can reduce this to a simpler expression.

Corollary 2.3.2. *Suppose in the extension given in (2.16) the quotient H is \mathbb{Z} . Then*

$$\text{CLF}_G(n) \leq \max \left\{ \text{RCL}_F^G(n), n + \mathcal{TCL}_F \left(2\delta_F^G(2n); A_G^{(n)} \right) \right\}$$

where $A_G^{(n)} = \{\varphi_u \in A_G \mid u \in G, |u| \leq n\}$.

Proof. Let u, v be conjugate in G such that $|u| + |v| \leq n$. Since $H = \mathbb{Z}$, the conjugacy length function of H is the zero function. Furthermore we can put $\rho(u) = |u| \leq n$. □

A central extension is another situation where the expression is significantly simplified. Unlike with the cyclic extensions, we retain the need to understand the function ρ . In particular, if F is contained in the centre of G then Theorem 2.3.1 reduces to

$$\text{CLF}_G(n) \leq \text{CLF}_H(n) + \rho_n.$$

However we can see from this an example of the limitations of this result. If we take the Heisenberg group,

$$H_3(\mathbb{Z}) = \langle x, y, z \mid [x, z] = [y, z] = e, [x, y] = z \rangle$$

then this fits into a central extension of the form of (2.16) with $F = \langle z \rangle$. However, it is not hard to see that the centraliser of x consists precisely of elements of the form $x^r z^s$, for any pair of integers r, s . Projecting this centraliser onto $H_3(\mathbb{Z})/\langle z \rangle \cong \mathbb{Z}^2$ gives a copy of \mathbb{Z} , implying that ρ_n cannot be finite and Theorem 2.3.1 does not apply.

2.3.2.a Solvable Baumslag–Solitar groups

In Section 2.3.3 we will look at abelian-by-cyclic groups. Solvable Baumslag–Solitar groups are examples of such groups and we look at their conjugacy length function here. In [Sal11] we showed that these groups have linear conjugacy length function. In this thesis we have modified the proof: the method we apply here is designed to use Theorem 2.3.1, and in particular Corollary 2.3.2. This represents a simplified version of the arguments used for a wider class of abelian-by-cyclic groups, which appear in Section 2.3.3.

Let $\text{BS}(1, m)$ be the solvable Baumslag–Solitar group with presentation

$$\langle a, b \mid aba^{-1} = b^m \rangle.$$

In Example 2.3.1.b we obtained an upper bound for the restricted conjugacy length function of $\langle b \rangle$ in $\text{BS}(1, m)$. This will be used below to obtain a linear upper bound for the restricted conjugacy length function of the normal closure $\langle\langle b \rangle\rangle$ in $\text{BS}(1, m)$. In the example we needed equation (2.15), which gives a control on the word length of $|b^r|$ for any $r \in \mathbb{Z}$. We will now prove this statement.

Lemma 2.3.3. *Let $r \in \mathbb{Z}$. Then*

$$\frac{1}{2} \log_m |r| \leq |b^r| \leq (m + 2) \log_m(r) + \frac{m}{2} + 1.$$

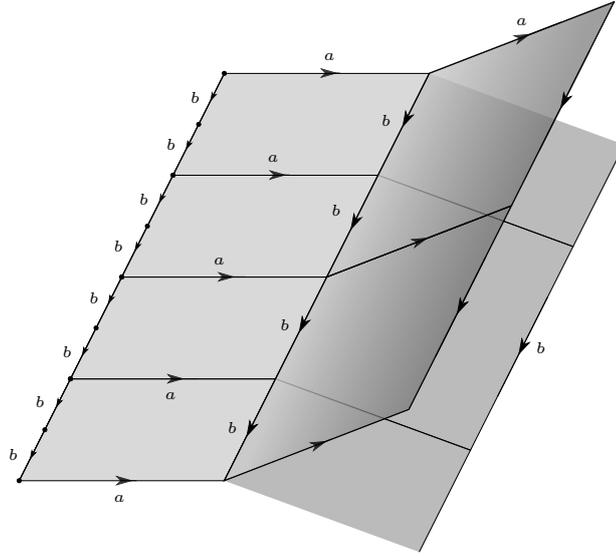


Figure 2.10: A section of the Cayley graph of $BS(1, 2)$.

Proof. We will prove this for $r > 0$. When r is negative we look at $|b^{-r}|$ instead. Let $k \in \mathbb{N}$ satisfy $m^{k-1} \leq r < m^k$. First observe that $|b^{m^k}| \leq 2k + 1 \leq 2 \log_m(r) + 3$. To get an upper bound on $|b^r|$ we need to relate $|b^r|$ to $|b^{m^k}|$. To do this, consider the path in the Cayley graph from 1 to b^r which begins by travelling to $a^{k-1}b$. To get to $b^{m^{k-1}}$ we would then travel backwards along $k - 1$ edges labelled by a . Instead, to get to b^r we should line ourselves up in each coset of $\langle b \rangle$ first, moving along at most m edges labelled by b before moving backwards along an edge labelled by a , see Figure 2.11. This gives an upper bound on $|b^r|$ of $2(k - 1) + (k - 1)m$. Since $|b^{m^k}| \geq 2k$ we get

$$|b^r| \leq \left(\frac{m}{2} + 1\right) |b^{m^k}| - (m + 2) \leq (m + 2) \log_m(r) + \frac{m}{2} + 1.$$

For the lower bound, write b^r as a geodesic word

$$b^r = a^{\alpha_0} b^{\beta_1} a^{\alpha_1} \dots b^{\beta_s} a^{\alpha_s}.$$

By shuffling the terms in a to the left or to the right according to whether α_i is positive or negative we can rewrite this. Let

$$A_i^- = \sum_{j=i}^s \min\{\alpha_j, 0\} \quad \text{and} \quad A_i^+ = \sum_{j=0}^{i-1} \max\{\alpha_j, 0\}.$$

Then

$$b^r = a^{-A} b^B a^A$$

where $A = -A_0^- = A_{s+1}^+$ and B can be expressed as

$$B = \beta_1 m^{A_1^+ - A_1^-} + \dots + \beta_s m^{A_s^+ - A_s^-}.$$

Since $|b^r| = \sum_{i=0}^s |\alpha_i| + \sum_{i=1}^s |\beta_i|$, we get an upper bound on $|B|$ as $|b^r| m^{|b^r|}$. Finally, note that $B = m^A r$ and $A \geq 0$, so $|r| \leq |B|$. This gives enough information to obtain the lower bound in the Lemma. \square

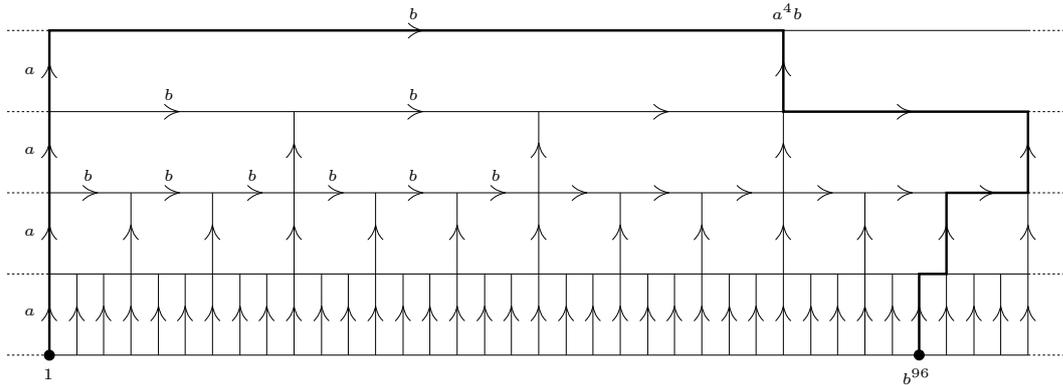


Figure 2.11: The path read out by the word $a^4 b a^{-1} b a^{-1} b^{-1} a^{-1} b^{-1} a^{-1}$ is shown in bold. Notice how from $a^4 b$ to b^{96} we drop down one level and then move along less than 3 edges before dropping down again.

Recall that G is the semidirect product $\mathbb{Z}[\frac{1}{m}] \rtimes_m \mathbb{Z}$. So in the notation above we have H infinite cyclic, generated by a , and $F = \mathbb{Z}[\frac{1}{m}]$, in which b is identified with 1. We can then see that $aba^{-1} = b^m$ (using multiplicative notation) and $\mathbb{Z}[\frac{1}{m}]$ is generated by the elements $a^{-r} b a^r$ for non-negative integers r . Since $\langle a^{-\alpha} b a^\alpha \rangle$ contains $\langle a^{-\beta} b a^\beta \rangle$ whenever $\beta \leq \alpha$, for any $u \in \mathbb{Z}[\frac{1}{m}]$ there exists a minimal integer p such that $u = a^{-p} b^r a^p$ for some $r \in \mathbb{Z}$.

The following Lemma provides a couple more useful tools for determining word lengths in $\text{BS}(1, m)$.

Lemma 2.3.4. *Suppose $u \in \mathbb{Z}[\frac{1}{m}] < \text{BS}(1, m)$ and we can write it as $a^{-p} b^r a^p$, with p the minimal such integer. Then*

(i) $|u| \geq |p|$; and

(ii) $|u| \geq \frac{1}{3} |b^r|$.

Proof. The Cayley graph of $\text{BS}(1, m)$ is a well-understood object, see for example [ECH⁺92, §7.4]. It is the horocyclic product of an $(m + 1)$ -regular tree T and a

2–complex C that is homeomorphic to \mathbb{R}^2 . The cells of C are all identical, they look like bricks with a being read in the same direction along two opposite edges, and for the other pair of edges one reads b along one edge and b^m along the other. Figure 2.10 shows a small section of the Cayley graph of $\text{BS}(1, 2)$.

Suppose that $p \geq 0$. We can consider the projection of the path which the word $a^{-p}b^r a^p$ determines in the Cayley graph onto the tree T . Because we have chosen p to be minimal, m cannot divide r . Hence at some vertex in the coset $\langle b \rangle a^p$ the path will have to move into a different copy of C , meaning the projection onto T will take a different branch. One can see that any path from 1 to u in the Cayley graph will project onto a path in T of length at least $2p$. Since this projection cannot increase path length, we see that $|u| \geq 2p$.

Now suppose that $p < 0$. Then $u = a^{-p}b^r a^p = b^{rm-p}$ and by Lemma 2.3.3, $|u| \geq \frac{1}{2} \log_m |r| + |p|$. Since we may assume that $r \neq 0$, we get that $|u| \geq |p|$ and we have that part (i) holds.

Finally, (ii) follows by writing $b^r = a^p u a^{-p}$, applying the triangle inequality and part (i). \square

We now apply Theorem 2.3.1 to the solvable Baumslag–Solitar groups, obtaining a linear conjugacy length function.

Theorem 2.3.5. *Let $G = \text{BS}(1, m)$. Then there exists a constant $\kappa_m > 0$ depending on m such that*

$$\text{CLF}_G(n) \leq \kappa_m n.$$

REMARK: The keen reader will observe that the upper bound obtained here for the conjugacy length function of $\text{BS}(1, m)$ is very much dependent upon m . In [Sal11], where the author originally proved this result, the length of a conjugator was given in terms of the displacement of a basepoint in a space X which is quasi-isometric to $\text{BS}(1, m)$. The upper bound on this displacement was independent of m , but the precise quasi-isometric relationship between X and $\text{BS}(1, m)$ does indeed vary with m .

Proof of Theorem 2.3.5. In Example 2.3.1.b we considered the restricted conjugacy length function for $\langle b \rangle$ and showed it is linear. Now we need to consider this for the larger subgroup $\mathbb{Z}[\frac{1}{m}]$ instead.

Suppose $u, v \in \mathbb{Z}[\frac{1}{m}]$ are conjugate in $\text{BS}(1, m)$. Let p and q be the minimal integers such that $u = a^{-p}b^r a^p$ and $v = a^{-q}b^s a^q$ for some $r, s \in \mathbb{Z}$. Suppose x is a conjugator for u and v . This is equivalent to $a^p x a^{-q}$ being a conjugator for b^r and b^s . Thus we may take $x = a^{k+q-p}$ so a^k is a conjugator for b^r and b^s . By Example 2.3.1.b, k can be chosen so that

$$|a^k| \leq \text{RCL}_{(b)}^G(|b^r| + |b^s|) \leq |b^r| + |b^s|.$$

Using Lemma 2.3.4, the size of x is therefore bounded above by $4n$. So in particular:

$$\text{RCL}_{\mathbb{Z}[\frac{1}{m}]}^G(n) \leq 4n.$$

Now we need to understand the twisted conjugacy length function for $F = \mathbb{Z}[\frac{1}{m}]$, considering the automorphisms which correspond to multiplying by a power of m . The subgroup $\mathbb{Z}[\frac{1}{m}]$ is not finitely generated, so in the following we will take on it the metric induced from the word metric on G . Then $\delta_F^G(n) = n$.

Let φ denote multiplication by m , or equivalently conjugation by a , specifically $\varphi(x) = axa^{-1}$. For $u, v, x \in \mathbb{Z}[\frac{1}{m}]$, let p, q, y be minimal non-negative integers such that $u = a^{-p}b^r a^p$, $v = a^{-q}b^s a^q$ and $x = a^{-y}b^t a^y$. Then $u\varphi^i(x) = xv$ is equivalent to:

$$a^{-p}b^r a^{p-y+i}b^t a^{y-i} = a^{-y}b^t a^{y-q}b^s a^q.$$

After writing these in the normal form we observe the following:

$$p = y, \quad r + tm^i = t + sm^{p-q} \quad \text{if } p - y + i \geq 0, y - q \geq 0 \quad (2.18)$$

$$p = q, \quad r + tm^{p-y+i} = tm^{y-p} + s \quad \text{if } p - y + i \geq 0, y - q \leq 0 \quad (2.19)$$

$$y - i = y, \quad rm^{y-p} + t = t + sm^{y-q} \quad \text{if } p - y + i \leq 0, y - q \geq 0 \quad (2.20)$$

$$y - i = q, \quad rm^{q-p} + t = tm^{-i} + s \quad \text{if } p - y + i \leq 0, y - q \leq 0 \quad (2.21)$$

Case (2.20) implies that $i = 0$. In this case the automorphism becomes trivial and we are left calculating the conjugacy length function of an abelian group, which is zero. Hence we may assume that $i \neq 0$. Also, (2.18) and (2.21) are equivalent, but with the roles of r and s interchanged. Hence it is enough to consider just (2.18) and (2.19).

From (2.18), since $p = y$ we have $y \leq n = |u| + |v|$ by Lemma 2.3.4. In this case $i \geq 1$. This implies that $m^i - 1 \geq m - 1 \geq 1$, hence $|t| \leq |s|m^{p-q} + |r|$. We

may assume that r and s are both non-zero, otherwise u and v are both the identity. Hence²

$$\log_m |t| \leq \log_m |s| + p - q + \log_m |r| + \log_m 2 \leq 2n + \log_m 2.$$

Now suppose (2.19) holds instead. Firstly notice that y is bounded above by n since $y \leq q$. Next, if $y = p$ then equation (2.19) reduces to $t(m^i - 1) = s - r$. Since $i > 0$ and $m > 1$ we get

$$\log_m |t| \leq \log_m |s| + \log_m |r| + \log_m 2 \leq n + \log_m 2.$$

If instead $y > p$, then $m^{-p+y+i} - m^{-y+p} \geq m^{i+1} - m^{-1} \geq m^i$, so

$$|t| \leq \frac{|s| + |r|}{m^i}.$$

In particular, this also gives

$$\log_m |t| \leq -i + \log_m |s| + \log_m |r| + \log_m 2 \leq n + \log_m 2 + |i|.$$

Hence, by Lemma 2.3.3, in either case we get

$$|x| \leq 2y + (m+2) \log_m |t| + \frac{m}{2} + 1 \leq (2m+6)n + (m+2) |i| + m \left(\log_m 2 + \frac{1}{2} \right) + \log_m 2 + 1$$

and we have an upper bound for the twisted conjugacy length function:

$$\mathcal{TCL}_{\mathbb{Z}[\frac{1}{m}]}(n; \varphi^i) \leq (2m+6)n + (m+2) |i| + m \left(\log_m 2 + \frac{1}{2} \right) + \log_m 2 + 1$$

for any integer i . By applying Corollary 2.3.2, since we need only consider when $|i| \leq n$, we get that

$$\text{CLF}_G(n) \leq (5m+15)n + m \left(\log_m 2 + \frac{1}{2} \right) + \log_m 2 + 1.$$

□

2.3.3 Abelian-by-cyclic groups

An abelian-by-cyclic group Γ has a short exact sequence

$$1 \longrightarrow A \longrightarrow \Gamma \longrightarrow \mathbb{Z} \longrightarrow 1$$

²Suppose $x \geq y \geq 2$. Then $\frac{x}{y}(y-1) \geq 1$, hence $xy \geq x+y$. Thus $\log(2x+2y) \leq \log(2x) + \log(2y)$ for $x, y \geq 1$, and so $\log(x+y) \leq \log x + \log y + \log 2$.

where A is an abelian group. Following the work of Bieri and Strebel [BS78], the finitely presented, torsion-free, abelian-by-cyclic groups are given by presentations of the form

$$\Gamma_M = \langle t, a_1, \dots, a_d \mid [a_i, a_j] = 1, ta_it^{-1} = \varphi_M(a_i); i, j = 1, \dots, d \rangle$$

where $M = (m_{ij})$ is a $d \times d$ matrix with integer entries and non-zero determinant and $\varphi_M(a_i) = a_1^{m_{1i}} \dots a_d^{m_{di}}$ for each $i = 1, \dots, d$. The aim of this section is to give an exponential upper bound for the conjugacy length function of a certain family of abelian-by-cyclic groups:

Theorem 2.3.6. *Suppose M is a diagonalisable matrix, all of whose eigenvalues have absolute value greater than 1. Then there exists a constant C depending on M such that*

$$\text{CLF}_{\Gamma_M}(n) \leq C\lambda^{28n}$$

where λ is the largest absolute value of an eigenvalue of M .

Note that we have considered an example of such a group already in Theorem 2.3.5: the solvable Baumslag–Solitar groups. In Section 2.3.4 we look at the abelian-by-abelian groups of the form $\mathbb{Z}^d \rtimes \mathbb{Z}^k$, where the action on \mathbb{Z}^k corresponds to multiplication by matrices in an \mathbb{R} -split torus inside $\text{SL}_d(\mathbb{Z})$. This of course includes the abelian-by-cyclic groups Γ_M where M is a diagonal matrix in $\text{SL}_d(\mathbb{Z})$ whose eigenvalues are all real.

2.3.3.a Preliminaries

In order to find a control on the conjugacy length function we need to be able to estimate the word length of elements in Γ_M . The method we follow here is in analogy with what we did in Theorem 2.3.5 for $\text{BS}(1, m)$.

Denote by A_p the subgroup generated by $\{t^{-p}a_1t^p, \dots, t^{-p}a_dt^p\}$ for each integer p . Note that by the relation $ta_it^{-1} = \varphi_M(a_i)$, for each integer p we have $A_p \leq A_{p+1}$. Let $|\cdot|_{A_0}$ denote the word metric on A_0 with respect to the generating set $\{a_1, \dots, a_d\}$.

Lemma 2.3.7. *Let $u \in A$. Suppose that p is the minimal non-negative integer such that $u \in A_p$ and suppose $u = t^{-p}u_at^p$. Then*

$$2p \leq |u| \leq 2p + |u_a|_{A_0}.$$

Proof. The upper bound is obvious. For the lower bound consider the word of minimal length representing u :

$$u = t^{r_0} \alpha_1 t^{r_1} \alpha_2 \dots t^{r_{s-1}} \alpha_s t^{r_s}$$

where $\alpha_i \in A_0$ for each i . The word length of u in Γ_M is then equal to

$$|u| = \sum_{i=0}^s |r_i| + \sum_{i=1}^s |\alpha_i|_{A_0}.$$

Note that since u is in A the sum of the exponents of t in the above expression for u is zero. Shuffle the t 's to the left or right according to whether r_i is positive (move t^{r_i} to the right) or negative (move it left). This then gives us

$$u = t^{R^-} \alpha t^{R^+}, \text{ where } R^- = \sum_{i=0}^s \min\{r_i, 0\}, \quad R^+ = \sum_{i=0}^s \max\{r_i, 0\}.$$

Because this process does not change the exponent sum of the t 's in our word for u we notice that $R^+ + R^- = 0$. Furthermore, $2R^+ = R^+ - R^- = \sum |r_i|$. But this expression for u implies that $u \in A_{R^+}$. So by our choice of p we have:

$$p \leq R^+ = \frac{1}{2} \sum_{i=0}^r |r_i| \leq \frac{1}{2} |u|.$$

This proves the Lemma. □

By rewriting elements of A_0 in additive notation, one can see how repeated application of the automorphism φ_M corresponds to taking a power of M . That is, for all $k \in \mathbb{Z}$,

$$\varphi_M^k = \varphi_{M^k}. \tag{2.22}$$

We now investigate the distortion of the subgroup A_0 in Γ_M . Crucial to this is growth of the supremum norm of the matrices M^k , denoted by $\|M^k\|$. If M has an eigenvalue whose absolute value is greater than 1 then $\|M^k\|$ will grow exponentially with respect to k (we assume $k > 0$), while if all eigenvalues have absolute value equal to 1 the growth will be polynomial. See [BG96, Theorem 2.1] for a more accurate statement.

Lemma 2.3.8. *If M has an eigenvalue of absolute value different from 1, then A_0 is exponentially distorted in Γ_M . Otherwise the distortion is polynomial. Furthermore, for $u \in A_0$,*

$$|u|_{A_0} \leq (d \max\{\|M^{|u|}\|, \|M^{-|u|}\|\} + 1) |u|.$$

Proof. In equation (2.22) we saw how $t^k v_a t^{-k}$ is determined by M^k . In particular, for $i = 1, \dots, d$, the length of $t^k a_i t^{-k}$ as an element of A_0 is equal to the sum of the absolute values of the entries of the i -th column of M^k . Hence, for an appropriate choice of i we have $|t^k a_i t^{-k}|_{A_0} \geq \|M^k\|$. But $|t^k a_i t^{-k}| \leq 2k + 1$, so we in fact have that the distortion function of A_0 in Γ_M is bounded below by the growth of $\|M^k\|$ with respect to k .

The upper bound can be deduced, in certain cases, from the proof of Theorem 5.1 in [BP94]. In general though, write $u \in A_0$ as a shortest word

$$u = t^{r_1} u(1) t^{r_2} \dots u(s) t^{r_{s+1}}$$

where $u(i)$ is a word on $\{a_1, \dots, a_d\}$. If we set $R_i = r_1 + \dots + r_i$ then we can rewrite this as

$$u = (t^{R_1} u(1) t^{-R_1}) (t^{R_2} u(2) t^{-R_2}) \dots (t^{R_s} u(s) t^{-R_s}) t^{R_{s+1}}.$$

Since $u \in A_0$, the exponent sum of the stable letter t will be zero. That is, $R_{s+1} = 0$. By equation (2.22), an upper bound on the length of each $t^{R_i} u(i) t^{-R_i}$ in A_0 is given by $d \|M^{R_i}\| |u(i)|_{A_0}$. If we put $R = |r_1| + \dots + |r_{s+1}|$, then $|R_i| \leq R \leq |u|$ for each i and

$$\begin{aligned} |u|_{A_0} &\leq \sum_{i=1}^s d \|M^{R_i}\| |u(i)|_{A_0} \\ &\leq d \max\{\|M^{|u|}\|, \|M^{-|u|}\|\} \sum_{i=1}^s |u(i)|_{A_0} + R \\ &\leq (d \max\{\|M^{|u|}\|, \|M^{-|u|}\|\} + 1) |u|. \end{aligned}$$

Hence the distortion of A_0 in Γ_M is exponential if $\|M^k\|$ grows exponentially with k , or polynomial if the growth of $\|M^k\|$ is polynomial with respect to k . \square

We can obtain the case of the Baumslag–Solitar groups $BS(1, m)$ by setting $M = (m)$. Recall that when we considered elements in A_0 (which in the notation of section 2.3.2.a was the subgroup $\langle b \rangle$), we had a lower bound on their word length given by the logarithm of their length in A_0 . We need a similar result here, but to do so we must restrict which matrices M we are to work with.

Lemma 2.3.9. *Suppose M is diagonalisable with all eigenvalues having absolute value strictly greater than 1. Let λ_1 be the largest absolute value of an eigenvalue of M*

and λ_2 the minimal. If $u \in A_p$, with p the minimal such non-negative integer, and $u = t^{-p}u_a t^p$, then

$$|u_a|_{A_0} \leq \lambda_1^{\frac{3}{2}|u|} |u|.$$

In particular, for $u \in A_0$, there exists a positive constant μ_M depending on the eigenvalues of M such that

$$\log_{\lambda_2} |u|_{A_0} \leq \mu_M |u|.$$

Proof. Let u be represented by a geodesic word

$$u = t^{r_1}u(1)t^{r_2} \dots u(s)t^{r_{s+1}}$$

where each $u(i) \in A_0$. It follows that $|u| = |u(1)|_{A_0} + \dots + |u(s)|_{A_0} + |r_1| + \dots + |r_{s+1}|$. We can rearrange this word as

$$u = (t^{R_1}u(1)t^{-R_1})(t^{R_2}u(2)t^{-R_2}) \dots (t^{R_s}u(s)t^{-R_s})t^{R_{s+1}}$$

where $R_i = r_1 + \dots + r_i$ and $R_{s+1} = 0$ since the exponent sum of the t 's must be zero as u is in A . By Lemma 2.3.7, $|p| \leq \frac{1}{2}|u|$. Hence, using (2.22),

$$\begin{aligned} |u_a|_{A_0} = |t^p u t^{-p}|_{A_0} &\leq \sum_{i=1}^s \left| \varphi_M^{R_i+p}(u(i)) \right|_{A_0} \\ &\leq \sum_{i=1}^s \lambda_1^{|R_i|+|p|} |u(i)|_{A_0} \\ &\leq \lambda_1^{\frac{3}{2}|u|} |u| \end{aligned}$$

implying the first assertion. The second assertion follows by setting $\mu_M = \frac{5}{2 \log_{\lambda_1}(\lambda_2)}$. \square

2.3.3.b Restricted conjugacy length function of A in Γ_M

We will first find a control on the restricted conjugacy length function of A in Γ_M when M is diagonalisable and acts on each eigenspace by expansion.

Proposition 2.3.10. *Suppose M is diagonalisable with all eigenvalues having absolute value greater than 1. Then the restricted conjugacy length function of A in Γ_M satisfies*

$$\text{RCL}_A^{\Gamma_M}(n) \leq \mu_M n$$

where μ_M is as in Lemma 2.3.9.

Proof. Suppose u and v are distinct elements in A which are conjugate in Γ_M . Let p, q be minimal non-negative integers such that $u \in A_p$ and $v \in A_q$. Since A is abelian, u, v must be conjugate by t^k for some integer k . By reversing the roles of u, v if necessary, we may assume that k is non-negative and that $u = t^k v t^{-k} = \varphi_M^k(v)$. Since $\varphi_M(A_q) \subseteq A_{q-1}$, we see that $u \in A_{q-k}$. But by minimality of our choice of p we have either

1. $p \leq q - k$, so $k \leq q - p$; or
2. $q - k < 0 = p$.

Case (1) can be dealt with using Lemma 2.3.7. Thus from the first situation we get $|t^k| \leq \frac{1}{2}|v|$ and so we have a linear control on the conjugator length between u and v .

Case (2) can only occur if $p = 0$ and $k > q$. Suppose v can be written as $t^{-q} v_a t^q$ and $v_a = a_1^{v_1} \dots a_d^{v_d}$. Then

$$u = t^{k-q} v_a t^{q-k} = \varphi_M^{k-q}(v_a).$$

If $u = a_1^{u_1} \dots a_d^{u_d}$, then equation (2.22) tells us that

$$(u_1, \dots, u_d)^T = M^{k-q} \cdot (v_1, \dots, v_d)^T.$$

This gives us that $|u|_{A_0} \geq \lambda_2^{k-q} |v_a|_{A_0} \geq \lambda_2^{k-q}$, where λ_2 is equal to the minimal absolute value of an eigenvalue of M . Then, by Lemma 2.3.9, we get

$$k \leq \mu_M |u| + q.$$

Note that the value of μ_M calculated in the proof of Lemma 2.3.9 is bounded below by $\frac{5}{2}$, so applying Lemma 2.3.7 gives us the conclusion that $k \leq \mu_M n$, where $n = |u| + |v|$. □

2.3.3.c Twisted conjugacy in A

Maintaining the assumption that M is diagonalisable and has eigenvalues with absolute value greater than 1 we obtain the following result regarding twisted conjugacy in the subgroup A .

Proposition 2.3.11. *Suppose M is diagonalisable with all eigenvalues having absolute value greater than 1. Then for all $i \in \mathbb{Z}$ the twisted conjugacy length function of A satisfies*

$$\mathcal{TCL}_A(n; \varphi_M^i) \leq (\lambda^{|i|} + 1) \left(\lambda^{\frac{5}{2}n} + \lambda^{\frac{3}{2}n} \right) n + n$$

where λ is the largest absolute value of an eigenvalue of M .

Proof. Let $u \in A_p$, $v \in A_q$ and $x \in A_y$ with p, q, y minimal such non-negative integers. Suppose they satisfy the twisted conjugacy relationship $u\varphi_M^i(x) = xv$. This is equivalent to

$$t^{-p}u_a t^p t^{i-y}x_a t^{y-i} = t^{-y}x_a t^y t^{-q}v_a t^q$$

where $u = t^{-p}u_a t^p$, $v = t^{-q}v_a t^q$ and $x = t^{-y}x_a t^y$. We use the relations $t a t^{-1} = \varphi_M(a)$, for $a \in A_0$, to shuffle the occurrences of t in the middle to one end of the word on each side. As with the Baumslag–Solitar Case, the direction it moves depends on whether it has positive or negative exponent. We have four cases:

$$p = y, \quad u_a \varphi_M^i(x_a) = x_a \varphi_M^{p-q}(v_a) \quad \text{if } p + i \geq y \geq q \quad (2.23)$$

$$p = q, \quad u_a \varphi_M^{p+i-y}(x_a) = \varphi_M^{-y+p}(x_a) v_a \quad \text{if } p + i \geq y \leq q \quad (2.24)$$

$$y - i = y, \quad \varphi_M^{y-p}(u_a) x_a = x_a \varphi_M^{y-q}(v_a) \quad \text{if } p + i \leq y \geq q \quad (2.25)$$

$$y - i = q, \quad \varphi_M^{q-p}(u_a) x_a = \varphi_M^{-i}(x_a) v_a \quad \text{if } p + i \leq y \leq q \quad (2.26)$$

Case (2.25) implies $i = 0$, leaving us to calculate the conjugacy length function of an abelian group, which is zero. Hence we may assume that $i \neq 0$ and eliminate case (2.25).

Let $u_a = a^{u_1} \dots a_d^{u_d}$, $v_a = a^{v_1} \dots a_d^{v_d}$ and $x_a = a^{x_1} \dots a_d^{x_d}$. With a slight abuse of notation let $\mathbf{u} = (u_1, \dots, u_d)^T$, $\mathbf{v} = (v_1, \dots, v_d)^T$ and $\mathbf{x} = (x_1, \dots, x_d)^T$. Then the equation in (2.23) gives $i \geq 0$ and:

$$\mathbf{u} + M^i \mathbf{x} = \mathbf{x} + M^{p-q} \mathbf{v}$$

and hence, since 1 is not an eigenvalue of M^i ,

$$\mathbf{x} = (M^i - 1)^{-1} (M^{p-q} \mathbf{v} - \mathbf{u}).$$

This therefore gives us the upper bound

$$|x_a|_{A_0} \leq (\lambda^i + 1) (\lambda^{p-q} |v_a|_{A_0} + |u_a|_{A_0})$$

where $\lambda = \lambda_1$ is the largest absolute value of an eigenvalue of M . By Lemmas 2.3.7 and 2.3.9, and since $y = p$ and $0 \leq p - q \leq n$, we are lead to an upper bound on $|x|$ as follows:

$$\begin{aligned} |x| &\leq |x_a|_{A_0} + 2y \\ &\leq (\lambda^i + 1) \left(\lambda^n \lambda^{\frac{3}{2}|v|} |v| + \lambda^{\frac{3}{2}|u|} |u| \right) + n \\ &\leq (\lambda^i + 1) \left(\lambda^{\frac{5}{2}n} + \lambda^{\frac{3}{2}n} \right) n + n. \end{aligned}$$

Meanwhile, case (2.24) reduces to the following vector equation:

$$\mathbf{x} = (M^i - 1)^{-1} M^{y-p} (\mathbf{v} - \mathbf{u}).$$

Since $y-p \leq 0$, M^{y-p} will only act by contractions. Thus $\|M^{y-p}(\mathbf{v} - \mathbf{u})\| \leq \|\mathbf{v} - \mathbf{u}\|$. Hence

$$|x_a|_{A_0} \leq (\lambda^{|i|} + 1) (|v_a|_{A_0} + |u_a|_{A_0})$$

which implies an upper bound on $|x|$ as follows:

$$|x| \leq (\lambda^{|i|} + 1) \left(\lambda^{\frac{3}{2}|v|} |v| + \lambda^{\frac{3}{2}|u|} |u| \right) + 2y \leq (\lambda^{|i|} + 1) \lambda^{\frac{3}{2}n} n + n.$$

Finally, case (2.26) is very similar to case (2.23). We obtain the equation

$$\mathbf{x} = (M^{-i} - 1)^{-1} (M^{q-p} \mathbf{u} - \mathbf{v}).$$

Note that $0 \neq i = y - q \leq 0$, and $0 \leq q - p \leq n$. As we did for case (2.23) we obtain the upper bound

$$|x| \leq (\lambda^{|i|} + 1) \left(\lambda^{\frac{5}{2}n} + \lambda^{\frac{3}{2}n} \right) n + n.$$

This completes the proof, giving us an upper bound for the twisted conjugacy length function. \square

2.3.3.d Conjugacy length in Γ_M

We now obtain the upper bound for the conjugacy length function of Γ_M when M is a diagonalisable matrix with all eigenvalues greater than 1.

Proof of Theorem 2.3.6. This is a straight-forward application of Corollary 2.3.2. The distortion function $\delta_A^{\Gamma_M}$ is the identity map since we calculated the twisted conjugacy length function in A with respect to the word metric on Γ_M . The set of automorphisms $A_G^{(n)}$ that we are to consider consists of those automorphisms φ_x , where $x \in \Gamma_M$ with

$|x| \leq n$ and $\varphi_x(a) = xax^{-1}$ for $a \in A$. This clearly includes all the automorphisms φ_M^i for $|i| \leq n$. What's more, if $x = t^k a$ for some $a \in A$ and $k \in \mathbb{Z}$, then $\varphi_x = \varphi_{t^k}$. Hence

$$A_G^{(n)} = \{\varphi_M^i \mid |i| \leq n\}.$$

Theorem 2.3.6 now follows from Propositions 2.3.10 and 2.3.11, choosing the constant C appropriately. \square

2.3.4 Semidirect products $\mathbb{Z}^n \rtimes \mathbb{Z}^k$

We will now turn our attention to a class of polycyclic abelian-by-abelian groups $\mathbb{Z}^d \rtimes \mathbb{Z}^k$, where $d > k$. In [Sal11] we showed that these groups, under certain restrictions on the semidirect product, have a linear upper bound on the conjugacy length function if $k = 1$, or an exponential upper bound otherwise. We give the proof again here, modifying it to fit with the framework of Theorem 2.3.1.

We will express an element of $\Gamma = \mathbb{Z}^d \rtimes \mathbb{Z}^k$ as a pair (x, y) , with $x \in \mathbb{Z}^d$ and $y \in \mathbb{Z}^k$. The action of \mathbb{Z}^k on \mathbb{Z}^d is via matrices in an \mathbb{R} -split torus in $\mathrm{SL}_d(\mathbb{Z})$. That is, the semidirect product is defined by $\varphi : \mathbb{Z}^k \rightarrow \mathrm{SL}_d(\mathbb{Z})$ such that the image $\varphi(\mathbb{Z}^k)$ consists of matrices which are simultaneously diagonalisable over \mathbb{R} , all of whose eigenvalues are positive. Thus we can choose a basis of \mathbb{R}^d which consists of common eigenvectors for the matrices in the image $\varphi(\mathbb{Z}^k)$.

Throughout we will let $\|\cdot\|$ denote the ℓ^1 norm on either \mathbb{Z}^d or \mathbb{Z}^k . We begin by giving a method to relate the size of a component of $u \in \mathbb{Z}^d$, with respect to one of the eigenvectors, to the size of $\|u\|$.

Lemma 2.3.12. *Let $u = (u_1, \dots, u_d) \in \mathbb{R}^d$, with the coordinates given with respect to a basis of eigenvectors of the matrices in $\varphi(\mathbb{Z}^k)$. Then there exists a constant α_φ such that for each i such that $u_i \neq 0$*

$$|\log(|u_i|)| \leq \alpha_\varphi \log\|u\|.$$

Proof. We will prove the Lemma for $i = 1$. Let E_1, \dots, E_d be the one-dimensional spaces spanned by each eigenvector for the matrices in $\varphi(\mathbb{Z}^k)$. Then $|u_1|$ corresponds to the distance from u to the hyperplane $E_2 \oplus \dots \oplus E_d$ in a direction parallel to E_1 . In order to obtain a lower bound on $\log|u_1|$ we need to find a lower bound on the distance from u to $E_2 \oplus \dots \oplus E_d$. This lower bound follows from the subspace theorem of Schmidt [Sch72]. Since E_2, \dots, E_m are eigenspaces for $\varphi(y) \in \mathrm{SL}_d(\mathbb{Z})$, there exists

algebraic numbers $\alpha_1, \dots, \alpha_d$ such that $E_2 \oplus \dots \oplus E_d = \{x \in \mathbb{R}^d \mid x \cdot (\alpha_1, \dots, \alpha_d) = 0\}$. Then, for $u \in \mathbb{Z}^d$, we have

$$d(u, E_2 \oplus \dots \oplus E_d) = \frac{|u \cdot (\alpha_1, \dots, \alpha_d)|}{\|(\alpha_1, \dots, \alpha_d)\|}.$$

Thus, by the subspace theorem, for every $\varepsilon > 0$ there exists a positive constant C such that for every $u \in \mathbb{Z}^d$ we have the following bound on the distance to the hyperplane:

$$d(u, E_2 \oplus \dots \oplus E_d) \geq \frac{C}{\|u\|^{d-1+\varepsilon}}.$$

In particular this gives us a lower bound on $|u_1|$ and hence

$$\log |u_1| \leq (d-1+\varepsilon) \log \|u\| - \log C.$$

By a trigonometric argument we can obtain an upper bound on $\log |u_1|$ which will depend on the angles between the eigenspaces. Hence, combining this with the lower bound, there exists a positive constant α_φ , determined by d , ε and φ , such that $|\log(|u_1|)| \leq \alpha_\varphi \log \|u\|$. \square

It is important to understand the distortion of the \mathbb{Z}^d component in Γ . The following two lemmas give us a handle on this. The second, Lemma 2.3.14, gives an exponential upper bound for the distortion function while the first, Lemma 2.3.13, shows that for any element of the form $(x, 0)$ in Γ we can take a significant shortcut to get to x from 0 by using the action of one of the matrices in $\varphi(\mathbb{Z}^k)$.

Lemma 2.3.13. *Suppose that $\varphi(\mathbb{Z}^k)$ contains a matrix whose eigenvalues are all distinct from 1. Then there exists a constant $A_\varphi > 0$ such that for every $x \in \mathbb{Z}^d$*

$$|(x, 0)|_\Gamma \leq A_\varphi \log \|x\|.$$

Proof. We will use that fact that $\Gamma = \mathbb{Z}^d \rtimes_\varphi \mathbb{Z}^k$ is a uniform lattice in $G = \mathbb{R}^d \rtimes_\varphi \mathbb{R}^k$, finding an upper bound for $|(x, 0)|_G$ and thus obtaining an upper bound on $|(x, 0)|_\Gamma$. Here, by $|(x, 0)|_G$ we mean the distance $d_G(1, (x, 0))$ where d_G is a left-invariant Riemannian metric on G .

First, let y be a minimal length vector in \mathbb{Z}^k with the property that $\varphi(y)$ has no eigenvalues equal to 1. Write x in coordinates (x_1, \dots, x_d) with respect to a basis of eigenvectors for $\varphi(y)$, which have been chosen so that $x_i \geq 0$ for each i . Let e_i denote the vector with 0's everywhere except in the i -th coordinate where we put a 1 and

let λ_i be the eigenvalue of $\varphi(y)$ corresponding to the eigenvector e_i . For $i = 1, \dots, d$ let $\alpha_i \in \mathbb{R}$ be such that $\lambda_i^{\alpha_i} = x_i$. Then

$$(x, 0) = (0, \alpha_1 y)(e_1, \alpha_2 y - \alpha_1 y)(e_2, \alpha_3 y - \alpha_2 y) \dots (e_d, -\alpha_d y).$$

Calculating the length of each term in the product gives us an upper bound on $|(x, 0)|_G$ as

$$\begin{aligned} |(x, 0)|_G &\leq |\alpha_1| \|y\| + 1 + |\alpha_2 - \alpha_1| \|y\| + 1 + |\alpha_3 - \alpha_2| \|y\| + \dots + 1 + |\alpha_d| \|y\| \\ &\leq d + 2\|y\| \sum_{i=1}^d |\alpha_i|. \end{aligned}$$

But $\alpha_i = \frac{\log(x_i)}{\log(\lambda_i)}$, so by applying Lemma 2.3.12 we get $|\alpha_i| \leq \alpha_\varphi \frac{\log\|x\|}{|\log(\lambda_i)|}$. Hence

$$|(x, 0)|_G \leq d + 2\|y\| \alpha_\varphi \log\|x\| \sum_{i=1}^d \frac{1}{|\log(\lambda_i)|}.$$

Combining this with the fact that Γ is a uniform lattice in G gives us the existence of the constant A_φ given in the Lemma. \square

Lemma 2.3.14. *Suppose the image $\varphi(\mathbb{Z}^k)$ is generated by matrices $\varphi_1, \dots, \varphi_k$ and that λ is the largest eigenvalue of any of the matrices $\varphi_1, \varphi_1^{-1}, \dots, \varphi_k, \varphi_k^{-1}$. Let $x \in \mathbb{Z}^d$. Then*

$$\|x\| \leq |(x, 0)|_\Gamma (\lambda^{|(x, 0)|_\Gamma} + 1).$$

In particular this implies that the distortion function is bounded above by an exponential:

$$\delta_{\mathbb{Z}^d}^\Gamma(n) \leq n(\lambda^n + 1).$$

Proof. The generators of Γ are taken to be the set of elements of the form either $(e_i, 0)$ or $(0, e_j)$ where $e_i \in \mathbb{Z}^d$, $e_j \in \mathbb{Z}^k$ are elements of the standard bases. We can write $(x, 0)$ as a geodesic word on these generators, grouping together the generators of the form $(e_i, 0)$ and the generators of the form $(0, e_j)$:

$$\begin{aligned} (x, 0) &= (\alpha_1, 0)(0, \beta_1)(\alpha_2, 0) \dots (\alpha_r, 0)(0, \beta_r) \\ &= (\alpha_1, \beta_1)(\alpha_2, \beta_2) \dots (\alpha_r, \beta_r) \end{aligned}$$

where α_i and β_i are non-zero for all $1 \leq i \leq r$, except possibly for α_1 and β_r . First note that $|(x, 0)|_\Gamma = \sum_{i=1}^r (\|\alpha_i\| + \|\beta_i\|)$. We also obtain the following expression for x :

$$x = \alpha_1 + \varphi(\beta_1)(\alpha_2) + \varphi(\beta_1 + \beta_2)(\alpha_3) + \dots + \varphi(\beta_1 + \dots + \beta_{r-1})(\alpha_r).$$

Since $\|\beta_1 + \dots + \beta_i\| \leq |(x, 0)|_\Gamma$ for each i , we get an upper bound on the norm of x :

$$\|x\| \leq \|\alpha_1\| + \lambda^{|(x,0)|_\Gamma} (\|x_2\| + \dots + \|x_r\|) \leq |(x, 0)|_\Gamma + \lambda^{|(x,0)|_\Gamma} |(x, 0)|_\Gamma.$$

Thus the lemma is proved. \square

We now give an upper bound on the conjugacy length function of $\mathbb{Z}^d \rtimes \mathbb{Z}^k$. We will see that when $k = 1$ we can produce a linear upper bound, but when $k > 1$ we have to settle for exponential. The main obstacle that prevents us from finding a better than exponential upper bound is the nature of the projection of centralisers of elements into the \mathbb{Z}^k coordinate. In the language of Theorem 2.3.1, this is measured by the function ρ .

Theorem 2.3.15. *Let $\Gamma = \mathbb{Z}^d \rtimes_\varphi \mathbb{Z}^k$, where the image of $\varphi : \mathbb{Z}^k \hookrightarrow \mathrm{SL}_d(\mathbb{Z})$ is contained in an \mathbb{R} -split torus T . Then there exist constants $A > 1$ and $B > 0$ such that*

$$(1) \text{ if } k = 1 \text{ then } \mathrm{CLF}_\Gamma(n) \leq Bn;$$

$$(2) \text{ if } k > 1 \text{ then } \mathrm{CLF}_\Gamma(n) \leq A^n.$$

Proof. We will apply Theorem 2.3.1. Since Γ is abelian-by-abelian we need to find bounds only for the values of $\mathrm{RCL}_{\mathbb{Z}^d}^\Gamma(n)$, $\mathcal{TC}\mathcal{L}_{\mathbb{Z}^d}(n; \varphi)$ and $\rho(u, v)$, for $(u, v) \in \Gamma$, where ρ is the function as defined in Theorem 2.3.1.

STEP 1: Estimating $\mathrm{RCL}_{\mathbb{Z}^d}^\Gamma(n)$.

Consider $(u, 0), (w, 0) \in \mathbb{Z}^d$. Then $(u, 0)(x, y) = (x, y)(w, 0)$ if and only if $u = \varphi(y)(w)$. Note that we can immediately set x to be zero.

Suppose $\varphi(\mathbb{Z}^k)$ is generated by matrices $\varphi_1, \dots, \varphi_k$, so that if $y = (y_1, \dots, y_k)$ then

$$\varphi(y) = \varphi_1^{y_1} \dots \varphi_k^{y_k}.$$

Fix a basis of eigenvectors of the matrices in T . With respect to this basis, let u, v be represented with coordinates $(u_1, \dots, u_d), (w_1, \dots, w_d)$ respectively. Suppose the eigenvalues of φ_i are $\lambda_{j,i}$ for $j = 1, \dots, d$ and $i = 1, \dots, k$. Then from $u = \varphi(y)(w)$ we get the following system:

$$u_j = \left(\prod_{i=1}^k \lambda_{j,i}^{y_i} \right) w_j.$$

By taking logarithms we see that this system is equivalent to the matrix equation $Ly = a$, where L is the $d \times k$ matrix with (r, s) -entry equal to $\log |\lambda_{r,s}|$ and a is the vector with j^{th} entry equal to $\log |u_j| - \log |w_j|$. Since the matrices $\varphi_1, \dots, \varphi_k$ generate a copy of \mathbb{Z}^k , the columns of L are linearly independent. Hence we may take a non-singular $k \times k$ minor L' and we get a matrix equation $L'y = a'$. By Cramer's Rule, for each $i = 1, \dots, k$ we have

$$y_i = \frac{\det(L^{(i)})}{\det(L')}$$

where $L^{(i)}$ is obtained from L' by replacing the i^{th} column with a' . Hence $|y_i|$ is bounded by a linear expression in the terms $|\log(|u_j|)| + |\log(|w_j|)|$, for $j = 1, \dots, k$, and the coefficients are determined by the choice of φ . Therefore, by Lemma 2.3.12, we obtain an upper bound for each $|y_i|$ as linear sum of $\log(|u|)$ and $\log(|w|)$.

To reach an upper bound for the restricted conjugacy length function, we observe that we are able to find $y \in \mathbb{Z}^k$ such that $(u, 0)(0, y) = (0, y)(w, 0)$ and $\|y\| \leq B_1(\log\|u\| + \log\|w\|)$ for some constant $B_1 > 0$, determined by φ and independent of u, w . Furthermore, because the first coordinate in Γ is exponentially distorted, in particular by using Lemma 2.3.14, we indeed have a linear upper bound on conjugator length:

$$|(0, y)|_{\Gamma} \leq \|y\| \leq B_2 n$$

for some $B_2 > 0$ independent of u, w , and where $n = |(u, 0)|_{\Gamma} + |(w, 0)|_{\Gamma}$.

STEP 2: Estimating $\mathcal{TC}\mathcal{L}_{\mathbb{Z}^d}(n; \varphi(v))$.

This is precisely the situation dealt with in Example 2.3.1.a, where we take $\varphi(v)$ in place of φ . It gives us

$$\mathcal{TC}\mathcal{L}_{\mathbb{Z}^d}(n; \varphi(v)) \leq (1 + \lambda_v)n$$

where λ_v is the largest eigenvalue of $\varphi(v)$. Suppose λ is the largest eigenvalue of any of the generating matrices $\varphi_1, \dots, \varphi_k$ or their inverses. Then $\lambda_v \leq \lambda^{\|v\|}$.

STEP 3: Estimating $\rho(u, v)$.

If $k = 1$ then we are in the case of a cyclic extension, so we need only recall that $\rho(u, v) = n$ will suffice here. Now suppose $k > 1$. Note that (a, b) is in the centraliser $Z_{\Gamma}(u, v)$ if and only if

$$a = (\text{Id} - \varphi(v))^{-1}(\text{Id} - \varphi(b))u \in \mathbb{Z}^d.$$

We will show that given any $b \in \mathbb{Z}^k$, there exists a constant m which is bounded by an exponential in $\|v\|$ and such that $(\text{Id} - \varphi(v))^{-1}(\text{Id} - \varphi(mb))u \in \mathbb{Z}^d$. This will give an exponential upper bound on $\rho(u, v)$.

Let $L := (\text{Id} - \varphi(v))\mathbb{Z}^d$. Denote by c the absolute value of the determinant of $\text{Id} - \varphi(v)$. Then c is the index of L in \mathbb{Z}^d . Since $\varphi(b)$ commutes with $\varphi(v)$, and hence $\text{Id} - \varphi(v)$, it follows that $\varphi(b)L = L$. Let \bar{u} be the image of u in \mathbb{Z}^d/L . Then there exists some $m \leq c$ such that $\varphi(mb)\bar{u} = \bar{u}$. In particular $(\text{Id} - \varphi(mb))u \in L$.

In the above, if we let b be one of the canonical generators of \mathbb{Z}^d , then we see that we can control each coordinate and obtain an upper bound on $\rho(u, v)$ as $d\lambda^{d\|v\|}$, since $|\det(\text{Id} - \varphi(v))| \leq (1 + \lambda_v)^d$, and $\lambda_v \leq \lambda^{\|v\|}$, with λ_v and λ as in step 2. Hence, in particular, $\rho(u, v) \leq d(1 + \lambda^{(u,v)_\Gamma})^d$.

STEP 4: Estimating $\text{CLF}_\Gamma(n)$.

By Theorem 2.3.1, we use the above bounds on $\text{RCL}_{\mathbb{Z}^d}^\Gamma(n)$, $\mathcal{TC}\mathcal{L}_{\mathbb{Z}^d}(n; \varphi(v))$ and $\rho(u, v)$ to obtain an upper bound on the conjugacy length function of Γ as

$$\text{CLF}_\Gamma(n) \leq \max \{ B_2 n, \rho_n + 2(1 + \lambda^n) \delta_{\mathbb{Z}^d}^\Gamma(n + \rho_n) \}$$

where $\rho_n = n$ if $k = 1$ or $\rho_n = d(1 + \lambda^n)^d$ if $k > 1$. We can improve this however by using Lemma 2.3.13. In Theorem 2.3.1, the term involving the twisted conjugacy length function actually gives the upper bound of the size of an element when calculated with respect to the word metric on the subgroup. In this case the subgroup is \mathbb{Z}^d and we know from Lemma 2.3.13 that every element is at least exponentially distorted. Hence we can undo the effect of the distortion.

To see this, let $g = ha$ be the conjugator that we obtained from Theorem 2.3.1, with $a \in \mathbb{Z}^d$. Then $|h| \leq \rho_n$ and, firstly when $k = 1$, we get

$$|a|_{\mathbb{Z}^d} \leq \mathcal{TC}\mathcal{L}_{\mathbb{Z}^d}(2\delta_{\mathbb{Z}^d}^\Gamma(n + \rho_n); A_\Gamma^{(n)}) \leq 4n(1 + \lambda^n)(1 + \lambda^{2n}) \leq C_1 \lambda^n$$

for some constant $C_1 > 0$ depending on λ . Applying Lemma 2.3.13 then gives us $|a|_\Gamma \leq A_\varphi \log |a|_{\mathbb{Z}^d} \leq A_\varphi \log(\lambda)n + \log(C_1)$. This then leads to the upper bound of the conjugacy length function as

$$\text{CLF}_\Gamma(n) \leq \max \{ B_2 n, n + A_\varphi \log(\lambda)n + \log(C_1) \}.$$

Hence, when $k = 1$ there exists some positive constant B such that $\text{CLF}_\Gamma(n) \leq Bn$.

Now assume $k > 1$. Then a similar process yields a positive constant C_2 such that $|a|_{\mathbb{Z}^d} \leq C_2 \lambda^{\lambda^n}$ and hence, by applying Lemma 2.3.13 to undo one of the exponential functions, we obtain the exponential upper bound

$$\text{CLF}_{\Gamma}(n) \leq A^n$$

for some $A > 1$ when $k > 1$. □

We will now discuss some applications of Theorem 2.3.15 to other situations, firstly to the fundamental group of prime 3-manifolds and secondly to Hilbert modular groups.

Let M be a prime 3-manifold with fundamental group G . Recent work of Behrstock and Druţu [BD11, §7.2] has shown that, when M is non-geometric, there exists a positive constant K such that two elements u, v of G are conjugate only if there is a conjugator whose length is bounded above by $K(|u| + |v|)^2$. Theorem 2.3.15 in the case when $d = 2$ and $k = 1$ deals with the solmanifold case, while a result of Ji, Ogle and Ramsey [JOR10, §2.1] gives a quadratic upper bound for nilmanifolds. These, together with the result of Behrstock and Druţu, give the following:

Theorem 2.3.16. *Let M be a prime 3-manifold with fundamental group G . For each word metric on G there exists a positive constant K such that two elements u, v are conjugate in G if and only if there exists $g \in G$ such that $ug = gv$ and*

$$|g| \leq K(|u| + |v|)^2.$$

The exponential bound in Theorem 2.3.15 arises because of the way the projection of $Z_{\Gamma}(u, v)$ onto the \mathbb{Z}^k -component lies inside \mathbb{Z}^k . In particular, one may ask if the exponential upper bound is sharp:

QUESTION: Can one find a pair of conjugate elements in Γ whose shortest conjugator is exponential in the sum of the lengths of the two given elements?

We now apply Theorem 2.3.15 to the conjugacy of elements in parabolic subgroups of Hilbert modular groups. Such subgroups are isomorphic to a semidirect product $\mathbb{Z}^n \rtimes_{\varphi} \mathbb{Z}^{n-1}$, where φ depends on the choice of Hilbert modular group and the boundary point determining the parabolic subgroup (see for example either [vdG88] or [Hir73]). Because there is a finite number of cusps (see for example [Shi63] or [vdG88]), for each Hilbert modular group there are only finitely many φ to choose from. Hence,

by Theorem 2.3.15, any two elements in a parabolic subgroup are conjugate if and only if there exists a conjugator whose size is bounded exponentially in the sum of the sizes of the two given elements. More specifically:

Corollary 2.3.17. *Let $\Gamma = \mathrm{SL}_2(\mathcal{O}_K)$ be the Hilbert modular group corresponding to a finite, totally real field extension K over \mathbb{Q} of degree n . There exists a positive constant A , depending only on Γ , such that a pair of elements u, v in the same parabolic subgroup of Γ are conjugate in Γ if and only if there exists a conjugator $\gamma \in \Gamma$ such that*

$$d_\Gamma(1, \gamma) \leq A^{d_\Gamma(1, u) + d_\Gamma(1, v)}.$$

Furthermore, if u, v are actually unipotent elements in Γ , then this upper bound is linear.

Proof. Since u, v are in the same parabolic subgroup of Γ then Theorem 2.3.15 gives the first conclusion. The second conclusion, for unipotent elements, follows from the linear upper bound on the restricted conjugacy length function in the proof of Theorem 2.3.15. \square

In Section 3.3 we look at conjugacy inside the unipotent subgroups of general semisimple real Lie groups. We are able to show that certain pairs of elements enjoy a linear conjugacy relationship, partially extending the result for unipotent elements of Corollary 2.3.17.

2.3.5 Behaviour under finite extensions

Collins and Miller [CM77] provided examples of finitely presented groups in which the solution of the conjugacy problem is not stable when looking at index 2 subgroups. Firstly they construct a specific HNN extension in which the conjugacy problem is solvable among words in which the stable letter appears an even number of times, but unsolvable if it appears an odd number of times. Taking the appropriate index 2 subgroup then gives a group with solvable conjugacy problem. Secondly they construct a group L , which is a free product with amalgamation, and show that this does not have solvable conjugacy problem. They explain how a split extension of L using a well-chosen order 2 automorphism of L removes the source of unsolvability for the conjugacy problem, providing an example of a group with solvable conjugacy problem but containing an index 2 subgroup in which it is unsolvable.

It follows from [BMV10, Theorem 3.1] that if one can solve the twisted conjugacy problem in F , then the solubility of the conjugacy problem will pass to finite extensions. When we consider conjugacy length we obtain the following Corollary to Theorem 2.3.1:

Corollary 2.3.18. *Consider the short exact sequence*

$$1 \longrightarrow F \xrightarrow{\alpha} G \xrightarrow{\beta} H \longrightarrow 1$$

in which H is a finite group and F is finitely generated. Then

$$\text{CLF}_G(n) \preceq \mathcal{TCL}_F(n; A_H)$$

where $A_H \subset \text{Aut}(F)$ consists of those automorphisms determined by conjugation by elements in G .

Proof. Let k be such that every element in H has size bounded above by k in the quotient metric. It is clear then that $\text{CLF}_H(n) \leq k$ and $\rho(u) \leq k$ for all $n \in \mathbb{N}$ and $u \in G$.

Let g_1, \dots, g_r be a collection of coset representatives for F in G . Let $u, v \in F$ and suppose they are conjugate via xg_i for some i and some $x \in F$. So $uxg_i = xg_iv$, which is equivalent to $x^{-1}ux = g_ivg_i^{-1}$. The minimal x satisfying this relationship will have size

$$|x| \leq \text{CLF}_F(|u| + |g_ivg_i^{-1}|) \leq \text{CLF}_F(n + 2k).$$

In particular this gives a conjugator xg_i between u, v satisfying

$$|xg_i| \leq k + \text{CLF}_F(n + 2k).$$

This deals with the restricted conjugacy length function of F from G . Then by applying Theorem 2.3.1 we get

$$\text{CLF}_G(n) \leq k + \max\{k + \text{CLF}_F(n + 2k), k + \mathcal{TCL}_F(8n; A_H)\}.$$

Since $\text{CLF}_F(n) = \mathcal{TCL}_F(n; \text{Id})$, the proof is complete. \square

When we take a finite extension of an abelian group, for example, Corollary 2.3.18 and Example 2.3.1.a imply the following:

Corollary 2.3.19. *The conjugacy length function of a virtually abelian group G is at most linear.*

QUESTION: To what extent can \mathcal{TCL}_F be a lower bound for CLF_G when G is an extension of F ? What about when it is a finite extension?

Chapter 3

Higher Rank Lattices

Let Γ be a lattice in a higher-rank semisimple real Lie group G . Elements of Γ can be classified into various types (see Section 3.1.3) which are preserved by conjugation and choice of representation. We will focus our attention on two particular types of elements in Γ .

The first type that we consider are real hyperbolic elements. Geometrically, hyperbolic elements can be characterised as isometries of the symmetric space X which translate points along some geodesic. They may have rotational components, for example one may consider an isometry of \mathbb{R}^3 which acts on the x -axis by translation but acts on the y, z -plane by rotating around the origin. If the rotational component is trivial then we say the isometry is *real hyperbolic*.

We take a very geometric approach to study the conjugacy of real hyperbolic elements. Their centralisers are well understood — they stabilise a maximal flat in X , or a family of maximal flats containing a common geodesic. We study these flats, showing how the distance between them influences the length of a conjugator between two real hyperbolic elements in Γ .

The other type of elements we consider are *unipotent* elements. These are elements $g \in G$ for which $\text{Ad}(g)$ is a unipotent linear transformation — with respect to some basis of \mathfrak{g} , the matrix representing $\text{Ad}(g)$ is upper-triangular with 1's on the diagonal.

We gave here the algebraic definition of a unipotent element because this mirrors the approach taken in Section 3.3. We take advantage of the underlying root system of a semisimple real Lie group and show how it can be used to study the conjugacy of unipotent elements in G .

In both cases, for a pair of conjugate elements $u, v \in \Gamma$, whether they are real

hyperbolic elements or unipotent elements¹ which lie in the same maximal unipotent subgroup, we find a conjugator g which lies in the ambient group G having linearly bounded size:

$$d_G(1, g) \leq K(d_G(1, u) + d_G(1, v))$$

where K is a constant which may depend on certain choices made: in the real hyperbolic case K will depend on the slope of geodesics translated by u, v (this is a notion defined in Section 3.1.4.b); in the unipotent case it will depend on how Γ intersects the maximal unipotent subgroup containing u, v .

This chapter contains three sections. The first section introduces the background material, with particular emphasis on the relationship between Lie groups and their associated symmetric spaces. Section 3.2 deals with the real hyperbolic case while in Section 3.3 we focus on the conjugacy of unipotent elements.

3.1 Symmetric Spaces

Before we look at the conjugacy of elements in a lattice we will give some background on the subject. An important tool when studying the conjugacy of real hyperbolic elements is the geometry of flats in the symmetric space on which the lattice acts. We will describe the relationship between symmetric spaces and Lie groups (Section 3.1.1) and describe some of the structure of symmetric spaces which we use later on (Section 3.1.4). In Section 3.1.2 we look at lattices and see how, thanks to the famous result of Lubotzky, Mozes and Raghunathan [LMR00], we are able to study lattices via their action on the symmetric space. We classify the elements of a semisimple Lie group in Section 3.1.3, comparing the terminologies coming from the theory of algebraic groups and the theory of isometries of CAT(0) spaces. When studying real hyperbolic elements we use an asymptotic cone of the symmetric space. In Section 3.1.5 we define these and state a result of Kleiner and Leeb [KL97] which describes their structure as Euclidean buildings.

3.1.1 Symmetric spaces and Lie groups

Associated to each semisimple real Lie group G is a symmetric space X on which the group acts. We will be considering non-uniform irreducible lattices Γ of G so the

¹For the unipotent case we have added conditions: firstly the Lie algebra of G needs to be split; secondly the elements must lie in a unique maximal unipotent subgroup.

action of Γ on X will not be cocompact. We can however still use properties of X to study conjugacy in Γ and in particular we use the fact that it has non-positive curvature.

Let X be a complete connected Riemannian manifold with a metric $d_X : X \times X \rightarrow [0, \infty)$. At each point p in X we can define a geodesic involution s_p as follows: let $c : \mathbb{R} \rightarrow X$ be a unit speed geodesic such that $c(0) = p$. Then for $t \in \mathbb{R}$ we define $s_p(c(t)) := c(-t)$. If such a geodesic involution can be defined at every point in X and if moreover it is a global isometry of X then we say X is a (globally) *symmetric space*. When X has non-positive sectional curvature and no Euclidean de Rham factors we say it is of *noncompact type*.

To explain why symmetric spaces will be useful we need to see their relationship with semisimple real Lie groups (see for example [Ebe96, Sections 2.1 and 2.2]). First let X be a symmetric space of noncompact type. Then the identity component of its group of isometries $G = \text{Isom}_o(X)$ is a connected semisimple real Lie group with trivial centre and no compact factors. Take K to be the isotropy subgroup at a point p in X , that is $K := G_p = \{k \in G \mid kp = p\}$. Then K is a maximal compact subgroup of G and we can identify X with the quotient G/K . Conversely let G be a connected semisimple real Lie group with trivial centre and no compact factors. Then we can associate to G a unique symmetric space of noncompact type [Hel63, Ch. VI, Thm. 1.1]. Indeed, let K be a maximal compact subgroup of G . We can then put on G/K a Riemannian structure, induced from the Killing form of the Lie algebra of G , which makes it a symmetric space of noncompact type.

Throughout, we will assume X is a symmetric space of noncompact type with associated connected semisimple real Lie group G . For an arbitrary point $q \in X$, we let $G_q = \{g \in G \mid gq = q\}$, and we will often also let K denote the isotropy subgroup G_p of G for the action on X at a fixed basepoint p .

3.1.2 Lattices

A *lattice* in a semisimple real Lie group G is a discrete subgroup Γ such that $\Gamma \backslash G$ has finite volume with respect to the Haar measure on G . If the quotient $\Gamma \backslash G$ is compact, then we say Γ is a *cocompact* or *uniform* lattice in G . Otherwise we say it is *non-uniform*.

A lattice is said to be *irreducible* if it cannot be virtually decomposed as a product of lattices $\Gamma_1 \times \Gamma_2$ in a product of Lie groups $G_1 \times G_2$ such that Γ_i is a lattice in G_i

for each $i = 1, 2$. Equivalently, when G is given as a product of Lie groups $G_1 \times G_2$, then the projections of Γ into each factor are dense if and only if Γ is irreducible.

Lattices are finitely presented groups so we can define on them a word metric $d_\Gamma : \Gamma \times \Gamma \rightarrow \mathbb{R}$ with respect to some finite generating set for Γ . We can also consider the size of an element of Γ using the Riemannian metric on G , $d_G : G \times G \rightarrow \mathbb{R}$. By the following Theorem, provided the real rank of G is at least 2 (we will define this in Section 3.1.4), it does not matter which we use:

Theorem 3.1.1 (Lubotzky–Mozes–Raghunathan [LMR00]). *The word metric d_Γ on an irreducible lattice Γ in a semisimple group G of real rank at least 2 is Lipschitz equivalent to the Riemannian metric d_G on G restricted to $\Gamma \times \Gamma$.*

From here on in we will deal with irreducible lattices. Hence, in light of the result of Lubotzky, Mozes and Raghunathan, determining the size of $\gamma \in \Gamma$ can be done by considering the displacement $d_X(p, \gamma p)$ by γ of a fixed basepoint p in the associated symmetric space X . In particular we have the following consequence of Theorem 3.1.1:

Corollary 3.1.2. *For each $p \in X$ there exists a constant $C_p > 0$ such that*

$$d_\Gamma(1, \gamma) \leq C_p d_X(p, \gamma p)$$

for every $\gamma \in \Gamma$.

3.1.3 Classifications of elements: two terminologies

Let X be a symmetric space of noncompact type. There are two ways to view an isometry of X . One could either see it in its literal sense, as an isometry of a CAT(0) space, or one could use the fact that $G = \text{Isom}_o(X)$ is a semisimple linear algebraic group and use terminology from here. We shall see that in the most part confusion need not arise, however we will need to take care when using the term “hyperbolic” in particular.

We will first give a classification of the elements when considered as isometries (see also [BGS85, Section 6] or [BH99, Ch. II.6]). Given $g \in \text{Isom}_o(X)$ we can consider the distance it moves each point in X by defining the displacement function $d_g(q) := d_X(q, gq)$ for $q \in X$. In particular we consider the infimum of d_g over X and define the set

$$\text{MIN}(g) := \{x \in X \mid d_g(x) = \inf_{q \in X} d_g(q)\}.$$

We use this set to classify the isometries of X :

- if $\text{MIN}(g) \neq \emptyset$ and g has a fixed point in X then we say g is *elliptic*;
- if $\text{MIN}(g) \neq \emptyset$ and g does not have a fixed point then we say g is *hyperbolic* (or *CAT(0)-hyperbolic* or *axial*);
- if $\text{MIN}(g) = \emptyset$ then we say g is *parabolic*.

We say that an isometry is *semisimple* if it is either elliptic or hyperbolic in the above sense. A parabolic element where the infimum is zero is said to be *strictly parabolic*. The above classification can be applied to isometries of any CAT(0) space.

We will now classify the elements of G when considered as members of a semisimple algebraic group (see Onishchik and Vinberg [OV90, Ch.3 §2]). Let V be a finite dimensional real vector space. First recall the following classification of linear transformations $T : V \rightarrow V$. We say T is:

- *nilpotent* if there exists a positive integer k such that $T^k = 0$;
- *unipotent* if $T - I$ is nilpotent, where I is the identity transformation;
- *semisimple* if T is diagonalisable over \mathbb{C} .

We can extend the classification of semisimple transformations by putting conditions on their eigenvalues. We say a semisimple transformation T is:

- *real semisimple* if all the eigenvalues of T are real;
- *hyperbolic* if all the eigenvalues of T are positive;
- *elliptic* if all the eigenvalues of T have modulus one.

An element g of a semisimple group G is said to be *unipotent* (respectively *semisimple*) if there exists a finite-dimensional vector space V and a faithful linear representation $\rho : G \rightarrow \text{GL}(V)$ of G such that $\rho(g)$ is unipotent (respectively semisimple). In fact an equivalent definition of these types of elements would be to say that the image of g in every faithful linear representation is unipotent (respectively semisimple).

Since $\text{Isom}_0(X)$ is a semisimple real Lie group with trivial centre the adjoint representation satisfies the above criteria: it is certainly a linear representation and its kernel is the centre of G , hence it is faithful. Therefore we can say that an isometry g is unipotent (or semisimple) if and only if $\text{Ad}(g)$ is unipotent (or semisimple).

With the representation fixed we can define the *real semisimple*, *hyperbolic* and *elliptic* elements of G to be those $g \in G$ for which $\text{Ad}(g)$ is real semisimple, hyperbolic or elliptic respectively.

We can see now that several terms are used in both contexts. We should ask whether the two terminologies coincide.

By fixing a basis of the Lie algebra \mathfrak{g} we can identify $\text{Ad}(G)$ with a closed, connected subgroup of $\text{GL}_n(\mathbb{C})$, where n is the dimension of \mathfrak{g} . We can consider a decomposition of $\text{GL}_n(\mathbb{C})$ into the product of subgroups KAN , where K consists of orthogonal matrices, A consists of diagonal matrices with all entries positive and N consists of all upper-triangular matrices with 1's on the diagonal. This is an example of the Iwasawa decomposition, which we will see again in Section 3.1.4.

Proposition 3.1.3. *Let g be an element of $G = \text{Isom}_o(X)$. Then:*

1. $\text{Ad}(g)$ is elliptic if and only if $\text{Ad}(g)$ is conjugate in $\text{Ad}(G)$ to an element of K ;
2. $\text{Ad}(g)$ is hyperbolic if and only if $\text{Ad}(g)$ is conjugate in $\text{Ad}(G)$ to an element of A ;
3. $\text{Ad}(g)$ is unipotent if and only if $\text{Ad}(g)$ is conjugate in $\text{Ad}(G)$ to an element of N

This proposition leads to the following relationships between the two terminologies.

Theorem 3.1.4. *Let $G = \text{Isom}_o(X)$, where X is a symmetric space of noncompact type. Let g be any element in G . Then:*

1. g is elliptic as an isometry of a $\text{CAT}(0)$ space if and only if g is elliptic in the sense of semisimple groups;
2. g is hyperbolic as an isometry of a $\text{CAT}(0)$ space if and only if g is non-elliptic semisimple in the sense of semisimple groups;
3. g is semisimple as an isometry of a $\text{CAT}(0)$ space if and only if g is semisimple in the sense of semisimple groups;
4. if g is unipotent then g is strictly parabolic.

Proof. A proof can be found in [Ebe96, 2.19.18]. □

We can therefore use the words “semisimple” and “elliptic” without a problem. However we should take particular care when using the word “hyperbolic” to describe elements of G . Part (2) of Theorem 3.1.4 tells us that if $\text{Ad}(g)$ is hyperbolic then g is hyperbolic as an isometry of a $\text{CAT}(0)$ space. However the converse is not true in general. Therefore when describing an element of G as hyperbolic it makes sense to say $\text{CAT}(0)$ -hyperbolic when in particular we mean hyperbolic in the sense of isometries of $\text{CAT}(0)$ space, or *real hyperbolic* if and only if $\text{Ad}(g)$ is diagonalisable over \mathbb{R} with positive eigenvalues.

We finish this section with a characterisation of $\text{CAT}(0)$ -hyperbolic elements:

Proposition 3.1.5 (Corollary 2.19.19 of [Ebe96]). *Let g be $\text{CAT}(0)$ -hyperbolic in G . Then there exists $h, k \in G$ such that $g = kh = hk$, where h is real hyperbolic and k is elliptic.*

3.1.4 Structure of symmetric spaces

3.1.4.a Flats, the root-space decomposition and regular geodesics

A geodesic in the symmetric space X is an isometric copy of a closed interval in \mathbb{R} . When the interval is actually \mathbb{R} itself then we say the geodesic is *bi-infinite*. If the interval is bounded then we say we have a *geodesic segment*. In the remaining cases, when the interval is of the form $(-\infty, a]$ or $[a, \infty)$, we have a *geodesic ray* in X .

FLATS:

Geodesics are example of 1-dimensional flats. Let r be a positive integer. Then an r -flat in X is a complete, totally geodesic submanifold $F \subset X$ which is isometric to \mathbb{R}^r . The maximal such r is called the *rank* of X and flats of this dimension are called *maximal flats*. The *real rank* of the Lie group $G = \text{Isom}_o(X)$, denoted $\text{rank}_{\mathbb{R}}(G)$, is equal to the rank of X .

Proposition 3.1.6. *The group G acts transitively on the set of flats in X . Furthermore, if F_1, F_2 are two maximal flats in X , and $p_1 \in F_1, p_2 \in F_2$, then there exists $g \in G$ such that:*

- $gp_1 = p_2$; and
- $gF_1 = F_2$.

Proof. This follows from [Hel01, Ch. V Thm 6.4]. □

CARTAN DECOMPOSITION:

The Lie algebra \mathfrak{g} of $G = \text{Isom}_o(X)$ admits a Cartan decomposition at each point $p \in X$, which we will define now. Let s_p be the geodesic involution defined in Section 3.1.1 and let $\sigma_p : G \rightarrow G$ be the isometry defined by

$$\sigma_p(g) = s_p \circ g \circ s_p$$

for $g \in G$. The derivative $\theta_p = ds_p$ of this map is known as the *Cartan involution*. It squares to the identity, so has eigenvalues ± 1 . The Lie algebra of G can be decomposed into a product of the eigenspaces

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$$

where $\mathfrak{k} = \{H \in \mathfrak{g} \mid \theta_p(H) = H\}$ and $\mathfrak{p} = \{H \in \mathfrak{g} \mid \theta_p(H) = -H\}$. The decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is known as the *Cartan decomposition* of \mathfrak{g} at p . The subalgebra \mathfrak{k} is the Lie algebra of the maximal compact subgroup $K = G_p$, while \mathfrak{p} can be identified with T_pX , the tangent space at p of X .

A geodesic $c : \mathbb{R} \rightarrow X$ such that $c(0) = p$ determines a vector in T_pX and hence an element $H \in \mathfrak{p}$. Consider the maximal abelian subspace \mathfrak{a} of \mathfrak{p} which contains H . The submanifold $\exp(\mathfrak{a})p$ is a maximal flat in X . This follows from the following two facts:

- [Hel01, Ch. IV Thm 7.2] for $\mathfrak{s} \subset \mathfrak{p}$, the submanifold $\exp(\mathfrak{s})p$ in X is totally geodesic in X if and only if \mathfrak{s} is a Lie triple system;
- [Hel01, Ch. IV Thm 4.2] for $Y_1, Y_2, Y_3 \in \mathfrak{p}$, the curvature tensor at p is given by $R_p(Y_1, Y_2)Y_3 = -[[Y_1, Y_2], Y_3]$.

Lemma 3.1.7. *Every maximal flat F containing p is of the form $F = \exp(\mathfrak{a})p$ for some maximal abelian subspace \mathfrak{a} of \mathfrak{p} .*

Proof. This follows from [Hel01, Ch. V Prop 6.1]. □

ROOT-SPACE DECOMPOSITION:

Fix a Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Let \mathfrak{a} be a maximal abelian subspace of \mathfrak{p} . Then, since the operators $\text{ad}(H)$, for $h \in \mathfrak{a}$, are simultaneously diagonalisable we can consider, for linear functionals $\lambda : \mathfrak{a} \rightarrow \mathbb{R}$, the eigenspaces

$$\mathfrak{g}_\lambda = \{Y \in \mathfrak{g} \mid \text{ad}(H)Y = \lambda(H)Y \text{ for all } H \in \mathfrak{a}\}.$$

Those λ for which \mathfrak{g}_λ is non-empty are called *roots* of \mathfrak{g} with respect to \mathfrak{a} and the spaces \mathfrak{g}_λ the corresponding *root-spaces*. Let Λ be the set of all non-zero roots of \mathfrak{g} with respect to \mathfrak{a} . We have the follow *root-space decomposition* of \mathfrak{g} :

$$\mathfrak{g} = \mathfrak{g}_0 + \sum_{\lambda \in \Lambda} \mathfrak{g}_\lambda.$$

The roots Λ form a root system in the dual space \mathfrak{a}^* . A subset Π of Λ is called a *base* if it is a basis for \mathfrak{a}^* and if any root λ can be written as

$$\lambda = \sum_{\alpha \in \Pi} c_\alpha \alpha$$

in such a way that either each c_α is non-negative or each c_α is non-positive. The elements of Π are called *simple roots* and those elements for which $c_\alpha \geq 0$ for each $\alpha \in \Pi$ are called *positive roots* with respect to Π . We denote the set of positive roots by Λ^+ . The root system plays an important role in the Section 3.3.

WEYL CHAMBERS; REGULAR AND SINGULAR GEODESICS:

Consider a flat F in X . By Lemma 3.1.7 there exists a maximal abelian subspace \mathfrak{a} of \mathfrak{p} such that $F = \exp(\mathfrak{a})p$. Let Λ be the corresponding set of roots, Π a set of simple roots and Λ^+ the corresponding positive roots. The set of elements $H \in \mathfrak{a}$ for which $\lambda(H) > 0$ for each $\lambda \in \Pi$ forms an open *Weyl chamber* in \mathfrak{a} , denoted \mathfrak{a}^+ . The corresponding set $\mathcal{C}_\Pi = \exp(\mathfrak{a}^+)p$ is called an open Weyl chamber in X . The choice of Π determines the Weyl chamber \mathfrak{a}^+ .

For each root $\lambda \in \Lambda$ the kernel is a hyperplane in \mathfrak{a} . These are called the *singular hyperplanes* of \mathfrak{a} . The *walls* of \mathcal{C}_Π are contained in the singular hyperplanes and are defined, for a subset $\Theta \subset \Pi$, as

$$\mathcal{C}_\Theta = \{\exp(H) \in \overline{\mathcal{C}_\Pi} \mid \lambda(H) = 0 \text{ for } H \in \Pi \setminus \Theta\}$$

where $\overline{\mathcal{C}_\Pi}$ is the closure of \mathcal{C}_Π in F . The flat F is partitioned into Weyl chambers and walls. In fact, after removing all the singular hyperplanes from F , the connected components of what remains are all the Weyl chambers corresponding to the different choices for Π .

Let $c : \mathbb{R} \rightarrow X$ be a geodesic in F with $c(0) = p$. Then there exists $H \in \mathfrak{a}$ such that $c(t) = \exp(tH)$ for every $t \in \mathbb{R}$. If H is contained in a singular hyperplane of \mathfrak{a} then we say the geodesic c is *singular*. Otherwise H is contained in some Weyl chamber \mathcal{C}_Π and we call c a *regular geodesic* in X . Since every geodesic in X is

contained in some maximal flat this definition extends to all geodesics. The following is an equivalent definition of regular and singular geodesics:

Proposition 3.1.8. *A geodesic is regular if and only if it is contained in a unique maximal flat.*

Proof. See [Ebe96, §2.11]. □

Suppose $F = \exp(\mathfrak{a})p$, where \mathfrak{a} is the maximal abelian subspace from Lemma 3.1.7 and let $A = \exp(\mathfrak{a})$ and $K = G_p$. The subgroup

$$N_K(A) = \{k \in K \mid kA = Ak\}$$

consists of all elements in G fixing p and stabilising F . Meanwhile, its subgroup

$$Z_K(A) = \{k \in K \mid ka = ak \forall a \in A\}$$

contains all elements in G which fix every point in F . The quotient $N_K(A)/Z_K(A)$ of these groups is called the *Weyl group* of F at p . It is a finite group and acts transitively on the set of Weyl chambers in F .

3.1.4.b Properties of the boundary

IDEAL BOUNDARY AND SLOPES OF GEODESICS:

Given two geodesic rays ρ_1, ρ_2 in X , we say they are *asymptotic* if they are at finite Hausdorff distance from one-another. This defines an equivalence relation on geodesic rays in X , the equivalence classes of which form the *ideal boundary* $\partial_\infty X$ of X . The action of an isometry $g \in G$ on X can be extended to an action on $\partial_\infty X$ since ρ_1 and ρ_2 are asymptotic if and only if $g\rho_1$ and $g\rho_2$ are asymptotic. Hence we may consider the quotient of the action of G on $\partial_\infty X$. We denote this quotient by Δ_{mod} . The Δ_{mod} -*direction*, or *slope*, of a ray ρ is the image of ρ under the quotient maps.

Consider a bi-infinite geodesic $c : \mathbb{R} \rightarrow X$. This determines two boundary points, one for each end of the geodesic. Although, by the definition of a symmetric space, there is an isometry $\varphi = s_{c(0)}$ of X , the geodesic involution at $c(0)$, such that $\varphi c(t) = c(-t)$ for all $t \geq 0$, this isometry will not be in the connected component of the group of isometries of X , and thus not in G . Thus the two ends of c determine two ideal points corresponding to $c(\infty)$ and $c(-\infty)$, and these will usually give rise to distinct

slopes. We call the *slope of the bi-infinite geodesic* $c : \mathbb{R} \rightarrow X$ the projection of $\sigma(\infty)$ onto Δ_{mod} .

If the geodesic c is regular, then we say the corresponding Δ_{mod} -directions are regular, while if c is singular its slopes are said to be singular too. Equivalently, the regular Δ_{mod} -slopes are the ones contained in the interior of Δ_{mod} , while the singular slopes are those in the boundary of Δ_{mod} .

TITS BOUNDARY:

We begin by defining an *angular metric* on $\partial_\infty X$. For any ideal point $\xi \in \partial_\infty X$, at each $p \in X$ there is a unique ray ρ_ξ in the equivalence class ξ which emanates from p . Given $\xi, \eta \in \partial_\infty X$, we can define the angle at p between ξ and η , denoted $\angle_p(\xi, \eta)$, to be the angle between the initial vectors $H_\xi, H_\eta \in T_p X$ which determine the rays ρ_ξ and ρ_η . That is to say, $\rho_\xi(t) = \exp(tH_\xi)$ and $\rho_\eta(t) = \exp(tH_\eta)$ for all $t > 0$. The angular metric on $\partial_\infty X$ is then defined to be

$$\angle(\xi, \eta) = \sup_{p \in X} \angle_p(\xi, \eta).$$

The length metric associated to the angular metric is called the *Tits metric* on $\partial_\infty X$ and will be denoted by d_T .

We will use the metric on Δ_{mod} induced from the Tits metric when comparing the slopes of geodesics in X .

SPHERICAL BUILDING STRUCTURE:

Take a flat F in X and consider the boundary $\partial_\infty F$ of F at infinity. This can be viewed as a subspace of $\partial_\infty X$: each point in $\partial_\infty F$ is an equivalence class of asymptotic rays in F , which will be contained in an equivalence class in $\partial_\infty X$. By identifying the elements of $\partial_\infty F$ with the equivalence class they are contained in from $\partial_\infty X$, we can see $\partial_\infty F$ as a subspace of $\partial_\infty X$. Under the Tits metric, $\partial_\infty F$ is isometric to a euclidean sphere of dimension equal to the rank of X . These spheres will form the apartments of the spherical building.

Each flat is divided into Weyl chambers and walls. Each Weyl chamber or wall \mathcal{C} determines a subspace $\partial_\infty \mathcal{C}$ of $\partial_\infty F$. We can partition $\partial_\infty F$ into chambers and walls by taking the boundaries of the Weyl chambers and their walls in F .

For more details on why this gives a spherical building structure see either [BH99, Ch. II Thm 10.71] or [BGS85, Appendix 5].

3.1.4.c Families of parallel geodesics

We defined in Section 3.1.4.b the slope of a geodesic. Given a geodesic $c : \mathbb{R} \rightarrow X$, its slope ξ belongs to the set Δ_{mod} , which is the closure of a model chamber in the boundary of X . If the slope of c is contained in the interior of Δ_{mod} then it is regular and therefore contained in a unique maximal flat F . In fact, by the Flat Strip Theorem [BH99, Pg. 182], any geodesic parallel to c will also be contained in F . Hence F is equal to the subspace of X containing all geodesics parallel to c .

If c is a singular geodesic in X then it will be contained in a whole family of maximal flats. Again, using the Flat Strip Theorem, any geodesic parallel to c must be contained in one of these flats.

More formally, let $P(c)$ denote the subspace of X consisting of all geodesics that are parallel to c . So when c is regular, $P(c)$ is equal to the unique maximal flat F described above. While if c is singular $P(c)$ will be the union of (infinitely many) maximal flats. The structure of these sets is discussed in more detail in [Ebe96, §2.20].

Lemma 3.1.9. *G acts transitively on the set of subspaces of the form $P(\sigma)$, where σ varies over geodesics with the same slope.*

Proof. Let $\sigma, \tau : \mathbb{R} \rightarrow X$ be non-parallel geodesics of the same slope and let F_1, F_2 be any pair of maximal flats containing σ, τ respectively. By Proposition 3.1.6 there exists $g \in G$ such that $gF_1 = F_2$ and $g\sigma(0) = \tau(0)$. We now have two geodesics, $g\sigma$ and τ , which are contained in the same maximal flat and have the same slope. If the positive rays, that is $g\sigma[0, \infty)$ and $\tau[0, \infty)$, are in the same Weyl chamber, then having the same slope implies the geodesics must coincide, hence $g\sigma = \tau$. If they are not in the same Weyl chamber then we apply an element of the Weyl group of F_2 , which acts transitively on the Weyl chambers, so that they end up in the same Weyl chamber. An element of the Weyl group is a coset of the point-wise stabiliser of F_2 . So by choosing a representative of this coset we have $k \in K = G_{\tau(0)}$ such that $kg\sigma = \tau$. Finally, if σ' is any geodesic parallel to σ , then $kg\sigma'$ will be parallel to $kg\sigma = \tau$. Hence $kgP(\sigma) \subseteq P(\tau)$ and equality follows by symmetry. \square

3.1.4.d Iwasawa and Jordan decompositions

Recall that in Section 3.1.4.a we described the root-space decomposition of \mathfrak{g} with respect to a maximal abelian subspace \mathfrak{a} of \mathfrak{p} :

$$\mathfrak{g} = \mathfrak{g}_0 + \sum_{\lambda \in \Lambda} \mathfrak{g}_\lambda$$

where $\mathfrak{g}_\lambda = \{Y \in \mathfrak{g} \mid \text{ad}(H)Y = \lambda(H)Y \ \forall H \in \mathfrak{a}\}$. The set of roots Λ form a root system and the choice of a base Π for Λ gives us a set of positive roots Λ^+ . Let

$$\mathfrak{n} = \sum_{\lambda \in \Lambda^+} \mathfrak{g}_\lambda.$$

The *Iwasawa decomposition* of the Lie algebra is $\mathfrak{g} = \mathfrak{k} + \mathfrak{a} + \mathfrak{n}$ and the corresponding Iwasawa decomposition of G is

$$G = KAN$$

where $A = \exp(\mathfrak{a})$ and $N = \exp(\mathfrak{n})$. The decomposition of each element in G in this form is unique. The Iwasawa decomposition of G itself, on the other hand, is not unique and we can see geometrically how it is determined. The choice of positive roots Λ^+ corresponds to choosing a Weyl chamber in \mathfrak{a} , or equivalently a chamber in $\partial_\infty(Ap)$. Hence the Iwasawa decomposition of G is determined by a choice of basepoint $p \in X$ and a choice of a maximal chamber \mathcal{C} in $\partial_\infty X$. Once the chamber is chosen, A is determined by the unique maximal flat in X which contains p and is asymptotic to \mathcal{C} .

Given an isometry g of X , by using any Iwasawa decomposition of G , we can always write it as $g = kan$, where k is elliptic, a is real hyperbolic and n is unipotent. What's more, we can always choose p and \mathcal{C} so that the components in the corresponding Iwasawa decomposition, k, a and n , will commute. In this case, the decomposition $g = kan$ is called the (*complete*) *Jordan decomposition* of g .

3.1.4.e Parabolic subgroups

Let $\xi \in \partial_\infty X$. The subgroup of G which fixes ξ ,

$$G_\xi = \{g \in G \mid g\xi = \xi\}$$

is called the *parabolic subgroup* of G at ξ . Parabolic subgroups are discussed in more detail in [Ebe96, §2.17] — we will just recall here some basic properties they possess.

The structure of the family of parabolic subgroups reflects the Tits geometry of the boundary of X . In particular, given ξ and η in $\partial_\infty X$, these are contained in the same chamber or wall in $\partial_\infty X$ if and only if $G_\xi = G_\eta$. Furthermore, suppose that ξ is contained in a wall or chamber \mathcal{C}_ξ and η in \mathcal{C}_η . Then $G_\xi \subseteq G_\eta$ if and only if \mathcal{C}_η is contained in the closure of \mathcal{C}_ξ .

The parabolic subgroups act transitively on X and so G can be expressed as $G = KG_\xi$ for any maximal compact subgroup K . The Levi decomposition of a parabolic subgroup is $G_\xi = Z_\xi N_\xi$, where N_ξ is the unipotent radical of G_ξ and Z_ξ is a closed reductive subgroup.

We can describe the subgroup Z_ξ more accurately: let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be the Cartan decomposition at p and let $H \in \mathfrak{p}$ be such that the geodesic ray $\rho(t) = \exp(tH)p$ is asymptotic to ξ . Let $A_\xi = \exp(Z_{\mathfrak{g}}(H) \cap \mathfrak{p})$ and $K_\xi = K \cap Z_\xi$. Then $Z_\xi = K_\xi A_\xi$, leading to a decomposition of G as $G = K A_\xi N_\xi$, which Eberlein describes as a ‘‘generalised Iwasawa decomposition’’ since in the case when ξ is a regular ideal point this agrees with the Iwasawa decomposition of G as defined above.

3.1.5 The asymptotic cone of a symmetric space

We briefly discuss here asymptotic cones and a result of Kleiner and Leeb about the asymptotic cones of a symmetric space. Before we define an asymptotic cone we should discuss ultralimits and ultrafilters.

A *non-principal ultrafilter* ω on \mathbb{N} is a finitely additive probability measure on \mathbb{N} which takes values of either 0 or 1 and all finite sets have zero measure. Given a sequence $(a_n)_{n \in \mathbb{N}}$ of real numbers the *ultralimit* of this sequence is $a = \lim_\omega (a_n) \in \mathbb{R}$ which has the property that $\omega\{n \in \mathbb{N} \mid |a_n - a| < \varepsilon\} = 1$ for every $\varepsilon > 0$.

Let X be a metric space with metric d and let $p = (p_n)_{n \in \mathbb{N}}$ be a sequence of points in X . Let $(d_n)_{n \in \mathbb{N}}$ be a sequence in $(0, \infty)$ which diverges to infinity. Given a non-principal ultrafilter ω on \mathbb{N} we can define the *asymptotic cone* $\text{Cone}_\omega(X, p, d_n)$ to be the quotient space of sequences $(x_n)_{n \in \mathbb{N}}$ such that $\lim_\omega \frac{d(p_n, x_n)}{d_n} < \infty$ under the equivalence relation saying that two sequences (x_n) and (y_n) are equivalent if and only if $\lim_\omega \frac{d(x_n, y_n)}{d_n} = 0$. We can define a metric d_ω on the cone by setting $d_\omega((x_n), (y_n)) = \lim_\omega \frac{d(x_n, y_n)}{d_n}$. A point $x \in \text{Cone}_\omega(X, p, d_n)$ is said to be the ultralimit of a sequence of points (x_n) in X if (x_n) is a member of the equivalence class determining x .

As explained in Section 3.1.4.b, the ideal boundary of a symmetric space can be given a spherical building structure. If we take an asymptotic cone of a symmetric

space X then, as Kleiner and Leeb put it in [KL97], intuitively speaking we are “pulling the spherical building structure from infinity to the space of directions.” Spherical buildings have a useful property, which can be described as *rigidity of angles*. This means that given two points in a spherical building, their location inside their respective chambers determines a finite set of possible (angular) distances between them. When this property is pulled to the tangent space, by taking the asymptotic cone of our symmetric space, it transfers to a similar statement regarding the angle between intersecting geodesics.

We now recall the key aspects of the definition of a Euclidean building given by Kleiner and Leeb. We refer the reader to [KL97] for a more complete discussion. Note that in her PhD thesis [Par00] Parreau showed that Kleiner and Leeb’s axioms are equivalent to the definition of a building given by Tits.

A Euclidean Coxeter complex is a pair (E, W) where E is finite-dimensional Euclidean space and W is a group of isometries of E generated by reflections such that, if $\rho : \text{Isom}(E) \rightarrow \text{Isom}(\partial_\infty E)$ is the canonical map associating to every Euclidean isometry its rotational part, then $\rho(W)$ is finite. Define the anisotropy polyhedron $\Delta_E = (\partial_\infty E)/\rho(W)$. A metric space Y is a *Euclidean building* modelled on (E, W) if we are given a collection \mathcal{A} of isometric embeddings $\iota : E \rightarrow Y$, whose images are called *apartments*, and the following axioms are satisfied:

- EB1. We can define a map θ from the set of all directed geodesics in Y to Δ_E , the image of the map being the Δ_E -direction of the geodesic. The map θ satisfies the property that the difference between the Δ_E -direction of two geodesic segments starting at any point $x \in Y$ is less than or equal to the comparison angle between the two segments.
- EB2. For $\delta_1, \delta_2 \in \Delta_E$, define $D(\delta_1, \delta_2)$ to be the set of (angular) distances in $\partial_\infty E$ between points $\xi_1, \xi_2 \in \partial_\infty E$ such that $\rho(W)(\xi_i) = \delta_i$ for $i = 1, 2$. The set $D(\delta_1, \delta_2)$ will always be finite. For any three points $x, y, z \in Y$, the angle $\angle_x(y, z)$ between the geodesics $[x, y]$ and $[x, z]$ belongs to $D(\theta[x, y], \theta[x, z])$.
- EB3. Every geodesic, whether a segment, a ray or bi-infinite, is contained in an apartment.
- EB4. For $\iota_1, \iota_2 \in \mathcal{A}$ the composition $\iota_1^{-1} \circ \iota_2$ is the restriction of an isometry in W .

Given a maximal flat E in a symmetric space X we can determine a group of isometries M on E for which the pair (E, M) becomes a Euclidean Coxeter complex. By taking M to be the quotient of the set-wise stabiliser $\text{Stab}_G(E)$ by the point-wise stabiliser $\text{Fix}_G(E)$ we obtain a group which acts on E by isometries and maps Weyl chambers to Weyl chambers and walls to walls. This means that $\rho(M) \subset \text{Isom}(\partial_\infty E)$ is a finite group acting on the Euclidean sphere $\partial_\infty E$. The anisotropy polyhedron Δ_E will essentially be the closure of the boundary of a Weyl chamber, so will be isometric to Δ_{mod} as defined in Section 3.1.4.b. We therefore drop the Δ_E notation in favour of Δ_{mod} .

The key axiom which we will take advantage of is EB2 — described in [KL97] as “angle rigidity.” We explained above the intuition behind why the asymptotic cone of a symmetric space should satisfy axiom EB2. The remaining axioms are very similar to properties held by a symmetric space. The apartments in the asymptotic cone are the ultralimits of flats in X , so here we just need that the transitivity of the action of G on the set of all maximal flats in X gives a subset of G which can be interpreted as an atlas. We have already described in Section 3.1.4.b how to assign a Δ_{mod} -direction to geodesics in X . It is also not hard to see that every geodesic is contained in a maximal flat and since singular and regular geodesics are preserved by isometries the atlas maps are compatible with the Weyl group.

The following is the aforementioned result of Kleiner and Leeb:

Theorem 3.1.10 (Kleiner-Leeb, 1997). *Let X be a symmetric space of noncompact type. For any sequence of positive numbers (d_n) diverging to infinity and for any sequence of points $p = (p_n)$ in X the asymptotic cone $\text{Cone}_\omega(X, p, d_n)$ is a Euclidean building modelled on the Euclidean Coxeter complex (E, M) , where E is $\text{rank}(X)$ -dimensional Euclidean space and M is the quotient of the set-wise stabiliser $\text{Stab}_G(E)$ by the point-wise stabiliser $\text{Fix}_G(E)$.*

3.2 Real Hyperbolic Elements

3.2.1 Centralisers of real hyperbolic elements

Recall that for an isometry g of the symmetric space X the set of elements of minimal displacement by g is denoted by

$$\text{MIN}(g) = \left\{ x \in X \mid d_X(x, gx) = \inf_{y \in X} d_X(y, gy) \right\}.$$

We will see how when g is a hyperbolic element of G these sets provide us with a geometric interpretation of their centralisers.

The first part of the following is essentially a selection of results from Gromov's lecture notes [BGS85], many of which can also be found in [BH99], and as such most of it can be applied to more general situations where X is not necessarily a symmetric space but is a simply connected, complete, non-positively curved Riemannian manifold.

We say that g *translates* a geodesic $c : \mathbb{R} \rightarrow X$ if there exists non-zero $t \in \mathbb{R}$ such that $gc(s) = c(s + t)$ for all $s \in \mathbb{R}$. Suppose $q \in \text{MIN}(g)$ and let c be the bi-infinite geodesic passing through q and gq . We will show that c is translated by g . The geodesic segment $[q, gq]$ is mapped under g to a geodesic segment $[gq, g^2q]$. Hence if we take any point $x \in [q, gq]$ and consider the geodesic triangle with vertices x, gx, gp , see Figure 3.1, then we observe:

$$\begin{aligned} d_X(x, gx) &\leq d_X(x, gq) + d_X(gq, gx) \\ &= d_X(q, gq) \end{aligned}$$

But q was chosen so that $d_X(q, gq)$ was minimal, hence $d_X(x, gx) = d_X(q, gq)$ and $x \in \text{MIN}(g)$. Since geodesics in X are unique gq lies on the geodesic segment $[x, gx]$ and hence gx and g^2q lie on the geodesic c . We can then extend this result across the whole of c and we learn that the geodesic c is translated by g .

The above gives us the following:

Lemma 3.2.1. *Let g be a CAT(0)-hyperbolic isometry of X . If $q \in \text{MIN}(g)$ then the geodesic c passing through q and gq is translated by g .*

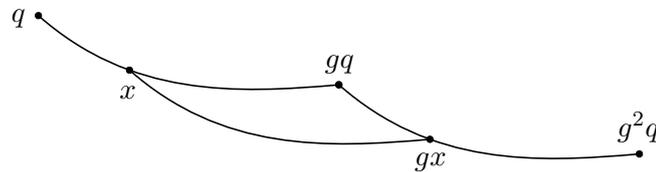


Figure 3.1: Given $q \in \text{MIN}(g)$ and the geodesic c through q and gq we can show $g^n q \in c$ for every integer n and every point inside the geodesic segment $[g^n q, g^{n+1} q]$ lies in $\text{MIN}(g)$.

We say that two unit speed geodesics $c_1 : \mathbb{R} \rightarrow X$ and $c_2 : \mathbb{R} \rightarrow X$ are *parallel* if they are at finite Hausdorff distance from each other.

Lemma 3.2.2. *Let g be a CAT(0)–hyperbolic isometry of X and let c be a geodesic in $\text{MIN}(g)$ translated by g . The set $\text{MIN}(g)$ consists precisely of all geodesics translated by g , which are all parallel to c .*

Proof. Let c' be a geodesic translated by g and let π be the orthogonal projection of X onto c . Let x be a point on c' and suppose $d_X(x, gx) = D + \varepsilon \geq D = \inf\{d_X(q, gq) \mid q \in X\}$. For every positive integer n , the distance from x to $g^n x$ is equal to $nD + n\varepsilon$. However, by travelling via the geodesic c we notice that $nD + n\varepsilon \leq 2d_X(x, c) + nD$. In particular $\varepsilon \leq \frac{2}{n}d_X(x, c)$ for every n and hence must be zero. Thus $c' \subseteq \text{MIN}(g)$.

Next we show that c' is parallel to c . Using the invariance of c' under g , any point $c(t)$ is within a bounded distance of at most $D/2$ of a point on c of the form $g^n c(0)$ for some integer n . Then we observe

$$d_X(g^n c(0), c') = d_X(c(0), g^{-n} c') = d_X(c(0), c') \leq d_X(c(0), c'(0)).$$

Hence $d_X(c(t), c') \leq D/2 + d_X(c(0), c'(0))$ and so c is contained in a neighbourhood of c' of radius $D/2 + d_X(c(0), c'(0))$. We obtain a similar result for c' and hence have that c and c' are at a finite Hausdorff distance from each other, and hence are parallel geodesics. \square

Eberlein [Ebe96, Prop 2.19.18] gives the following characterisation of real hyperbolic elements: g is real hyperbolic if and only if $g = \exp(H)$ where $H \in \mathfrak{p}$ for some Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. We use this in the proof of the following:

Lemma 3.2.3. *Let g be a real hyperbolic element of G and let c be any geodesic translated by g . Then*

$$\text{MIN}(g) = P(c)$$

where $P(c)$ is the subspace of X containing all geodesics parallel to c .

Proof. Suppose $g = \exp(H)$, for $H \in \mathfrak{p}$, where $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is the Cartan decomposition at a point $q \in X$. Let c be the geodesic $c(t) = \exp(tH)q$, for $t \in \mathbb{R}$. Lemma 3.2.2 gives us $\text{MIN}(g) \subseteq P(c)$. Let c' be any geodesic parallel to c . Then c and c' are contained in a flat in X . In particular, there exists a maximal abelian subspace \mathfrak{a} of \mathfrak{p} containing H such that $c, c' \subset \exp(\mathfrak{a})q$. Let $Y \in \mathfrak{a}$ be such that $\exp(Y)c(t) = c'(t)$ for every $t \in \mathbb{R}$. The geodesic c' is also translated by g since $\exp(Y)\exp(tH) = \exp(tH)\exp(Y)$ implies $c'(t) = \exp(tH)\exp(Y)q$ and therefore $gc'(t) = c'(t+1)$ for every $t \in \mathbb{R}$. Hence c' is contained in $\text{MIN}(g)$ and the Lemma follows. \square

Let g be a CAT(0)–hyperbolic isometry of X and suppose $\sigma : \mathbb{R} \rightarrow X$ is a geodesic translated by g , oriented so that $g\sigma(0) = \sigma(t)$ for some $t > 0$. We define the *slope* of g to be the Δ_{mod} –direction of σ corresponding to the positive direction, $\sigma(\infty)$. Of course parallel geodesics are asymptotic and so we get the same Δ_{mod} –direction regardless of which geodesic we consider. We say an isometry g is *regular semisimple* if it is CAT(0)–hyperbolic and its slope is regular in Δ_{mod} . When g is CAT(0)–hyperbolic but its slope is singular we say g is a *singular semisimple* isometry.

Lemma 3.2.4. *Let a be a regular hyperbolic element in G , contained in a maximal torus A . Then:*

- (1) *there exists a unique maximal flat F_a in X which is stabilised by a ;*
- (2) *for any $p \in F_a$, let K be the stabiliser of p , then $\text{Stab}(F_a) = Z_G(a)N_K(A)$, where $Z_G(a)$ is the centraliser of a in G and $N_K(A)$ is the normaliser of A in K .*

Proof. Suppose a is regular hyperbolic, translating a geodesic c . We can decompose a as $a = hk = kh$ where h is real hyperbolic and k is elliptic. The real hyperbolic component will translate c , hence $\text{MIN}(h)$ is equal to the maximal flat F_a . Consider the action of $k = h^{-1}a$ on F_a . It will fix c pointwise and hence must act trivially on F_a since any non-trivial elliptic isometry on F_a will permute the Weyl chambers and therefore cannot fix a regular geodesic. But this tells us that the action of a on F_a is precisely the same as that of h , thus $\text{MIN}(a) = F_a$.

Suppose that another flat F' is stabilised by a . Since a is not elliptic, it must act on F' hyperbolically, by translating some geodesic in F' . This geodesic must therefore be contained in $\text{MIN}(a) = F$. So F and F' intersect in a regular geodesic, hence must be equal. This proves (1).

Let p, K be as in (2) and note that $F_a = Ap$. Let $z \in Z_G(a)$ and $w \in N_K(A)$. Then $zwF_a = zF_a$ and $azF_a = zaF_a = zF_a$, so by (1) $zF_a = F_a$. Hence $Z_G(a)N_K(A) \subseteq \text{Stab}(F_a)$. Now let $g \in \text{Stab}(F_a)$. Then there exists $b \in A$ such that $gp = bp$. Thus $b^{-1}g = k \in K$ and stabilises F_a . Since $kAk^{-1}p = F_a$ we see that kAk^{-1} is a maximal torus stabilising the flat F_a . Such a torus is unique, so $k \in N_K(A)$. Hence $g = bk \in Z_G(a)N_K(A)$ and (2) holds. \square

We can see that the orbit of the centraliser of a regular hyperbolic element will be a maximal flat. When we take a singular real hyperbolic element instead, then the orbit of its centraliser will contain many maximal flats.

Lemma 3.2.5. *Let a be a singular real hyperbolic element in G . Then:*

- (1) *the subspace $\text{MIN}(a)$ is precisely the set of all geodesics translated by a , which is the Riemannian product of a Euclidean space and a symmetric space of non-compact type; and*
- (2) *for any $p \in \text{MIN}(a)$, let K be the stabiliser of p , then*

$$Z_G(a) \subseteq \text{Stab}(\text{MIN}(a)) \subseteq Z_G(a)K$$

where $Z_G(a)$ is the centraliser of a in G .

Proof. The first part of assertion (1) follows from Lemma 3.2.3, while for a proof of the second part we refer the reader to [Ebe96, 2.11.4].

For (2), take $b \in Z_G(a)$ and let c' be any geodesic translated by a . Then $bc' = bac' = abc'$ implies that bc' is translated by a , hence is contained in $\text{MIN}(a)$. Now let g stabilise $\text{MIN}(a)$. Let c_1, c_2 be the geodesics translated by a such that $c_1(0) = p$ and $c_2(0) = gp$ respectively. Since c_1 and c_2 are parallel they are contained in a common flat $F = Ap$, where A is a maximal abelian Lie subgroup of G which contains a . There exists $b \in A$ such that $bc_2 = c_1$ and $bgp = p$. Hence $bg \in K$ and in particular $g \in Z_G(a)K$. \square

3.2.2 Bounding the size of a conjugator from G

The aim of this section is to obtain a control on the length of a conjugator between two real hyperbolic elements a, b in G . The control will be linear, but the constant in the upper bound will depend on the slope of a and b , and hence their conjugacy class. Our method of demonstrating this is to first show that a conjugator corresponds to an isometry that maps $\text{MIN}(a)$ to $\text{MIN}(b)$. Then, by obtaining a control on the distance from an arbitrary basepoint p to $\text{MIN}(a)$ in terms of $d_X(p, ap)$, we can obtain a control on the length of a conjugator from G .

3.2.2.a Relating conjugators to maps between flats

Here we show why we can obtain a short conjugator by understanding the distance to $\text{MIN}(a)$ and $\text{MIN}(b)$. In the following let π_a be the orthogonal projection of X onto $\text{MIN}(a)$ and for $x \in X$ let $G_q = \{k \in G \mid gq = q\}$.

Proposition 3.2.6. *Let a, b be conjugate semisimple elements in G . Then:*

(1) for $g \in G$, if $gag^{-1} = b$ then $g\text{MIN}(a) = \text{MIN}(b)$;

(2) for $h \in G$, if $h\text{MIN}(a) = \text{MIN}(b)$ then there exists $x \in G_{\pi_a(p)}$ such that $(hx)a(hx)^{-1} = b$.

Proof. For (1), let c be any geodesic stabilised by a . Since $g = bga^{-1}$ we see that the geodesic gc is translated by b and so is contained in $\text{MIN}(b)$. Hence $g\text{MIN}(a) \subseteq \text{MIN}(b)$. Similarly we get $g^{-1}\text{MIN}(b) \subseteq \text{MIN}(a)$ and (1) is proved.

Next suppose that $g, h \in G$ are such that $gag^{-1} = b$ and $h\text{MIN}(a) = \text{MIN}(b)$. By the first part we observe $g^{-1}h\text{MIN}(a) = \text{MIN}(a)$, so $g^{-1}h \in \text{Stab}(\text{MIN}(a)) \subseteq Z_G(a)K$, with the latter relationship coming from Lemma 3.2.5, where we take $K = G_{\pi_a(p)}$. Then there exists $x \in K$ such that $g^{-1}hx \in Z_G(a)$. This implies $(g^{-1}hx)a(g^{-1}hx)^{-1} = a$ and so $(hx)a(hx)^{-1} = gag^{-1} = b$, proving (2). \square

If a and b are regular semisimple elements in G then they each stabilise a unique maximal flat in X . Suppose a, b stabilise maximal flats F_a and F_b respectively. By taking our basepoint p to be in F_a we can build a quadrilateral which has two vertices in F_a and two vertices in F_b , as in Figure 3.2. In light of 3.2.6, the aim is to find an element $g \in G$ of a controlled size which maps the flat F_a to F_b . We will then obtain a conjugator of controlled size.

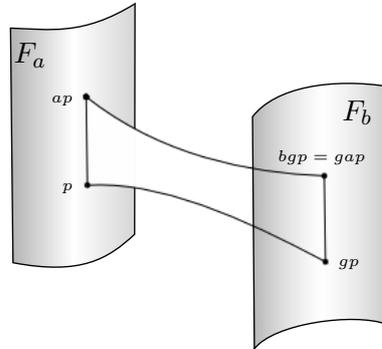


Figure 3.2: A quadrilateral in X demonstrating the conjugacy of a and b in G .

Proposition 3.2.7. *Let p be our basepoint in X . Suppose for all semisimple a in G we can find a constant $\ell(a)$ such that:*

$$d_X(p, \text{MIN}(a)) \leq \ell(a)d_X(p, ap).$$

Then for a, b conjugate hyperbolic elements in G there exists a conjugator $g \in G$ such that:

$$d_X(p, gp) \leq \ell(a)d_X(p, ap) + \ell(b)(p, bp).$$

Proof. Choose points $p_a \in \text{MIN}(a)$ and $p_b \in \text{MIN}(b)$ such that $d_X(p, p_a) \leq \ell(a)d_X(p, ap)$ and $d_X(p, p_b) \leq \ell(b)d_X(p, bp)$. Let $g_1 \in G$ be such that $g_1\text{MIN}(a) = \text{MIN}(b)$ and $g_1p_a = p_b$. By Lemma 3.2.6 there exists x in G_{p_b} such that $(xg_1)a(xg_1)^{-1} = b$. Let $g = xg_1$.

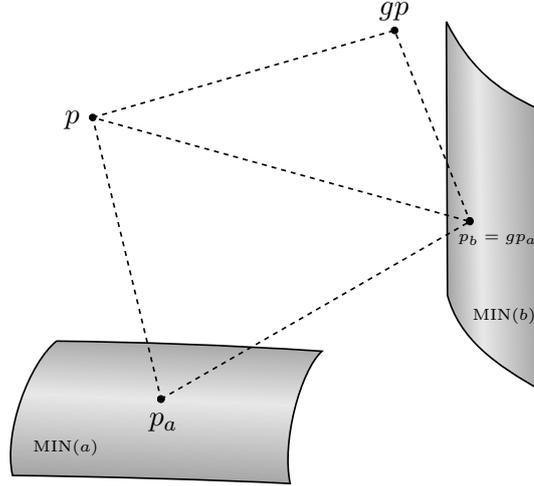


Figure 3.3: Obtaining an upper bound on $d_X(p, gp)$.

To finish the proof we need to check that we have the required upper bound on $d_X(p, gp)$. By the triangle inequality:

$$\begin{aligned}
 d_X(p, gp) &\leq d_X(p, p_b) + d_X(p_b, gp) \\
 &= d_X(p, p_b) + d_X(gp_a, gp) \\
 &= d_X(p, p_b) + d_X(p, p_a) \\
 &\leq \ell(a)d_X(p, ap) + \ell(b)d_X(p, bp).
 \end{aligned}$$

□

3.2.2.b Bounding the distance to a flat

In Proposition 3.2.7 we saw how finding some constant $\ell(a)$ such that $d_X(p, \text{MIN}(a)) \leq \ell(a)d_X(p, ap)$ helps us to control the size of a conjugator in G between a and another element b . The aim of this section is to find such constants for the case when a is real hyperbolic. In fact, we will determine a value for $\ell(a)$ which depends only on the slope of a . Note that we have the following:

Lemma 3.2.8. *If a is conjugate to b then a and b have the same slope.*

Proof. Let a have slope $\xi \in \Delta_{\text{mod}}$. This means that the geodesic segment $[p, ap]$, where p is a point in $\text{MIN}(a)$, has Δ_{mod} -direction ξ . Suppose $b = gag^{-1}$ for some $g \in G$. Then g maps the bi-infinite geodesic through p and ap to a bi-infinite geodesic through gp and $gap = bgp$. This geodesic is translated by b , so the slope of b is the Δ_{mod} -direction of the geodesic segment $[gp, bgp] = g[p, ap]$. Since the Δ_{mod} -direction is defined to be G -invariant we have that the slope of b is ξ . \square

In order to determine the value of $\ell(a)$ we will use an asymptotic cone of X , which, by a result of Kleiner and Leeb [KL97], is a Euclidean building. It is helpful therefore to first determine the corresponding value in a Euclidean building. This will then be useful to find the value for symmetric spaces. Recall Lemma 3.2.3 which asserted, for a real hyperbolic element a , that $\text{MIN}(a) = P(\sigma)$, where σ is any geodesic translated by a and $P(\sigma)$ is the subspace of X consisting of all geodesics parallel to σ . We are therefore interested in the distance to similarly defined subspaces of a Euclidean building. Also recall that, for a Euclidean building Y , the quotient of $\partial_\infty Y$ by the group of isometries of Y is denoted by Δ_{mod} , and $\theta : \partial_\infty Y \rightarrow \Delta_{\text{mod}}$ is the natural map.

Lemma 3.2.9 (see [HKM10, Lemma 5.1]). *Let Y be a Euclidean building, $\delta \in \partial_\infty Y$, c be a geodesic in Y with one end asymptotic to δ and E be the subset of Y consisting of all geodesics parallel to c . Then for any point $p \in Y$ the geodesic ray emanating from p which is asymptotic to δ enters the set E in finite distance.*

REMARK: The proof of this Lemma given in [HKM10, Lemma 5.1] only covers algebraic Euclidean buildings. It is worth noting that by [KT04] the asymptotic cone of a symmetric space is an algebraic Euclidean building, so their proof applies to the buildings which we are concerned with.

Proposition 3.2.10. *Let a be an isometry of a Euclidean building Y such that a translates all geodesics parallel to a geodesic c and let E be the subset of Y containing all these geodesics. Suppose the ray $c[0, \infty)$ satisfies $ac[0, \infty) \subset c[0, \infty)$ and is asymptotic to $\delta \in \partial_\infty E$, where $\theta(\delta) = \xi \in \Delta_{\text{mod}}$. Then there exists a constant ℓ_ξ such that for any basepoint p in Y the following holds:*

$$d(p, E) \leq \ell_\xi d(p, ap).$$

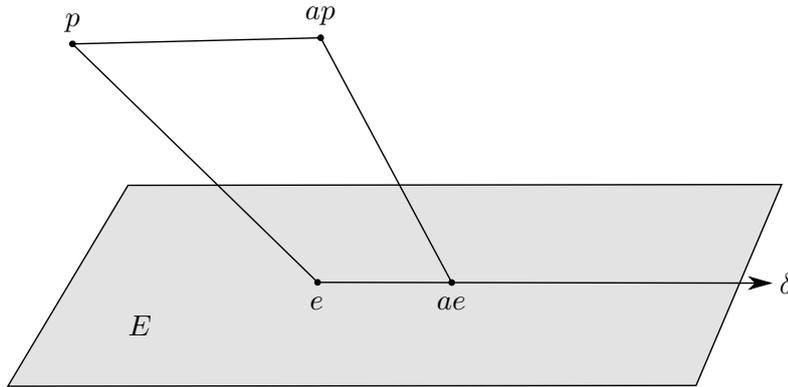


Figure 3.4: The angle $\angle_{ae}(e, ap)$ is rigid, that is the angle is contained in a finite set which is determined by ξ . This leads to a bound on the distance from p to E .

Proof. Consider the ray emanating from p which represents δ . By Lemma 3.2.9 this ray enters E . Let e be the first point along this ray such that $e \in E$. By design ae also lies on this ray. Now translate the ray by a . What we get is a geodesic triangle in Y , as seen in Figure 3.4, with vertices p, ae, ap . By the rigidity of angles in Y , $\angle_{ae}(p, ap)$ belongs to the finite set $D(\xi)$. Let ϕ be minimal in this set. Then since Y is a CAT(0) space we have that $d(p, ap) \geq d(ap, ae) \sin \phi$. Hence we put $\ell_\xi = \frac{1}{\sin \phi}$ and the proposition holds. \square

Lemma 3.2.11. *Let the Tits building structure on $\partial_\infty X$ have anisotropy polyhedron Δ_{mod} . Fix a basepoint p in X . Then for each element $\xi \in \Delta_{mod}$ there exists positive constants ℓ_ξ and d_ξ such that for each $a \in G$ which is real hyperbolic of slope ξ and such that $d_X(p, ap) > d_\xi$ the following holds:*

$$d_X(p, \text{MIN}(a)) \leq 2\ell_\xi d_X(p, ap).$$

Proof [regular slopes]. The proof which follows applies to the case when ξ is regular. The proof for singular slopes is analogous, with slight modifications which are described at the end of this proof.

First note that the constant ℓ_ξ will be the same constant that we obtained in Proposition 3.2.10.

We proceed by contradiction, supposing the statement is false. We then obtain a sequence of regular semisimple elements a_n in G , each of slope ξ , such that $d_X(p, a_n p)$ diverges to infinity and which satisfies:

$$d_X(p, F_n) > 2\ell_\xi d_X(p, a_n p) \tag{3.1}$$

where F_n is the unique maximal flat stabilised by a_n . Write $d_n := d_X(p, a_n p)$ and $D_n := d_X(p, F_n)$. Let $\pi_n : X \rightarrow F_n$ be the orthogonal projection onto F_n and let t_n be the translation length of a_n , that is $t_n := d_X(\pi_n(p), a_n \pi_n(p))$. We split the proof into three parts depending on the limits of the ratios t_n/D_n and d_n/D_n .

CASE 1: $\lim_{\omega}(t_n/D_n) \neq 0 \neq \lim_{\omega}(d_n/D_n)$.

In the first part of the proof we build a Euclidean building and use Proposition 3.2.10 to obtain a contradiction under the assumption that t_n/D_n does not converge to zero.

Since d_n diverges to infinity, it follows from (3.1) that D_n does too. Hence we pick a non-principal ultrafilter ω and consider the asymptotic cone $Y = \text{Cone}_{\omega}(X, p, D_n)$. By choice of scalars it follows that the ultralimit E of the sequence of flats F_n is contained in Y and lies a distance 1 away from the point p (when we view p as an element of the cone Y).

Define the map $g : Y \rightarrow Y$ by sending $(x_n) \in Y$ to $(a_n x_n)$. To check it is well-defined on Y we need only observe that it moves p a bounded distance:

$$\begin{aligned} d_{\omega}(p, gp) &= \lim_{\omega} \left(\frac{d_X(p, a_n p)}{D_n} \right) \\ &= \lim_{\omega} \left(\frac{d_n}{D_n} \right) \\ &\leq \lim_{\omega} \left(\frac{1}{2\ell_{\xi}} \right) \\ &= \frac{1}{2\ell_{\xi}}. \end{aligned}$$

Furthermore, since a_n acts on X by isometries for each n it follows that g acts on Y by isometries.

By assumption t_n/D_n does not converge to zero, hence $d_{\omega}(\pi(p), g\pi(p)) > 0$. It implies, since $t_n \leq d_n$, that $d_X(p, gp) > 0$. Under these conditions we may apply Proposition 3.2.10, since g acts on E by translating along geodesics towards a boundary point ξ . This gives us the following contradiction:

$$1 = d_{\omega}(p, E) \leq \ell_{\xi} d_{\omega}(p, gp) \leq \frac{\ell_{\xi}}{2\ell_{\xi}} = \frac{1}{2}.$$

CASE 2: $\lim_{\omega}(t_n/D_n) = 0 = \lim_{\omega}(d_n/D_n)$.

We must therefore assume the ω -limit of t_n/D_n is zero. We assume this for the second part of the proof and we also assume that the ω -limit of d_n/D_n is zero. We will build

a sequence of quadrilaterals and take their Hausdorff limit. The limiting quadrilateral will be flat and intersecting a flat F only in one edge, along a regular geodesic. This will give the contradiction.

Fix a flat F in X and a point $q \in F$. For each n consider an isometry $g_n \in G$ which sends $\pi_n(p)$ to q and F_n to F . The first thing to note is that each geodesic segment $[\pi_n(p), a_n\pi_n(p)]$ is mapped to a geodesic segment $T := [q, g_n a_n g_n^{-1}q]$ in F of Δ_{mod} -direction ξ . Consider the projection $\tau : [g_n p, g_n a_n p] \rightarrow T$. For a fixed constant h and for large enough n we may pick a subsegment S of T of length h such that the pre-image $\tau^{-1}(S)$ in $[g_n p, g_n a_n p]$ has length at most $d_n h/t_n$.

Label points $b_n^{(1)}, b_n^{(2)}$ on the geodesic $[g_n p, a_n g_n p]$ which are mapped under τ to each end of the segment S , see Figure 3.5. Let $L_n^{(i)} := d_X(b_n^{(i)}, T)$ and without loss of generality assume $L_n^{(1)} \leq L_n^{(2)}$. Observe:

$$D_n = d_X(g_n p, T) \leq d_X(g_n p, b_n^{(1)}) + d_X(b_n^{(1)}, T)$$

and replacing $g_n p$ by $g_n a_n p$ if necessary we get:

$$L_n^{(1)} \geq D_n - \frac{d_n}{2}.$$

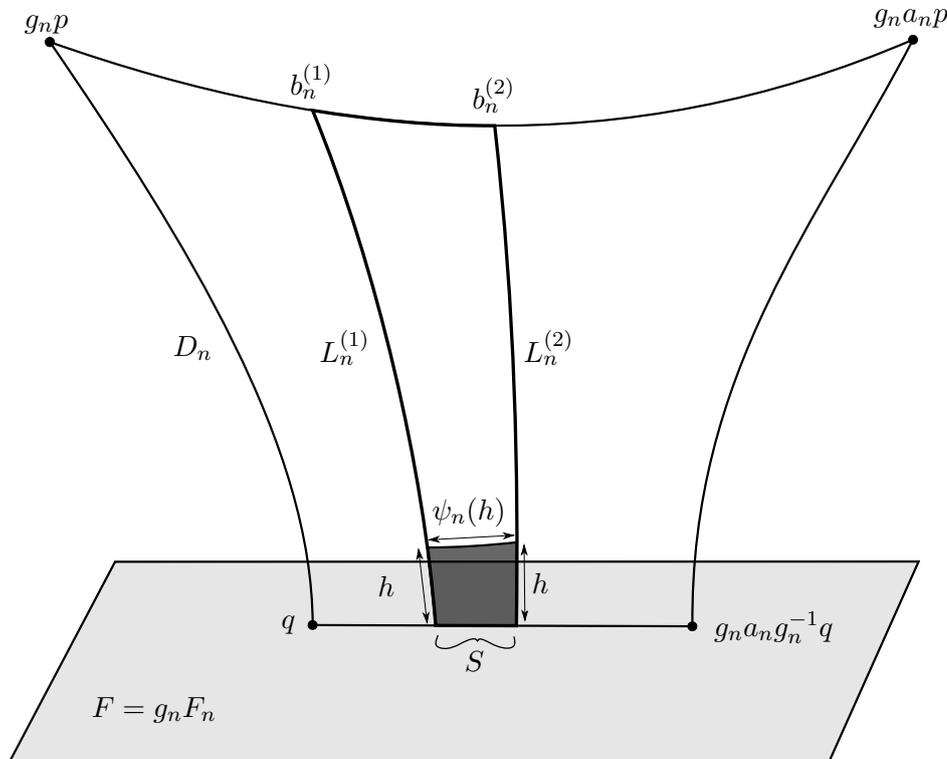


Figure 3.5: The quadrilateral Q_n with side S .

Using our hypothesis we therefore get $L_n^{(1)} > (2\ell_\xi - \frac{1}{2}) d_n$. So, referring back to the value of ℓ_ξ obtained in Proposition 3.2.10, since $2\ell_\xi - \frac{1}{2} > 0$ for any choice of ξ , we get that $L_n^{(1)}$ diverges to infinity.

Now define quadrilaterals Q_n as follows. We take one edge to be the segment S and the two adjacent edges are those subsegments of $[b_n^{(i)}, \tau(b_n^{(i)})]$ of length h which include the points $\tau(b_n^{(i)})$, for $i = 1, 2$. The quadrilateral Q_n has three sides of length h . Let $\psi_n(h)$ be the length of the fourth side. Define a map $\hat{\psi}_n : [0, L_n^{(1)}] \rightarrow \mathbb{R}$ measuring the distance across the flat rhombus (see Figure 3.6).

Using the CAT(0) property of X we see that

$$\begin{aligned} h \leq \psi_n(h) &\leq \hat{\psi}_n(h) \\ &= \left(\frac{L_n^{(1)} - h}{L_n^{(1)}} \right) h + \frac{h^2 d_n}{L_n^{(1)} t_n} \\ &= h - \frac{h^2}{L_n^{(1)}} + h^2 \left(\frac{d_n}{D_n t_n} \right) \left(\frac{D_n}{L_n^{(1)}} \right) \end{aligned}$$

Above we showed that $L_n^{(1)} \geq D_n - \frac{d_n}{2}$. We can use this to show that $D_n/L_n^{(1)}$ is bounded above by $\left(1 - \frac{1}{4\ell_\xi}\right)^{-1}$. This is therefore enough, since we have the assumption that d_n/D_n converges to zero, to conclude that $\psi_n(h)$ converges to h as n tends to infinity. After translating the quadrilaterals Q_n along the geodesics to q we can find the Hausdorff limit Q of a convergent subsequence of the quadrilaterals Q_n . The quadrilateral Q will have four sides with length h , two right-angles and hence must be a flat square. But Q intersects F only through the regular geodesic segment T . This gives a contradiction.

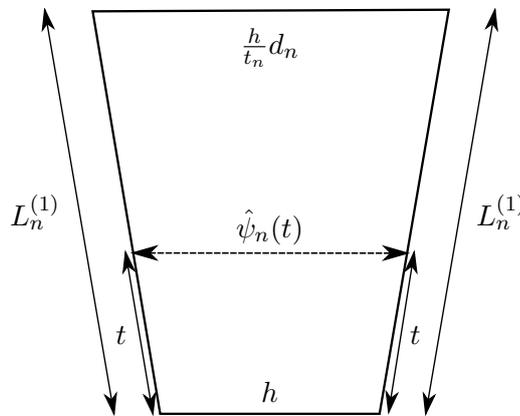


Figure 3.6: Defining $\hat{\psi}_n(t)$.

CASE 3: $\lim_{\omega}(t_n/D_n) = 0 \neq \lim_{\omega}(d_n/D_n)$.

We conclude by combining both of the above arguments into one in order to obtain a contradiction when t_n/D_n converges to zero but d_n/D_n does not. We start by looking at the situation inside the Euclidean building Y that we built in case 1. It is constructed so that $d_{\omega}(p, E) = 1$, but in what follows we will show that in this case we would have the contradiction $d_{\omega}(p, E) < 1$.

In Y , take the two geodesic rays asymptotic to ξ which begin at p and at gp respectively. Note that the second ray is the image of the first under g . Also recall that both rays will enter the apartment E . Since g fixes E pointwise we see that the two rays must come together at some point y . In particular either y is the point where the rays enter E or it is not in E . We will show that $d_{\omega}(y, E) \leq (4\ell_{\xi})^{-1}$.

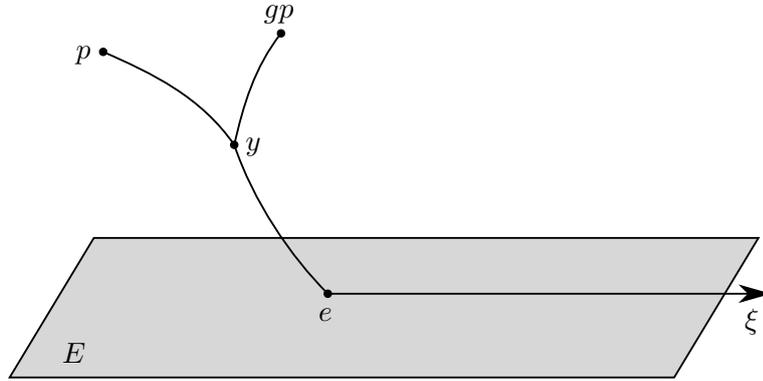


Figure 3.7: Two geodesic rays asymptotic to ξ , entering the apartment E at a point e and merging at the point y .

Suppose that $d_{\omega}(y, E) > (4\ell_{\xi})^{-1}$ and let (y_n) be a sequence of points in X which represent y in Y . Then there exists ε such that $(4\ell_{\xi})^{-1} < \varepsilon < d_{\omega}(y, E)$ and $\omega\{n \in \mathbb{N} \mid d_X(y_n, F_n) \geq \varepsilon D_n\} = 1$. We now proceed as in case 2, for each n applying the isometry g_n and constructing quadrilaterals Q_n with one edge in the fixed flat F . As before we pick a segment S of length h from $[q, g_n a_n g_n^{-1} q]$ whose pre-image under the projection onto $[q, g_n a_n g_n^{-1} q]$ intersects $[g_n y_n, g_n a_n y_n]$ in a segment of length at least $\frac{h}{t_n} d_X(y_n, a_n y_n)$. Let $L_n^{(i)}$ be the distances between the corresponding end-points of these subsegments and suppose $L_n^{(1)} \leq L_n^{(2)}$. Then in order to proceed as before we need that the function:

$$\hat{\psi}_n(t) = \left(\frac{L_n^{(1)} - t}{L_n^{(1)}} \right) h + \frac{t}{L_n^{(1)}} \frac{h}{t_n} d_X(y_n, a_n y_n)$$

converges to h when we put $t = h$. We need to check two things: firstly that $L_n^{(1)}$ diverges to infinity and secondly that $\frac{d_X(y_n, a_n y_n)}{L_n^{(1)}}$ converges to zero. For the former we note that for all but finitely many $n \in \mathbb{N}$ we have the following:

$$\begin{aligned} L_n^{(1)} &\geq d_X(y_n, F_n) - \frac{1}{2}d_X(y_n, a_n y_n) \\ &> \left(2\ell_\xi \varepsilon - \frac{1}{2}\right) d_X(y_n, a_n y_n) \end{aligned}$$

If $d_X(y_n, a_n y_n)$ is bounded we use the first line to show $L_n^{(1)}$ is unbounded. Otherwise we use the second line, recalling that $\varepsilon > \frac{1}{4\ell_\xi}$. To prove that $\frac{d_X(y_n, a_n y_n)}{L_n^{(1)}}$ converges to zero we first check that $\frac{D_n}{L_n^{(1)}}$ is bounded. This is so because for all but finitely many n we have the following:

$$\begin{aligned} L_n^{(1)} &\geq d_X(y_n, F_n) - \frac{1}{2}d_X(y_n, a_n y_n) \\ &> \varepsilon D_n - \frac{1}{2}d_n \\ \frac{L_n^{(1)}}{D_n} &> \varepsilon - \frac{d_n}{2D_n} \\ \frac{D_n}{L_n^{(1)}} &< \left(\varepsilon - \frac{1}{4\ell_\xi}\right)^{-1} \end{aligned}$$

Hence we see that $\frac{d_X(y_n, a_n y_n)}{L_n^{(1)}} = \frac{d_X(y_n, a_n y_n)}{D_n} \frac{D_n}{L_n^{(1)}}$ converges to zero by our choice of y_n . Then as before, after translating the quadrilaterals Q_n so they each have a vertex at q , we take the Hausdorff limit of a convergent subsequence of these quadrilaterals and obtain a flat quadrilateral which intersects F only through a regular geodesic segment. Here we have our contradiction and conclude that $d_\omega(y, E) \leq (4\ell_\xi)^{-1}$.

To finish the argument we look at the triangle in Y with vertices p, gp, y . In a similar manner to the proof of Proposition 3.2.10 we use the fact that Y is a CAT(0) space to get $d_\omega(p, gp) \geq d_\omega(p, y)\ell_\xi^{-1}$, recalling that ℓ_ξ^{-1} is the sine of the minimal angle in the finite set $D(\xi)$ of possible angles between geodesics of Δ_{mod} -direction ξ .

Therefore we have the following:

$$\begin{aligned} d_\omega(p, E) &\leq d_\omega(p, y) + d_\omega(y, E) \\ &\leq \ell_\xi d_\omega(p, gp) + \frac{1}{4\ell_\xi} \\ &\leq \ell_\xi \frac{1}{2\ell_\xi} + \frac{1}{4\ell_\xi} \\ &= \frac{1}{2} + \frac{1}{4\ell_\xi} < 1 \end{aligned}$$

However, we also have $d_\omega(p, E) = \lim_\omega \left(\frac{D_n}{D_n} \right) = 1$, thus giving the contradiction and proving the Lemma. \square

Proof [singular slopes]. In order to modify the above proof to work for singular directions we need to make the following adjustments. First we replace the flats F_n by $\text{MIN}(a_n)$. Case 1 continues as above with no change. For cases 2 and 3, the contradiction we obtain will be similar. Instead of using a fixed flat F , we use a fixed subspace $P(c)$ which consists of a family of geodesics parallel to some geodesic c of slope ξ , for example we may take $M = \text{MIN}(a_1)$ and c any geodesic translated by a_1 . By Lemma 3.1.9 we know there exists $g_n \in G$ which sends $\text{MIN}(a_n)$ to M . From the proof of Lemma 3.1.9 it is also clear that g_n can be chosen so it sends a geodesic translated by a_n to a geodesic translated by a_1 . Furthermore, if we fix a point q in M , as we did in the above proof, then we can choose g_n so it sends $\pi_n(p)$ to q . Once we have this, we can find a flat quadrilateral Q in the same way as above, but it will intersect M only in one side, which is a geodesic segment of slope ξ . The opposite edge of Q will be a segment of a geodesic parallel to c , hence should be contained in M , but it is not. \square

In light of Lemma 3.2.11 we can put $\ell(a) = \ell(b) = 2\ell_\xi$, where ξ is the slope of a and b , into Proposition 3.2.7 to get the following:

Theorem 3.2.12. *Let ℓ_ξ and d_ξ be the constants from Lemma 3.2.11. Suppose a and b are conjugate real hyperbolic elements in G with slope $\xi \in \Delta_{\text{mod}}$ and such that $d_X(p, ap), d_X(p, bp) \geq d_\xi$. Then there exists a conjugator $g \in G$ such that:*

$$d_X(p, gp) \leq 2\ell_\xi(d_X(p, ap) + d_X(p, bp)).$$

The constant $2\ell_\xi$ that we have obtained will depend on the slope of a and b , and hence on the conjugacy class. However it is important to note that it is independent of the basepoint p that was chosen.

When we restrict our attention to a lattice Γ in G , Theorem 3.2.12 will apply to all but finitely many real hyperbolic elements of Γ . This leads to the following result:

Corollary 3.2.13. *Let Γ be a lattice in G . Then for each $\xi \in \Delta_{\text{mod}}$, there exists a constant L_ξ such that two elements $a, b \in \Gamma$ are conjugate in G if and only if there exists a conjugator $g \in G$ such that*

$$d_X(p, gp) \leq L_\xi(d_X(p, ap) + d_X(p, bp)).$$

We now offer a couple of corollaries to Theorem 3.2.12 which shed a little more light on the nature of short conjugators between real hyperbolic elements in a semisimple real Lie group. For a semisimple element a of G , the translation length of a is

$$\tau(a) = \inf\{d_X(x, ax) \mid x \in X\}.$$

We can reformulate Theorem 3.2.12 so that it applies to all real hyperbolic elements in G , provided their translation length isn't too small.

Corollary 3.2.14. *Let $0 < \varepsilon \leq d_\xi$ and suppose that a and b are real hyperbolic elements of G of slope ξ and with translation lengths $\tau(a), \tau(b) \geq \varepsilon$. Then a and b are conjugate in G if and only if there exists a conjugator $g \in G$ such that*

$$d_X(p, gp) \leq 2\ell_\xi \left(\frac{d_\xi}{\varepsilon} + 1 \right) (d_X(p, ap) + d_X(p, bp)).$$

Proof. Since a is real hyperbolic, $\tau(a) > 0$ and $\tau(a^k) = k\tau(a)$ for all $k \in \mathbb{N}$. Let k be the maximal positive integer such that $k\varepsilon < d_\xi$. In particular, maximality of k implies that $d_X(p, a^{k+1}p), d_X(p, b^{k+1}p) \geq d_\xi$ and we are able to apply Lemma 3.2.11, concluding that

$$d_X(p, \text{MIN}(a)) \leq 2\ell_\xi(k+1)d_X(p, ap), \quad d_X(p, \text{MIN}(b)) \leq 2\ell_\xi(k+1)d_X(p, bp).$$

We can then apply Lemma 3.2.7, taking $\ell(a) = \ell(b) = 2\ell_\xi(\frac{d_\xi}{\varepsilon} + 1)$. This gives the upper bound on the length of a conjugator g as required. \square

We complete this section with the following consequence of the above work. It says that if we restrict ourselves to looking at the majority of regular hyperbolic elements — that is, those with not too small translation length and of a slope which is not too close to being singular — then we can obtain a linear bound on the length of short conjugators.

Corollary 3.2.15. *For every $\varepsilon_1, \varepsilon_2 > 0$ there exists $\kappa = \kappa(\varepsilon_1, \varepsilon_2)$ with the following property: assume that a and b are conjugate hyperbolic elements with translation lengths $\tau(a), \tau(b) \geq \varepsilon_1$ and slope $\xi \in \Delta_{\text{mod}}$ such that the spherical distance from ξ to $\partial\Delta_{\text{mod}}$ is at least ε_2 . Then there exists a conjugator $g \in G$ such that:*

$$d_X(p, gp) \leq \kappa(d_X(p, ap) + d_X(p, bp)).$$

It is worth noting that while we have a precise expression for the constant ℓ_ξ in terms of the slope ξ , we have no grasp on the value taken by d_ξ .

3.2.3 Dependence on slope

The aim here is to show that any constant satisfying the linear relationship in Theorem 3.2.12 must depend on the common slope of a and b . To do this, we first show there is not a uniform constant $\ell > 0$ such that for all regular hyperbolic elements $a \in G$ the following holds:

$$d_X(p, \text{MIN}(a)) \leq \ell d_X(p, ap) \quad (3.2)$$

where p is an arbitrary basepoint in X . This will imply that the constants $\ell(a)$ required for Proposition 3.2.7 will have to depend somehow on a .

To do this, we will construct a sequence of regular hyperbolic elements which contradict the existence of such an ℓ . The sequence will converge to a singular hyperbolic element, agreeing with the intuition that the constant ℓ_ξ of Lemma 3.2.11 diverges to infinity as the slope ξ converges to a singular direction.

We first note that if (3.2) is true for a point $p \in X$ then it is true for every point $q \in X$. Indeed let $g \in G$ be any isometry such that $gp = q$. Then firstly:

$$\begin{aligned} d_X(q, \text{MIN}(a)) &= d_X(gp, \text{MIN}(a)) \\ &= d_X(p, g^{-1}\text{MIN}(a)) \\ &= d_X(p, \text{MIN}(g^{-1}ag)) \end{aligned}$$

and secondly:

$$\begin{aligned} d_X(q, aq) &= d_X(gp, agp) \\ &= d_X(p, g^{-1}agp). \end{aligned}$$

But since we assume (3.2) to be true for all hyperbolic elements it follows that it is true for $g^{-1}ag$ and thus:

$$d_X(q, \text{MIN}(a)) \leq \ell d_X(q, aq).$$

Fix a pair of distinct flats F, F' whose intersection is non-trivial and of dimension at least one. Let p_0 be a point in their intersection and let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be the corresponding Cartan decomposition. Let $\mathfrak{a}, \mathfrak{a}'$ be the maximal abelian subspaces of \mathfrak{p} such that:

$$F = \exp(\mathfrak{a})p_0$$

$$F' = \exp(\mathfrak{a}')p_0$$

Furthermore suppose $H \in \mathfrak{a} \cap \mathfrak{a}'$ has length 1 and $Y \in \mathfrak{a}' \setminus \mathfrak{a}$ is orthogonal to H and is also of length 1.

Let (H_n) be a sequence of regular unit vectors in $\mathfrak{a} \setminus \mathfrak{a}'$ which converge to H . Define $a_0 := \exp(H) \in G$ and $a_n := \exp(H_n) \in G$ for $n \in \mathbb{N}$ (see Figure 3.8).

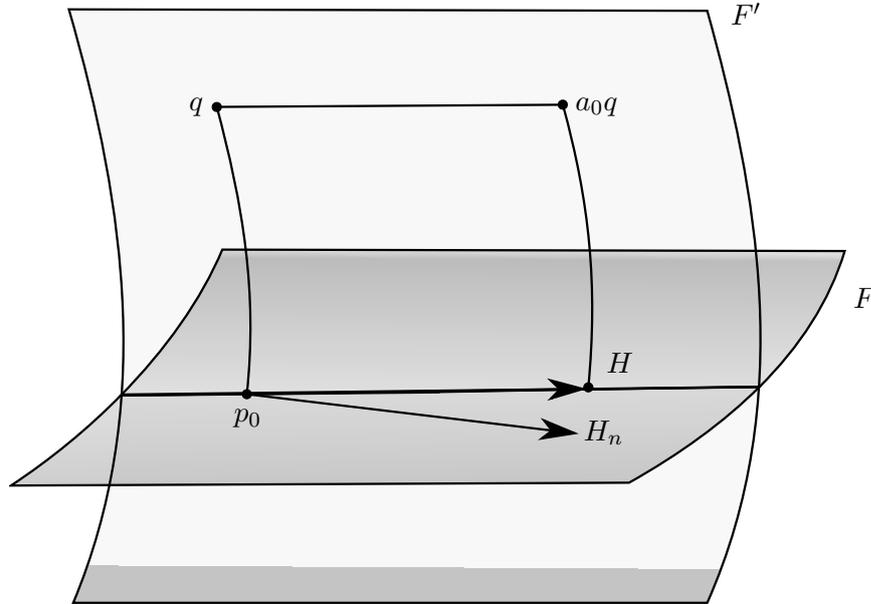


Figure 3.8: The two flats F and F' .

We suppose (3.2) holds for some $\ell > 0$. Let φ be the angle between Y and the flat F and set $q = \exp(\frac{2\ell}{\sin \varphi} Y)p_0$. Then

$$d_X(q, F) \geq \sin \varphi d_X(q, p_0) = 2\ell.$$

But for each $n \in \mathbb{N}$, by construction, $F = \text{MIN}(a_n)$. Thus $d_X(q, \text{MIN}(a_n)) \geq 2\ell$ for each n . Meanwhile the sequence of points $a_n q$ converges to the point $a_0 q$ as n tends to infinity. Hence $d_X(q, a_n q)$ converges to 1. This gives the following contradiction:

$$\begin{aligned} 2\ell &\leq d_X(q, \text{MIN}(a_n)) \\ &\leq \ell d_X(q, a_n q) \\ &\rightarrow \ell \end{aligned}$$

Hence there cannot exist such a constant ℓ which satisfies (3.2) for every hyperbolic element in G .

The following Lemma explains how to use the above to demonstrate the non-existence of a uniform constant for the linear control on conjugacy length among all regular hyperbolic elements in G .

Lemma 3.2.16. *Suppose for $p \in X$ there exists $\ell' > 0$ such that for every pair of conjugate regular hyperbolic elements a, b in G there exists a conjugator $g \in G$ such that*

$$d_X(p, gp) \leq \ell'(d_X(p, ap) + d_X(p, bp)).$$

Then (3.2) holds for $\ell = 2\ell'$.

Proof. Let a be regular hyperbolic in G such that $\text{MIN}(a) = F_a$ does not contain p . Let π_a be the orthogonal projection of X onto F_a and let m be the midpoint of the geodesic segment $[p, \pi_a(p)]$. Consider the geodesic symmetry s_m of X about m , that is the map $s_m : X \rightarrow X$ such that for any geodesic $c : \mathbb{R} \rightarrow X$ with $c(0) = m$, $s_m(c(t)) = c(-t)$ for all $t \in \mathbb{R}$. Since X is a symmetric space s_m is an isometry. Let F denote the flat $s_m(F_a)$. Then $p \in F$ and $[p, \pi_a(p)]$ meets both flats at right-angles, so this geodesic segment realises the distance between the flats.

Take $g \in G$ such that $gF = F_a$ and $gp = \pi_a(p)$. Then by the hypothesis of the Lemma, $d_X(p, \pi_a(p)) \leq \ell'(d_X(p, ap) + d_X(p, g^{-1}agp))$. But $d_X(p, g^{-1}agp)$ is the translation length of a , so is less than $d_X(p, ap)$. Hence $d_X(p, F_a) \leq 2\ell'd_X(p, ap)$. \square

Since (3.2) cannot hold, we have the following:

Corollary 3.2.17. *For $p \in X$ there does not exist $\ell' > 0$ such that for every pair of conjugate regular hyperbolic elements a, b in G there exists a conjugator $g \in G$ such that*

$$d_X(p, gp) \leq \ell'(d_X(p, ap) + d_X(p, bp)).$$

3.2.4 Finding a conjugator in Γ

In Corollary 3.2.13 we found a short conjugator between two real hyperbolic elements in Γ . However this conjugator lies in the ambient Lie group G . In order to improve our understanding of conjugacy length in Γ we need to work out how to move our conjugator from G so that it becomes a conjugator in Γ . The main obstacle here is in understanding how the lattice will intersect flats in the symmetric space.

Given a conjugator $g \in G$ for $a, b \in \Gamma$, the set of all conjugators is the coset $Z_G(a)g$ of the centraliser of a . We are therefore interested in the contents of the set $Z_G(a)g \cap \Gamma$, or equivalently $gZ_G(b) \cap \Gamma$. If we begin with the assumption that a conjugator for a, b from the lattice exists, then at least we know these sets are non-empty.

When a and b are real hyperbolic we have a good understanding of the geometry of their centralisers (see Lemmas 3.2.4 and 3.2.5). For example, when b is regular we can find a point $q \in X$ such that the orbit $Z_G(b)q$ is a maximal flat. Then $gZ_G(b)q$ is also a maximal flat and is in fact the unique maximal flat stabilised by a . Hence it is equal to $Z_G(a)gq$. Clearly $(Z_G(a)g \cap \Gamma)q$ is contained in this flat, and the question is how far is gq from this subset? Once we know this distance, we can shift our conjugator g , whose length we have an estimate for courtesy of Section 3.2.2, to a conjugator in Γ and keep track of the size of the new lattice conjugator.

Suppose $\gamma \in gZ_G(b) \cap \Gamma$. Then $gZ_G(b) \cap \Gamma = \gamma(Z_G(b) \cap \Gamma)$. It is therefore enough to look at how the set $(Z_G(b) \cap \Gamma)q$ sits inside $Z_G(b)q$. In the singular case $Z_G(b)q$ will be made up of a family of maximal flats. Each flat will be the orbit of a maximal torus T contained in $Z_G(b)$. If there exists some such torus T which satisfies the properties that $b \in T \cap \Gamma$ and $T \cap \Gamma$ is isomorphic to \mathbb{Z} , then we understand what the fundamental domain for the action of $T \cap \Gamma$ on Tq will look like: it will be an R -tubular neighbourhood of a hyperplane orthogonal to the geodesics translated by b , where $2R \leq d_X(q, bq)$.

This situation cannot arise if the \mathbb{Q} -rank of the lattice is too small: it must satisfy $\text{rank}_{\mathbb{Q}}(\Gamma) \geq \text{rank}_{\mathbb{R}}(G) - 1$. If $\text{rank}_{\mathbb{Q}}(\Gamma) = \text{rank}_{\mathbb{R}}(G) - 1$ then a maximal \mathbb{Q} -split torus S is a hyperplane inside a maximal \mathbb{R} -split torus T . Because Γ must intersect S in a finite set, Γ will intersect T in nothing more than a finite extension of \mathbb{Z} . In particular, if $q \in X$ is chosen so that Tq is flat, then $(\Gamma \cap T)q$ will look like a copy of \mathbb{Z} inside the flat. Furthermore, the fundamental domain for the action of $T \cap \Gamma$ on this flat will be a tubular neighbourhood of Sq .

In general, if $\text{rank}_{\mathbb{Q}}(\Gamma) = \text{rank}_{\mathbb{R}}(G) - d$, then the flats in X which have a non-trivial intersection with an orbit of Γ can do so only with copies of \mathbb{Z}^k for $\text{rank}_{\mathbb{R}}(G) \geq k \geq d$.

Lemma 3.2.18. *Suppose $\text{rank}_{\mathbb{Q}}(\Gamma) \geq \text{rank}_{\mathbb{R}}(G) - 1$ and let T be a maximal \mathbb{R} -split torus in G such that $T \cap \Gamma$ is a finite extension of \mathbb{Z} . Take a real hyperbolic element $b \in T \cap \Gamma$ and suppose it is conjugate in Γ to a . If a and b have slope ξ , then there exists a conjugator $\gamma \in \Gamma$ for a, b such that*

$$d_X(p, \gamma p) \leq (6L_\xi + 1)(d_X(p, ap) + d_X(p, bp))$$

where L_ξ is as in Corollary 3.2.13.

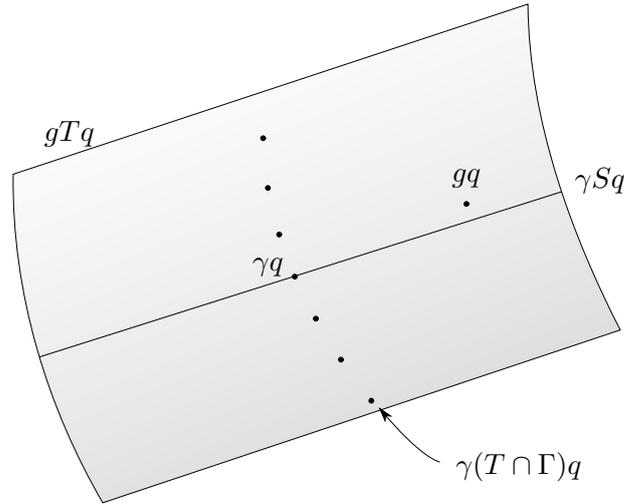


Figure 3.9: When $\text{rank}_{\mathbb{R}}(G) = 2$, the quotient of the maximal flat gTq by the action of $\Gamma \cap gTg^{-1}$ is a cylinder whose diameter is bounded above by $d_X(\gamma q, b\gamma q) = d_X(q, aq)$.

Proof. Suppose first that $\text{rank}_{\mathbb{Q}}(\Gamma) = \text{rank}_{\mathbb{R}}(G) - 1$ and let S be a maximal \mathbb{Q} -split torus contained in T . To adapt the following to the case when $\text{rank}_{\mathbb{Q}}(\Gamma) = \text{rank}_{\mathbb{R}}(G)$, we merely take S^+ to instead be the face of a Weyl chamber in T (which will be isometric to a \mathbb{Q} -Weyl chamber, so S^+ will still isometrically embed into $\Gamma \backslash X$).

Let q be any point in X such that Tq is a (maximal) flat. Note that Tq will be contained in $\text{MIN}(b)$. Choose $g \in G$ as in Theorem 3.2.12, taking $p = q$. Then g maps Tq to a maximal flat contained in $\text{MIN}(a)$ and in particular there exists $\gamma \in \Gamma$ such that gq lies in an R -tubular neighbourhood of γSq , where $2R = d_X(q, aq)$. In particular there is a \mathbb{Q} -Weyl chamber S^+ in S such that gq lies in an R -tubular neighbourhood of γS^+q .

The \mathbb{Q} -Weyl chamber maps isometrically into $\Gamma \backslash X$ (see [Leu04]), so in particular $d_X(\gamma q, gq) \leq d_X(\Gamma q, gq) + R$. Hence $d_X(q, \gamma q) \leq 2d_X(q, gq) + R$.

To finish, we need to translate it so it works for an arbitrary basepoint p . To do this we use Lemma 3.2.11 and apply the triangle inequality:

$$\begin{aligned} d_X(p, \gamma p) &\leq 2d_X(p, q) + d_X(q, \gamma q) \\ &\leq 4L_{\xi}d_X(p, ap) + 4L_{\xi}(d_X(q, aq) + d_X(q, bq)) + r \\ &\leq (6L_{\xi} + 1)(d_X(p, ap) + d_X(p, bp)) \end{aligned}$$

Note that in the above we have assumed that $d_X(p, ap) \leq d_X(p, bp)$; if this is not the case, we can just reverse the roles of a and b . \square

The preceding Lemma works because we know enough about the shape and size of a fundamental domain for the action of $\Gamma \cap T$ on the flat Tq . In this case, Γq intersects Tq in a copy of \mathbb{Z} , and we essentially have a tubular neighbourhood of a co-dimension 1 flat, the radius of the neighbourhood bounded above by the translation length of a .

QUESTION: When Γq intersects Tq in a copy of \mathbb{Z}^k , for some $k \geq 2$, what can we say about the dimensions of the fundamental domain for the action of $\Gamma \cap T$ on the k -dimensional flat inside Tq stabilised by $\Gamma \cap T$ in terms of $d_X(p, bp)$ and $d_X(p, \text{MIN}(b))$, where $b \in T \cap \Gamma$?

3.3 Unipotent Elements

We move on now to look at the conjugacy of unipotent elements in Γ . The partial result obtained here gives a linear bound on the length of a conjugator from G between two lattice elements which satisfy certain conditions and which are both contained in the same minimal unipotent subgroup N of G . The method used to prove it relies heavily on the Lie algebra and in particular on the root system corresponding to \mathfrak{g} .

Consider two conjugate elements u, v in $N \cap \Gamma$. When looking at the case when N is the subgroup of $\text{SL}_n(\mathbb{R})$ consisting of the unipotent upper triangular matrices, the condition we impose on u and v is equivalent to demanding that the super-diagonal entries in the matrix, that is the $(i, i + 1)$ -entries, are all non-zero. The short conjugator we obtain is built up by gradually knocking off entries in the matrix until you are left with a matrix with zeros above the super-diagonal. Doing this for both u and v gives two matrices that are then related via conjugation by a diagonal matrix. In Section 3.3.2 we give a few more details about how the process works for $\text{SL}_n(\mathbb{Z})$.

In the process used to knock off the extra terms in the matrix, the super-diagonal entries play an important role and it is crucial that they are non-zero. The underlying root system for $\text{SL}_n(\mathbb{R})$ is of type A_n , and each positive root corresponds to a particular entry of the matrix. The simple roots correspond to the super-diagonal entries of the matrix. Hence we use the term “simple case” to describe the situation in which we insist that the super-diagonal, or simple, entries of u and v are all non-zero.

The process works especially well in $\text{SL}_n(\mathbb{R})$. Here the root spaces \mathfrak{g}_λ all have dimension 1. In general this is not the case, however it is necessary for our method

to work. We are therefore restricted to looking at Lie groups in which the Lie algebra is split. This is discussed in Section 3.3.1.a.

The ideas laid out below for the simple case could potentially be extended to a more general situation. First one should find the largest parabolic subgroup of G which contains u in its unipotent radical. When looking at the root system, this corresponds to taking bites out of it — i.e. removing the linear spans of certain simple roots. We would then need to find a new subset of the set of positive roots which can play the role of the simple roots. Then conjugate by an element of the parabolic subgroup in order to ensure the corresponding entries are non-zero.

3.3.1 Preliminaries

Fix a point p in the symmetric space X and let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be the Cartan decomposition at p . Take a maximal abelian subspace \mathfrak{a} of \mathfrak{p} — note that this is equivalent to choosing a maximal flat F in X containing p . The (restricted) root-space decomposition of the Lie algebra with respect to \mathfrak{a} is

$$\mathfrak{g} = Z_{\mathfrak{g}}(\mathfrak{a}) \oplus \sum_{\lambda \in \Lambda} \mathfrak{g}_{\lambda}$$

where $\mathfrak{g}_{\lambda} = \{Y \in \mathfrak{g} \mid \text{ad}(H)Y = \lambda(H)Y, \forall H \in \mathfrak{a}\}$. The set Λ is a root system in the dual space of \mathfrak{a} . Let Λ^+ be a subset of positive roots of Λ with corresponding subset Π of simple roots. Then

$$\mathfrak{n} = \sum_{\lambda \in \Lambda^+} \mathfrak{g}_{\lambda}$$

is a nilpotent Lie algebra whose corresponding Lie subgroup N is a maximal unipotent subgroup of G .

Recall that the Iwasawa decomposition of the Lie group is $G = KAN$, where $A = \exp(\mathfrak{a})$ and $K = G_p$ is the Lie subgroup of G with Lie algebra \mathfrak{k} . The Iwasawa decomposition is not unique for G . Indeed we can see how it is determined geometrically. The point p determines the maximal compact subgroup K which appears, while the unipotent subgroup N corresponds to a unique chamber in $\partial_{\infty}X$. In fact, N is the unipotent radical of the (minimal) parabolic subgroup which is the stabiliser of this chamber. The factor A is then determined by the maximal flat which contains p and has this chamber in its boundary.

Since N is a connected nilpotent Lie group, the exponential map $\exp : \mathfrak{n} \rightarrow N$ is a diffeomorphism. In particular, every element u in N has a unique expression

$$u = \exp \left(\sum_{\lambda \in \Lambda^+} Y_\lambda \right)$$

where $Y_\lambda \in \mathfrak{g}_\lambda$ for each positive root λ . Based on this, we introduce some terminology. For u as above, the element Y_λ will be called the λ -entry of u . If λ is a simple root in Λ^+ , that is $\lambda \in \Pi$, then we will say that Y_λ is a *simple entry*.

When taking products of elements in N , a useful tool is the Baker–Campbell–Hausdorff formula. This is a polynomial map $P : \mathfrak{n} \times \mathfrak{n} \rightarrow \mathfrak{n}$ which satisfies:

$$\exp(Y) \exp(Z) = \exp(P(Y, Z)) \quad (3.3)$$

for $Y, Z \in \mathfrak{n}$. See [Var74, §2.5] for discussion on this.

Lemma 3.3.1. *The first few terms of the Baker–Campbell–Hausdorff formula are:*

$$\begin{aligned} P(Y, Z) = & Y + Z + \frac{1}{2}[Y, Z] + \frac{1}{12}[[Y, Z], Z] - \frac{1}{12}[[Y, Z], Y] \\ & - \frac{1}{48}[Z, [Y, [Y, Z]]] - \frac{1}{48}[Y, [Z, [Y, Z]]] + \dots \end{aligned}$$

3.3.1.a Split Lie algebras

The root-space decomposition of a complex semisimple Lie algebra has the useful property that each root-space is of dimension one. This fails to be the case when you replace the word “complex” with “real.” In certain cases however, namely when the Lie algebra is split, this property remains true in the real case.

Let $\mathfrak{g}^{\mathbb{C}}$ denote the complexification of \mathfrak{g} . A real form \mathfrak{g}_0 of $\mathfrak{g}^{\mathbb{C}}$ which satisfies the condition that, for every Cartan decomposition $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$, the space \mathfrak{p}_0 contains a maximal abelian subalgebra of \mathfrak{g}_0 is known as a *split* or *normal* real form. For each complex semisimple Lie algebra there is precisely one (unique up to isomorphism) split real form (see [Vin94, §4 Thm 4.4]). For example, the real special linear Lie algebra $\mathfrak{sl}_n(\mathbb{R})$ is the split real form of $\mathfrak{sl}_n(\mathbb{C})$.

By choosing a compact real form \mathfrak{u} of $\mathfrak{g}^{\mathbb{C}}$ we can determine a Cartan decomposition (see for example [Hel01, Ch.III §7]) of \mathfrak{g} as

$$\mathfrak{g} = \mathfrak{g} \cap \mathfrak{u} + \mathfrak{g} \cap (i\mathfrak{u})$$

and we denote $\mathfrak{k} = \mathfrak{g} \cap \mathfrak{u}$ and $\mathfrak{p} = \mathfrak{g} \cap (\mathfrak{iu})$. Let θ be the associated Cartan involution. If we fix a maximal abelian subspace $\mathfrak{a}_\mathfrak{p}$ of \mathfrak{p} and consider the maximal abelian subalgebra \mathfrak{a} of \mathfrak{g} containing $\mathfrak{a}_\mathfrak{p}$ then $\mathfrak{a}^\mathbb{C}$, the complexification of \mathfrak{a} , is a Cartan subalgebra of $\mathfrak{g}^\mathbb{C}$ (see for example [Hel01, Ch.VI, Lemma 3.2]). This in turn determines a reduced root system Δ inside the dual space of $\mathfrak{a}^\mathbb{C}$ and a root-space decomposition

$$\mathfrak{g}^\mathbb{C} = \mathfrak{a}^\mathbb{C} \oplus \sum_{\alpha \in \Delta} \mathfrak{g}_\alpha^\mathbb{C}$$

where $\mathfrak{g}_\alpha^\mathbb{C} := \{Y \in \mathfrak{g}^\mathbb{C} \mid \text{ad}(H)Y = \alpha(H)Y, \forall H \in \mathfrak{a}^\mathbb{C}\}$. For each root α , the root space $\mathfrak{g}_\alpha^\mathbb{C}$ has complex dimension equal to 1.

Returning to the real world, recall that we have, with respect to \mathfrak{a} , the (restricted) root-space decomposition

$$\mathfrak{g} = Z_\mathfrak{g}(\mathfrak{a}) \oplus \sum_{\lambda \in \Lambda} \mathfrak{g}_\lambda$$

where Λ is the root system, which is considered as a subset of the dual space of \mathfrak{a} . The restricted root-spaces though do not necessarily have dimension one. In fact we can calculate the dimension of \mathfrak{g}_λ in terms of the two root systems Δ and Λ . We can recognise the complexification of \mathfrak{g}_λ as the sum of the root-spaces $\mathfrak{g}_\alpha^\mathbb{C}$ of $\mathfrak{g}^\mathbb{C}$ such that, when restricted from $\mathfrak{a}^\mathbb{C}$ to \mathfrak{a} , the roots α are equal to λ (see [OV90, Ch.5, §4.2]). Hence the dimension of \mathfrak{g}_λ is equal to the number of such roots α . This number is called the *multiplicity* of the restricted root λ .

The following is Proposition 6.3 in Ch.IX of [Hel01]:

Proposition 3.3.2. *Let \mathfrak{g} be a semisimple Lie algebra which is the split real form of its complexification. Then the multiplicity of each restricted root is one.*

To summarise, we can associate to G a reduced root system Λ such that any element $u \in N = \exp(\mathfrak{n})$ can be expressed as

$$u = \exp \left(\sum_{\lambda \in \Lambda^+} Y_\lambda \right)$$

where $Y_\lambda \in \mathfrak{g}_\lambda$. Furthermore, if we assume that the Lie algebra \mathfrak{g} is split, then \mathfrak{g}_λ has dimension 1 for each $\lambda \in \Lambda$.

3.3.1.b The choice of metric on G

The Killing form on \mathfrak{g} is the symmetric bilinear form $B : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{R}$ given by $B(V, W) = \text{Trace}(\text{ad}(V)\text{ad}(W))$. Given $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, the Cartan decomposition at $p \in X$, the Killing form is negative definite on \mathfrak{k} and positive definite on \mathfrak{p} . Furthermore \mathfrak{p} is the orthogonal complement of \mathfrak{k} with respect to the Killing form. Let θ_p be the Cartan involution on \mathfrak{g} defined at p , that is θ_p acts on \mathfrak{k} as the identity and $\theta_p(Y) = -Y$ for any $Y \in \mathfrak{p}$. We can then define an inner product on \mathfrak{g} as follows (see [Ebe96, §2.7] or [Hel01, Ch. III Prop 7.4]):

$$\varphi_p(Y, Z) = -B(\theta_p Y, Z), \text{ for all } Y, Z \in \mathfrak{g}.$$

This then determines a left-invariant Riemannian metric d_G on G . Denote the norm on \mathfrak{g} corresponding to φ_p by $\|\cdot\|$.

We will be interested in the effect of the Lie bracket on the size of elements from the root-spaces of \mathfrak{g} .

Proposition 3.3.3. *Suppose \mathfrak{g} is a split real Lie algebra. Let $Y_\lambda \in \mathfrak{g}_\lambda$ and $Y_\mu \in \mathfrak{g}_\mu$, where $\lambda \in \Pi$ and $\mu \in \Lambda^+$ and $\lambda + \mu$ is a root. Then there exist constants $c_1 \geq c_0 > 0$, independent of the choice of $Y_\lambda, Y_\mu, \lambda, \mu$, such that*

$$c_1 \|Y_\lambda\| \|Y_\mu\| \geq \|[Y_\mu, Y_\lambda]\| \geq c_0 \|Y_\lambda\| \|Y_\mu\|.$$

Proof. This follows from the fact that the root-spaces have dimension one and also from the bilinearity of the Lie bracket and of the inner product φ_p . In particular, if we let Z_λ denote one of the two elements of \mathfrak{g}_λ such that $\|Z_\lambda\| = 1$, and similarly for Z_μ , then for $\alpha, \beta \in \mathbb{R}$ such that $Y_\lambda = \alpha Z_\lambda$ and $Y_\mu = \beta Z_\mu$:

$$\|[Y_\lambda, Y_\mu]\| = |\alpha| |\beta| c_{\lambda, \mu}$$

where $c_{\lambda, \mu} = \|[Z_\lambda, Z_\mu]\|$. By taking

$$c_0 = \min\{c_{\lambda, \mu} \mid \mu \in \Lambda^+, \lambda \in \Pi \text{ such that } \lambda + \mu \in \Lambda^+\}$$

$$c_1 = \max\{c_{\lambda, \mu} \mid \mu \in \Lambda^+, \lambda \in \Pi \text{ such that } \lambda + \mu \in \Lambda^+\}$$

we obtain the result, since $|\alpha| = \|Y_\lambda\|$ and $|\beta| = \|Y_\mu\|$. □

We will need to determine the size of an element in N . In order to do this we need a couple of preliminary results about the size of the entries of u . Throughout the following we will assume $u \in N$ is of the form

$$u = \exp \left(\sum_{\lambda \in \Lambda^+} Y_\lambda \right) \quad (3.4)$$

where $Y_\lambda \in \mathfrak{g}_\lambda$ for each $\lambda \in \Lambda^+$.

Lemma 3.3.4. *Let $u \in N$ be as in (3.4). For each $\lambda \in \Lambda^+$ we have $\|Y_\lambda\| \leq d_G(1, u)$.*

Proof. Since the root-spaces \mathfrak{g}_λ for $\lambda \in \Lambda^+$ are pairwise orthogonal with respect to the Killing form, and hence also the inner product φ_p , we observe that:

$$\|Y_\lambda\| \leq \left\| \sum_{\lambda \in \Lambda} Y_\lambda \right\| = d_G(1, u).$$

□

Lemma 3.3.5. *Let $u \in N$ be as in (3.4). Then there exists $\delta > 0$ such that if $u \in \Gamma$ then for each simple root $\lambda_i \in \Pi$ either $\|Y_{\lambda_i}\| \geq \delta$ or $Y_{\lambda_i} = 0$.*

Proof. Since $\Gamma \cap N$ is a discrete subgroup of N we know it is finitely generated (see, for example, Corollary 2 of Theorem 2.10 in [Rag72]). Let $\{\gamma_1, \dots, \gamma_r\}$ be a set of generators for $\Gamma \cap N$ and let $\gamma = \gamma_{i_1}^{\varepsilon_1} \dots \gamma_{i_s}^{\varepsilon_s} \in \Gamma \cap N$ where $i_j \in \{1, \dots, r\}$ and $\varepsilon_j \in \mathbb{Z} \setminus \{0\}$. We can write each generator as

$$\gamma_i = \exp \left(\sum_{\lambda \in \Lambda^+} Y_\lambda^{(i)} \right)$$

where $Y_\lambda^{(i)} \in \mathfrak{g}_\lambda$ for each i and each λ . Then, by using the Campbell–Baker–Hausdorff formula,

$$\gamma = \exp \left(\sum_{\lambda \in \Lambda^+} \sum_{j=1}^s \varepsilon_j Y_\lambda^{(i_j)} + \tilde{Y} \right)$$

where \tilde{Y} is a sum of terms from non-simple root-spaces. This tells us that each simple entry Y_{λ_i} of u belongs to the integer linear span of the set $\{Y_{\lambda_i}^{(1)}, \dots, Y_{\lambda_i}^{(r)}\}$, hence there is an element of minimal length for each simple root which can appear as an entry of an element in $\Gamma \cap N$. By taking the shortest of these lengths we obtain a positive value for δ . □

3.3.2 Outline of the method

The idea is to conjugate $u \in N$ by a sequence elements of the form $\exp(Z_\mu)$, where $Z_\mu \in \mathfrak{g}_\mu$, or by a commutator of two such elements (from two distinct root-spaces), each step in the sequence removing a λ -entry of u .

For example, when dealing with unipotent upper triangular matrices in $\mathrm{SL}_n(\mathbb{Z})$ each entry in the triangle above the diagonal corresponds to a root. The simple roots correspond to the super-diagonal entries, that is those which lie adjacent to the diagonal. Take

$$u := \begin{pmatrix} 1 & x_1 & y_1 & z_1 \\ 0 & 1 & x_2 & y_2 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Here the x_i entries are the simple entries of u , the y_i terms correspond to roots of height 2 and z_1 to the unique root of height 3. Suppose all these entries are non-zero.

We will conjugate u by an elementary matrix to make the y_1 term vanish:

$$\begin{pmatrix} 1 & \alpha & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} u \begin{pmatrix} 1 & -\alpha & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x_1 & y_1 + \alpha x_2 & z_1 + \alpha y_2 \\ 0 & 1 & x_2 & y_2 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

So if we set $\alpha = -\frac{y_1}{x_2}$ then the entry where y_1 was has now been made to be zero. Notice that all simple entries and the other entry of height 2 are unchanged by this conjugation — the only collateral damage is to entries corresponding to roots of strictly greater height than the entry we removed.

The idea is to repeat this process, next removing the other height 2 entry. This will again cause collateral damage, but it will similarly only effect the height 3 entry. This then is the last entry to be removed and is done so by one last conjugation, but in this case there is no root of greater height than 3, so there will be no collateral damage.

We have conjugated u to

$$u' := \begin{pmatrix} 1 & x_1 & 0 & 0 \\ 0 & 1 & x_2 & 0 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

via upper triangular matrices with rational entries. We do the same for another upper triangular matrix v , reducing to v' in a similar manner. If u and v are conjugate in

$SL_n(\mathbb{Z})$ then u' and v' must be conjugate in $SL_n(\mathbb{Q})$. In fact, when the simple entries in u and v are positive we can find a diagonal matrix over \mathbb{R} to do the job:

$$\begin{pmatrix} \alpha_1 & 0 & 0 & 0 \\ 0 & \alpha_2 & 0 & 0 \\ 0 & 0 & \alpha_3 & 0 \\ 0 & 0 & 0 & \alpha_4 \end{pmatrix} u' \begin{pmatrix} \alpha_1 & 0 & 0 & 0 \\ 0 & \alpha_2 & 0 & 0 \\ 0 & 0 & \alpha_3 & 0 \\ 0 & 0 & 0 & \alpha_4 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & w_1 & 0 & 0 \\ 0 & 1 & w_2 & 0 \\ 0 & 0 & 1 & w_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where $\alpha_1^4 = \frac{w_3^3 w_2^2 w_3}{x_1^3 x_2^2 x_3}$, $\alpha_2^4 = \frac{x_1 w_2^2 w_3}{w_1 x_2^2 x_3}$, $\alpha_3^4 = \frac{x_1 x_2^2 w_3}{w_1 w_2^2 x_3}$, and $\alpha_4^4 = \frac{x_1 x_2^2 x_3^3}{w_1 w_2^2 w_3^3}$.

From our point of view, the crucial aspect of this process is that we can keep track of the size of the conjugator in each step and also control the extent of the collateral damage occurring to entries of greater height.

3.3.3 Relating the root system to conjugation

If N is a maximal unipotent subgroup with corresponding root system Λ then any element in N can be written uniquely as

$$u = \exp\left(\sum_{\lambda \in \Lambda^+} Y_\lambda\right) \quad (3.5)$$

where $Y_\lambda \in \mathfrak{g}_\lambda$. We begin by studying the behaviour of the λ -entries under the action of conjugation by elements in N of the form $\exp(Z_\mu)$, where $Z_\mu \in \mathfrak{g}_\mu$, or a commutator of two such elements.

Lemma 3.3.6. *Let $u \in N$ be as in (3.5) and let $Z_\mu \in \mathfrak{g}_\mu$ for some $\mu \in \Lambda^+$. When we conjugate u by $\exp(Z_\mu)$ all entries of u are unchanged except (possibly) for the λ -entries where $\lambda = r\mu + \lambda'$ for some $r \in \mathbb{N}$ ($r \neq 0$) and $\lambda' \in \Lambda^+$ such that $Y_{\lambda'} \neq 0$. Furthermore, the λ -entry of $\exp(Z_\mu)u \exp(-Z_\mu)$ is*

$$\sum_{r\mu + \lambda' = \lambda} \frac{(\text{ad} Z_\mu)^r Y_{\lambda'}}{r!}$$

where the sum takes values of r from $\mathbb{N} \cup \{0\}$ and λ' from Λ^+ .

Proof. Let $Z = Z_\mu$. We observe that:

$$\begin{aligned} \exp(Z)u \exp(-Z) &= \exp\left(e^{\text{ad}(Z)} \sum_{\lambda' \in \Lambda^+} Y_{\lambda'}\right) \\ &= \exp\left(\sum_{r=0}^{\infty} \sum_{\lambda' \in \Lambda^+} \frac{(\text{ad} Z)^r Y_{\lambda'}}{r!}\right) \end{aligned}$$

Recall that if $r\mu + \lambda'$ is not a root then $(\text{ad}Z)^r Y_{\lambda'} = 0$. Otherwise $(\text{ad}Z)^r Y_{\lambda'} \in \mathfrak{g}_{r\mu + \lambda'}$. It follows that if λ cannot be written as $r\mu + \lambda'$ for any $r \neq 0$ or any λ' then the λ -entry of u is unchanged by this conjugation. \square

The preceding Proposition is important in recognising the link between conjugation of unipotent elements and the root system of G . In particular we can see that if the λ -entry of u is affected by conjugating by $\exp(Z_\mu)$ then the height of λ , denoted $\text{ht}\lambda$, must be greater than $\text{ht}\mu$. Furthermore, the affected entries whose height is precisely $\text{ht}\mu + 1$ will be in the set $\{\mu\} + \Pi$, where Π is the set of simple roots in Λ^+ . This is crucial for motivating Lemma 3.3.9.

To complete the picture which lies behind the scenes of Lemma 3.3.9 we must also consider conjugating by a commutator of two elements. Building up to this, which is Lemma 3.3.8, we give the following:

Lemma 3.3.7. *Let $u \in N$ be as in (3.5) and let $Z_1 \in \mathfrak{g}_{\mu_1}$, $Z_2 \in \mathfrak{g}_{\mu_2}$ for some $\mu_1, \mu_2 \in \Lambda^+$. When we conjugate u by $\exp(Z_1)\exp(Z_2)$ all entries of u are unchanged except (possibly) for the λ -entries where $\lambda = r\mu_1 + t\mu_2 + \lambda'$ for some $\lambda' \in \Lambda^+$ and non-negative integers r, t where at least one of r, t is non-zero.*

Proof. As in the proof of Lemma 3.3.6 we get:

$$\exp(Z_2)\exp(Z_1)u\exp(-Z_1)\exp(-Z_2) = \sum_{t=0}^{\infty} \sum_{r=0}^{\infty} \sum_{\lambda' \in \Lambda^+} \frac{(\text{ad}Z_2)^t (\text{ad}Z_1)^r}{r!t!} Y_{\lambda'}$$

Since $(\text{ad}Z_2)^t (\text{ad}Z_1)^r Y_{\lambda'} \in \mathfrak{g}_{r\mu_1 + t\mu_2 + \lambda'}$ if $r\mu_1 + t\mu_2 + \lambda'$ is a root, or is zero otherwise, it follows as in Lemma 3.3.6 that if λ cannot be expressed as $r\mu_1 + t\mu_2 + \lambda'$ for some λ' and non-negative integers r, t where one of r, t is non-zero, then the λ -entry of u is not affected by this conjugation process. \square

Rather than conjugating by $\exp(Z_1)\exp(Z_2)$ we will conjugate by their commutator $[\exp(Z_1)\exp(Z_2)]$. The extra terms in the product act to clean up any effect conjugating by $\exp(Z_1)\exp(Z_2)$ had on the entries of height less than or equal to $\text{ht}\mu_1 + \text{ht}\mu_2$. Observe that the λ -entry of the conjugate of u by $\exp(Z_1)\exp(Z_2)$ is:

$$\sum_{r\mu_1 + t\mu_2 + \lambda' = \lambda} \frac{(\text{ad}Z_2)^t (\text{ad}Z_1)^r}{r!t!} Y_{\lambda'}$$

Suppose $\text{ht}\lambda \leq \text{ht}\mu_1 + \text{ht}\mu_2$. Then in each term in the sum either $r = 0$ or $t = 0$. We can therefore rewrite it as:

$$\sum_{r\mu_1+\lambda'=\lambda} \frac{(\text{ad}Z_1)^r}{r!} Y_{\lambda'} + \sum_{t\mu_2+\lambda'=\lambda} \frac{(\text{ad}Z_2)^t}{t!} Y_{\lambda'} - Y_{\lambda}.$$

The extra $-Y_{\lambda}$ term is needed because when $r = t = 0$ we count Y_{λ} twice when it should only be counted once. Next we conjugate by $\exp(-Z_1)\exp(-Z_2)$ and we get the following for the λ -entry:

$$\begin{aligned} & \sum_{R\mu_1+\lambda'=\lambda} \frac{(\text{ad}(-Z_1))^R}{R!} \left(\sum_{r\mu_1+\lambda''=\lambda'} \frac{(\text{ad}Z_1)^r}{r!} Y_{\lambda''} + \sum_{t\mu_2+\lambda''=\lambda'} \frac{(\text{ad}Z_2)^t}{t!} Y_{\lambda''} - Y_{\lambda'} \right) \\ + & \sum_{T\mu_2+\lambda'=\lambda} \frac{(\text{ad}(-Z_2))^T}{T!} \left(\sum_{r\mu_1+\lambda''=\lambda'} \frac{(\text{ad}Z_1)^r}{r!} Y_{\lambda''} + \sum_{t\mu_2+\lambda''=\lambda'} \frac{(\text{ad}Z_2)^t}{t!} Y_{\lambda''} - Y_{\lambda'} \right) - Y_{\lambda} \end{aligned}$$

Since $\text{ht}\lambda \leq \text{ht}\mu_1 + \text{ht}\mu_2$, we cannot write $\lambda = R\mu_1 + t\mu_2 + \lambda'$ when both R, t are non-zero (and similarly for r and T). Hence this expression can be reduced to:

$$\begin{aligned} & \sum_{R\mu_1+\lambda'=\lambda} \frac{(\text{ad}(-Z_1))^R}{R!} \left(\sum_{r\mu_1+\lambda''=\lambda'} \frac{(\text{ad}Z_1)^r}{r!} Y_{\lambda''} \right) \\ + & \sum_{T\mu_2+\lambda'=\lambda} \frac{(\text{ad}(-Z_2))^T}{T!} \left(\sum_{t\mu_2+\lambda''=\lambda'} \frac{(\text{ad}Z_2)^t}{t!} Y_{\lambda''} \right) - Y_{\lambda} \end{aligned}$$

This can be rewritten as:

$$\begin{aligned} & \sum_{R\mu_1+\lambda'=\lambda} \left(\sum_{r\mu_1+\lambda''=\lambda'} \frac{(-1)^R}{R!r!} (\text{ad}Z_1)^{R+r} Y_{\lambda''} \right) \\ + & \sum_{T\mu_2+\lambda'=\lambda} \left(\sum_{t\mu_2+\lambda''=\lambda'} \frac{(-1)^T}{T!t!} (\text{ad}Z_2)^{T+t} Y_{\lambda''} \right) - Y_{\lambda} \end{aligned}$$

Notice that whenever $R + r \neq 0$ all the terms cancel, since if $R + r = k \neq 0$ then the coefficient of $(\text{ad}Z_1)^k Y_{\lambda''}$ is:

$$\sum_{R+r=k} \frac{(-1)^R}{R!r!} = 0.$$

A similar statement holds for $T + t \neq 0$. Hence, whenever $\text{ht}\lambda \leq \text{ht}\mu_1 + \text{ht}\mu_2$, the λ -entry is Y_{λ} . We use this in the following:

Lemma 3.3.8. *Let $u \in N$ be as in (3.5) and let $Z_1 \in \mathfrak{g}_{\mu_1}$, $Z_2 \in \mathfrak{g}_{\mu_2}$ for some $\mu_1, \mu_2 \in \Lambda^+$. When we conjugate u by $[\exp(Z_1), \exp(Z_2)]$ all entries of u are unchanged*

except (possibly) for the λ -entries where $\lambda = r\mu_1 + t\mu_2 + \lambda'$ for some $\lambda' \in \Lambda^+$ and non-negative integers r, t .

Furthermore, for such λ , the λ -entry of the conjugate is

$$\sum_{(r+t)\mu_1+(s+u)\mu_2+\lambda'=\lambda} \frac{\text{ad}(-Z_2)^u \text{ad}(-Z_1)^t \text{ad}(Z_2)^s \text{ad}(Z_1)^r Y'_\lambda}{r!s!t!u!}$$

where the summation takes non-negative integers r, s, t, u and positive roots λ' .

Proof. By repeating Lemma 3.3.6 we get that the λ -entry of the conjugate of u by $[\exp(Z_1), \exp(Z_2)]$ is given by

$$\sum_{(r+t)\mu_1+(s+u)\mu_2+\lambda'=\lambda} \frac{\text{ad}(-Z_2)^u \text{ad}(-Z_1)^t \text{ad}(Z_2)^s \text{ad}(Z_1)^r Y'_\lambda}{r!s!t!u!}$$

as required.

This, together with the argument preceding the statement of the Proposition, gives the result. □

The important difference between Lemma 3.3.8 and Lemma 3.3.7 is that when we conjugate by the commutator all entries of height no more than $\text{ht}\mu_1 + \text{ht}\mu_2$ are left unchanged. Furthermore the only (possibly) affected entries of height $\text{ht}\mu_1 + \text{ht}\mu_2 + 1$ are precisely those entries corresponding to roots in the set $\{\mu_1\} + \{\mu_2\} + \Pi$ where Π is the set of simple roots in Λ^+ .

3.3.4 An ordering on the root system

By “the simple case” we mean the case when u is given by

$$u = \exp\left(\sum_{\lambda \in \Lambda^+} Y_\lambda\right)$$

and $Y_\lambda \neq 0$ for each simple root λ . The aim is to find a sequence of elements like those considered in Lemmas 3.3.6 and 3.3.8 which reduce u to a form where the only non-zero λ -entries are those where λ is simple. The following Lemma is necessary to ensure that such a sequence of elements can be found in the simple case.

Lemma 3.3.9. *Let Λ^+ be a set of positive roots and Π the corresponding simple roots associated to a reduced root system Λ . We can assign to Λ^+ an ordering, which we will denote by $<$, such that for every $\lambda \in \Lambda^+ \setminus \Pi$ either:*

- (a) there exists some root μ such the set $\{\mu\} + \Pi$ contains λ and $\lambda \neq \lambda' \in \{\mu\} + \Pi$ implies $\lambda < \lambda'$; or
- (b) there exist roots $\mu_1, \mu_2 \in \Lambda^+$ such that $\{\mu_1\} + \{\mu_2\} + \Pi = \{\lambda\}$ and $\mu_1 + \mu_2$ is not a root.

REMARK: Case (a) corresponds to conjugation by something in $\exp(\mathfrak{g}_\mu)$, see Lemma 3.3.6. Case (b) corresponds to conjugating by a commutator as in Lemma 3.3.8. This Lemma, combined with Lemmas 3.3.6 and 3.3.8, tells us that we can always conjugate $u \in N$ by an element of N in such a way that we can choose the smallest entry of u which is affected by the conjugation.

Proof of Lemma 3.3.9. Before we proceed, note that if we find μ_1, μ_2 satisfying (b) but $\mu_1 + \mu_2$ is a root, then case (a) also applies.

Suppose that Λ is the sum of irreducible root systems $\Lambda_1, \dots, \Lambda_r$ and that $\Lambda^+ = \Lambda_1^+ \cup \dots \cup \Lambda_r^+$. Suppose also that on each Λ_i^+ we have an ordering $<_i$ which satisfies the Lemma. Then we can define an ordering $<$ on Λ^+ given by $\lambda < \mu$ if and only if

1. $\lambda \in \Lambda_i^+$ and $\mu \in \Lambda_j^+$ such that $i < j$; or
2. if $\lambda, \mu \in \Lambda_i^+$ for some i then $\lambda <_i \mu$.

If λ and μ are in different irreducible root systems inside Λ , then $\lambda + \mu$ cannot be a root. Hence it follows that if $<_i$ satisfies the Lemma for each i , then so does $<$. Thus it suffices to check the conditions of the Lemma for each irreducible root system.

In the classical root systems A_n, B_n, C_n, D_n , we make the base assumption that the simple roots are $\Pi = \{\lambda_1, \dots, \lambda_n\}$ and are ordered by $\lambda_1 > \lambda_2 > \dots > \lambda_n$. Note that because we will assume that the Lie algebra is split we do not need to consider BC_n root systems.

ROOT SYSTEMS OF TYPE A_n :

This is the root system associated to $SL_n(\mathbb{Z})$ so we expect this to be straightforward. The non-simple positive roots will be sums of consecutive simple roots:

$$\lambda_i + \lambda_{i+1} + \dots + \lambda_j$$

for $1 \leq i < j \leq n$. The ordering we assign is in two steps: primarily we order by height, then within each height we order the elements lexicographically. So if

$\lambda = \lambda_i + \dots + \lambda_j$ then we take $\mu = \lambda_i + \dots + \lambda_{j-1}$. It follows that:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \dots + \lambda_{j-1}\} & \text{if } i \neq 1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

and hence our chosen ordering satisfies the requirements of the lemma.

ROOT SYSTEMS OF TYPE B_n :

The non-simple positive roots are of the following forms:

$$\begin{array}{ll} \lambda_i + \dots + \lambda_j & \text{for } 1 \leq i < j \leq n \\ \lambda_i + \dots + \lambda_{j-1} + 2\lambda_j + \dots + 2\lambda_n & \text{for } 1 \leq i < j \leq n \end{array}$$

We order the roots as we did for type A_n : first by height, then order the elements of each height lexicographically. If we first take λ of the first form listed above, i.e. $\lambda = \lambda_i + \dots + \lambda_j$. Then we take $\mu = \lambda_i + \dots + \lambda_{j-1}$ and observe that:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \dots + \lambda_{j-1}\} & \text{if } i \neq 1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

satisfies the required conditions. If on the other hand we consider

$$\lambda = \lambda_i + \dots + \lambda_{j-1} + 2\lambda_j + \dots + 2\lambda_n$$

then we take

$$\mu = \begin{cases} \lambda_i + \dots + \lambda_j + 2\lambda_{j+1} + \dots + 2\lambda_n & \text{if } j \neq n \\ \lambda_i + \dots + \lambda_n & \text{if } j = n \end{cases}$$

and observe that:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \dots + \lambda_j + 2\lambda_{j+1} + \dots + 2\lambda_n\} & \text{if } i \neq 1, j \neq n \\ \{\lambda, \lambda_{i-1} + \dots + \lambda_n\} & \text{if } i \neq 1, j = n \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

satisfies the requirements for every choice of i, j .

ROOT SYSTEMS OF TYPE C_n :

The positive non-simple roots have one of the following forms:

$$\begin{array}{ll} \lambda_i + \dots + \lambda_j & \text{for } 1 \leq i < j \leq n \\ 2\lambda_i + \dots + 2\lambda_{n-1} + \lambda_n & \text{for } 1 \leq i \leq n-1 \\ \lambda_i + \dots + \lambda_{j-1} + 2\lambda_j + \dots + 2\lambda_{n-1} + \lambda_n & \text{for } 1 \leq i < j \leq n-1 \end{array}$$

We order these first by height, then order the elements of the same height by lexicographic ordering. We now give the choice for μ in each case.

First, for $1 \leq i < j \leq n$, let $\lambda = \lambda_i + \dots + \lambda_j$. Then we take $\mu = \lambda_i + \dots + \lambda_{j-1}$ and we have:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \dots + \lambda_{j-1}\} & \text{if } i \neq 1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

Under our chosen ordering these satisfy the requirements of the Lemma.

Second, for $1 \leq i \leq n-1$, let $\lambda = 2\lambda_i + \dots + 2\lambda_{n-1} + \lambda_n$. Then we take $\mu = \lambda_i + 2\lambda_{i+1} + \dots + 2\lambda_{n-1} + \lambda_n$ if $i \neq n-1$ or $\mu = \lambda_{n-1} + \lambda_n$ if $i = n-1$ and we have:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \lambda_i + 2\lambda_{i+1} + \dots + 2\lambda_{n-1} + \lambda_n\} & \text{if } i \neq n-1 \text{ and } i \neq 1 \\ \{\lambda, \lambda_{n-2} + \lambda_{n-1} + \lambda_n\} & \text{if } i = n-1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

In each case the elements of $\{\mu\} + \Pi$ are at least as big as λ in our chosen ordering, so the Lemma is satisfied in this case.

Finally, for $1 \leq i < j \leq n-1$, let $\lambda = \lambda_i + \dots + \lambda_{j-1} + 2\lambda_j + \dots + 2\lambda_{n-1} + \lambda_n$. We take $\mu = \lambda_i + \dots + \lambda_j + 2\lambda_{j+1} + \dots + 2\lambda_{n-1} + \lambda_n$ if $j \neq n-1$ or $\mu = \lambda_i + \dots + \lambda_n$ when $j = n-1$. Then:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \dots + \lambda_j + 2\lambda_{j+1} + \dots + 2\lambda_{n-1} + \lambda_n\} & \text{if } j \neq n-1 \text{ and } i \neq 1 \\ \{\lambda, \lambda_{i-1} + \dots + \lambda_n\} & \text{if } j = n-1 \text{ and } i \neq 1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

The requirements of the Lemma are satisfied in each case, and hence it follows that the Lemma holds for root systems of type C_n .

ROOT SYSTEMS OF TYPE D_n :

The non-simple positive roots in the root system D_n are of one of the following two types:

$$\begin{array}{ll} \lambda_i + \dots + \lambda_{j-1} & \text{if } 1 \leq i < j \leq n \\ \lambda_i + \dots + \lambda_{n-2} + \lambda_j + \dots + \lambda_n & \text{if } 1 \leq i < j \leq n \end{array}$$

Apply the same ordering to D_n as we applied to each of the preceding root systems: first order by height, then order the elements of the same height lexicographically.

In most instances we are able to satisfy the conditions of the Lemma by choosing a single $\mu \in \Lambda^+$. However there are some for which we must use the second allowable case, namely find two positive roots μ_1, μ_2 to satisfy the Lemma.

We first suppose $\lambda = \lambda_i + \dots + \lambda_{j-1}$ where $1 \leq i < j < n$. Then we take $\mu = \lambda_i + \dots + \lambda_{j-2}$ and observe:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \dots + \lambda_{j-2}\} & \text{if } i \neq 1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

Hence the conditions of the Lemma are satisfied in each case.

For $1 \leq i < j \leq n$ let $\lambda = \lambda_i + \dots + \lambda_{n-2} + \lambda_j + \dots + \lambda_n$. First assume $i \neq n-2$ and $j \neq n-1, n$. If we take $\mu = \lambda_i + \dots + \lambda_{n-2} + \lambda_{j+1} + \dots + \lambda_n$ then:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \dots + \lambda_{n-2} + \lambda_{j+1} + \dots + \lambda_n & \text{if } i \neq 1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

and the Lemma is satisfied.

Now suppose $j = n$, then $\lambda = \lambda_i + \dots + \lambda_{n-2} + \lambda_n$. Take $\mu = \lambda_i + \dots + \lambda_{n-2}$ then:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \dots + \lambda_{n-2}, \lambda_i + \dots + \lambda_{n-1}\} & \text{if } i \neq 1 \\ \{\lambda, \lambda_1 + \dots + \lambda_{n-1}\} & \text{if } i = 1 \end{cases}$$

and our choice of μ here satisfies the requirements of the Lemma.

We are left with the cases when $\lambda = \lambda_i + \dots + \lambda_{n-1}$ and when $\lambda = \lambda_{n-2} + \lambda_{n-1} + \lambda_n$. In the former case we take $\mu_1 = \lambda_i + \dots + \lambda_{n-3}$ and $\mu_2 = \lambda_{n-1}$ and observe the only way to make a root by adding μ_1, μ_2 and a simple root together is if the simple root is λ_{n-2} , thus giving λ . Hence $\{\mu_1\} + \{\mu_2\} + \Pi = \{\lambda\}$. In the latter case we take $\mu_1 = \lambda_{n-1}$ and $\mu_2 = \lambda_n$. Similarly, since the only simple root which we can add to $\mu_1 + \mu_2$ and still have a root is λ_{n-2} , we have $\{\mu_1\} + \{\mu_2\} + \Pi = \{\lambda\}$.

This completes the verification of the Lemma in the case when the root system is of type D_n .

ROOT SYSTEMS OF TYPE E_6, E_7, E_8, F_4, G_2 :

These are dealt with in the appendix. For root systems E_8 and F_4 a table is produced with an example of an ordering satisfying the Lemma. They also give suitable choices of μ or of μ_1 and μ_2 for each non-simple positive root. Table A.1 gives the ordering for E_8 , and hence for E_7 and E_6 by using the induced ordering. Table A.2 gives the ordering for F_4 . Figures A.1, A.2 and A.3 provide a visual method of checking in each case that the given root μ satisfies the requirements: given $\mu \in \Lambda^+$ one can quickly see what $\{\mu\} + \Pi$ will be by following all edges heading down the page from μ to the row below.

When dealing with G_2 , there is only one root of each height strictly greater than 1, hence we can order the roots by height alone. \square

3.3.5 Reduction of the simple case

From here on in we will assume that \mathfrak{g} is a split real Lie algebra, meaning that the root spaces \mathfrak{g}_λ are 1-dimensional. We first give an algorithm to reduce $u \in N$, all of

whose simple entries are non-zero, to $u' \in N$, all of whose non-simple entries are zero and the simple entries of u' are equal to those of u . Write u in terms of the elements from the root-spaces of \mathfrak{g} :

$$u = \exp \left(\sum_{\lambda \in \Lambda^+} Y_\lambda \right)$$

where $Y_\lambda \in \mathfrak{g}_\lambda$ for each $\lambda \in \Lambda^+$. Assign to Λ^+ the ordering from Lemma 3.3.9. The algorithm is based on an iteration of the following result:

Lemma 3.3.10. *Let λ_0 be the smallest non-simple root such that Y_{λ_0} is non-zero. Then there exists $g \in N$ and a positive constant $c_0 > 0$ such that:*

- (i) *the λ_0 -entry of gug^{-1} is zero and all entries corresponding to smaller roots are unchanged; and*
- (ii) *$d_G(1, g) \leq \frac{\|Y_{\lambda_0}\|}{c_0 \delta}$, where $\delta = \min\{\|Y_{\lambda_i}\| \mid \lambda_i \in \Pi\}$.*

Proof. We begin by applying Lemma 3.3.9 to λ_0 . This gives us either:

- (a) $\mu \in \Lambda^+$ such that λ_0 is minimal in $\{\mu\} + \Pi$; or
- (b) $\mu_1, \mu_2 \in \Lambda^+$ such that $\{\mu_1\} + \{\mu_2\} + \Pi = \{\lambda_0\}$ and $\mu_1 + \mu_2$ is not a root.

First suppose (a) holds. Take $g = \exp(Z_\mu)$ where $Z_\mu \in \mathfrak{g}_\mu$ is chosen so that

$$[Z_\mu, Y_{\lambda_i}] = -Y_{\lambda_0}$$

where λ_i is the simple root such that $\mu + \lambda_i = \lambda_0$. By Lemma 3.3.6, the λ_0 -entry of gug^{-1} is, by construction,

$$Y_{\lambda_0} + \text{ad}(Z_\mu)Y_{\lambda_i} = 0$$

and the other affected entries are of the form $r\mu + \lambda$ for some $\lambda \in \Lambda^+$. All of these are larger than λ_0 in the ordering from Lemma 3.3.9, hence the first part of the lemma is proved when case (a) holds.

Now suppose that instead case (b) holds. Then we take $g = [\exp(Z_1), \exp(Z_2)]$ where $Z_i \in \mathfrak{g}_{\mu_i}$ for $i = 1, 2$. By Lemma 3.3.8, the λ_0 -entry of gug^{-1} is

$$\sum_{(r+t)\mu_1 + (s+u)\mu_2 + \lambda' = \lambda_0} \frac{\text{ad}(-Z_2)^u \text{ad}(-Z_1)^t \text{ad}(Z_2)^s \text{ad}(Z_1)^r Y_{\lambda'}}{r!s!t!u!}$$

where the summation takes non-negative integers r, s, t, u and positive roots λ' . Since $Y_\lambda = 0$ for non-simple roots $\lambda < \lambda_0$, there is no other way to obtain a non-zero term

in the sum except by either taking $r = s = t = u = 0$ and $\lambda' = \lambda_0$ or with $\lambda' = \lambda_i \in \Pi$ such that $\mu_1 + \mu_2 + \lambda_i = \lambda_0$. In the latter case we know $r + t = 1 = s + u$. Hence there are only finitely many combinations to consider and the λ_0 entry becomes:

$$\begin{aligned} & \text{ad}(Z_2)\text{ad}(Z_1)Y_{\lambda_i} + \text{ad}(-Z_2)\text{ad}(Z_1)Y_{\lambda_i} + \\ & \text{ad}(-Z_1)\text{ad}(Z_2)Y_{\lambda_i} + \text{ad}(-Z_2)\text{ad}(-Z_1)Y_{\lambda_i} + Y_{\lambda_0} \end{aligned}$$

which simplifies to

$$\text{ad}(Z_2)\text{ad}(Z_1)Y_{\lambda_i} - \text{ad}(Z_1)\text{ad}(Z_2)Y_{\lambda_i} + Y_{\lambda_0}.$$

Finally, by application of the Jacobi identity, we see this is equal to

$$[[Z_2, Z_1], Y_{\lambda_i}] + Y_{\lambda_0}.$$

Hence, by choosing Z_1 and Z_2 so that $[[Z_2, Z_1], Y_{\lambda_i}] = -Y_{\lambda_0}$, the λ_0 -entry of gug^{-1} is zero.

Finally, Lemma 3.3.8 tells us that entries corresponding to roots of height less than or equal to $\text{ht}\mu_1 + \text{ht}\mu_2$ are unchanged. Since also $\{\mu_1\} + \{\mu_2\} + \Pi = \{\lambda_0\}$, all entries corresponding to roots smaller than λ_0 are unaffected. Thus we have proved (i).

Note that we have the flexibility to choose Z_μ, Z_1 and Z_2 as above because each root-space has dimension one so we only need to choose the appropriate scalar multiple of a basis element to get what we want.

Now we look at the size of g . If $g = \exp(Z_\mu)$ arises from a situation like (a) then, since we chose Z_μ to satisfy $[Z_\mu, Y_{\lambda_i}] = -Y_{\lambda_0}$, we can use Proposition 3.3.3 to show:

$$\begin{aligned} d_G(1, g) &= \|Z_\mu\| \\ &\leq \frac{\|Y_{\lambda_0}\|}{c_0\|Y_{\lambda_i}\|} \\ &\leq \frac{\|Y_{\lambda_0}\|}{c_0\delta} \end{aligned}$$

Suppose instead that $g = [\exp(Z_1), \exp(Z_2)]$, as is necessary for case (b). Using the Baker–Campbell–Hausdorff formula (Lemma 3.3.1), $g = \exp([Z_1, Z_2])$ since $\mu_1 + \mu_2$ is a not a root. Then, again using Proposition 3.3.3 and our choice of Z_1, Z_2 such that

$[[Z_2, Z_1], Y_{\lambda_i}] = -Y_{\lambda_0}$, we see that:

$$\begin{aligned} d_G(1, g) &= d_G(1, \exp[Z_1, Z_2]) \\ &= \|[Z_1, Z_2]\| \\ &\leq \frac{\|Y_{\lambda_0}\|}{c_0 \|Y_{\lambda_i}\|} \\ &\leq \frac{\|Y_{\lambda_0}\|}{c_0 \delta} \end{aligned}$$

This completes (ii). □

The following algorithm describes how, in the simple case, we can reduce $u \in N$ to $u' \in N$, where u' has no non-simple entries.

Algorithm A. Let $u \in N$ be given by

$$u = \exp \left(\sum_{\lambda \in \Lambda^+} Y_\lambda \right).$$

We define a sequence of elements $u(i) \in N$ where $u(0) = u$ and $u(i+1)$ has one fewer non-zero non-simple entry than $u(i)$ and is obtained by

$$u(i) := g(i)u(i-1)g(i)^{-1}, \text{ for } i \geq 1$$

where $g(i)$ is determined by Lemma 3.3.10. This process clearly terminates as Λ^+ is a finite set. Let $g(1), \dots, g(r)$ be the complete set of conjugators obtained. Define $g := g(r) \dots g(1)$. Then $u' := u(r) = gug^{-1}$, which has no non-zero non-simple entries.

3.3.6 The collateral damage of Algorithm A

Suppose now that $u \in N \cap \Gamma$. Before determining the size of a short conjugator in G we need to determine the effect each step of Algorithm A has on the entries of u . This is a notion we described in Section 3.3.2 as collateral damage. We showed in Lemmas 3.3.6 and 3.3.8 that while removing the λ_0 entry of u it was possible that some of the entries of greater height could be altered in the process. We will call those entries affected by one of the steps of Algorithm A, other than the intended target entry, the *collateral damage* of this step.

In general we expect collateral damage. We can, nonetheless, use an iterative method, bounding the size of each $u(i)$ in the sequence. By applying Lemmas 3.3.10 and 3.3.4 we see that the first conjugator $g(1)$ will satisfy

$$d_G(1, g(1)) \leq \frac{1}{c_0 \delta} d_G(1, u) \quad (3.6)$$

where δ is the constant from Lemma 3.3.5. The collateral damage of conjugating u by $g(1)$ includes elements of height greater than that of the smallest non-simple non-zero entry of u . Suppose $g(1), \dots, g(t_1)$ correspond to the steps to remove all entries of height 2. Since conjugating by any of these will not effect any height 2 entry of u , each $g(i)$, for $1 \leq i \leq t_1$, will satisfy inequality (3.6) in place of $g(1)$. Let $g_{\text{ht}}(2) = g(t_1) \dots g(1)$. Then

$$d_G(1, g_{\text{ht}}(2)) \leq \frac{R_2}{c_0 \delta} d_G(1, u)$$

where R_2 is equal to the number of roots of height 2. After the first t_1 steps of Algorithm A we obtain an element $u_{\text{ht}}(2) = g_{\text{ht}}(2) u g_{\text{ht}}(2)^{-1}$ whose entries of height 2 are all zero. Furthermore, by the triangle inequality

$$d_G(1, u_{\text{ht}}(2)) \leq \left(\frac{2R_2}{c_0 \delta} + 1 \right) d_G(1, u).$$

Suppose the λ -entry of $u_{\text{ht}}(2)$ is $Y_\lambda^{(2)}$. Then by Lemma 3.3.4

$$\|Y_\lambda^{(2)}\| \leq \left(\frac{2R_2}{c_0 \delta} + 1 \right) d_G(1, u).$$

By Lemma 3.3.10, the size of the next conjugator will be bounded above:

$$d_G(1, g(t_1 + 1)) \leq \frac{1}{c_0 \delta} \left(\frac{2R_2}{c_0 \delta} + 1 \right) d_G(1, u)$$

noting that we can still use δ as in Lemma 3.3.5 since the simple entries of $u_{\text{ht}}(2)$ are exactly those of u . Let $g_{\text{ht}}(3) = g(t_2) \dots g(t_1 + 1)$, where $g(t_1 + 1), \dots, g(t_2)$ are those conjugators from Algorithm A corresponding to the removal of height 3 entries of u . Then, as in the height 2 case, we get

$$d_G(1, g_{\text{ht}}(3)) \leq \frac{R_3}{c_0 \delta} \left(\frac{2R_2}{c_0 \delta} + 1 \right) d_G(1, u)$$

where R_3 is the number of roots of height 2. Then $u_{\text{ht}}(3) = g_{\text{ht}}(3) u_{\text{ht}}(2) g_{\text{ht}}(3)^{-1}$ has no entries of height 2 or 3, and it satisfies

$$d_G(1, u_{\text{ht}}(3)) \leq \left(\frac{2R_3}{c_0 \delta} + 1 \right) \left(\frac{2R_2}{c_0 \delta} + 1 \right) d_G(1, u).$$

Continuing in this way, if r is the greatest height of a root in Λ^+ , then for each $2 \leq i \leq r$ we have

$$d_G(1, g_{\text{ht}}(i)) \leq \frac{R_i}{c_0\delta} \prod_{j=2}^{i-1} \left(\frac{2R_j}{c_0\delta} + 1 \right) d_G(1, u).$$

Let $g = g_{\text{ht}}(r) \dots g_{\text{ht}}(2)$. Then g is the element obtained from Algorithm A and conjugates u to an element u' whose non-simple entries are all zero, while its simple entries are the same as for u . Finally, we see that the size of g is bounded linearly by the size of u :

Proposition 3.3.11. *Let g be the conjugator obtained by Algorithm A such that the non-simple entries of gug^{-1} are all zero. Then*

$$d_G(1, g) \leq K d_G(1, u)$$

where

$$K = \sum_{i=2}^r \frac{R_i}{c_0\delta} \prod_{j=2}^{i-1} \left(\frac{2R_j}{c_0\delta} + 1 \right).$$

3.3.7 The last step towards finding a short conjugator

Let $v = \exp\left(\sum_{\lambda \in \Lambda^+} W_\lambda\right)$ be an element in N conjugate to u . By applying Algorithm A we may assume that $Y_\lambda = 0 = W_\lambda$ for all non-simple roots $\lambda \in \Lambda^+$. Then, by choosing $H \in \mathfrak{a}$ appropriately, we can conjugate u to v using $\exp(H)$. To be precise:

$$\begin{aligned} gug^{-1} &= \exp(H) \exp\left(\sum_{\lambda \in \Pi} Y_\lambda\right) \exp(-H) \\ &= \exp\left(\sum_{\lambda \in \Pi} e^{\lambda(H)} Y_\lambda\right). \end{aligned}$$

Hence our choice of H needs to be such that $e^{\lambda(H)} Y_\lambda = W_\lambda$. We might ask, what if we need negative scalars? The following Proposition answers this question:

Proposition 3.3.12. *Let u and v be unipotent elements contained in the same maximal unipotent subgroup N of G . Suppose that u is conjugate to v in G and furthermore suppose that the non-simple entries of u and v are all trivial while the simple entries are all non-zero. Then there exists $H_0 \in \mathfrak{a}$ such that*

$$\exp(H_0)u \exp(-H_0) = v.$$

Proof. Let $g \in G$ be such that $gug^{-1} = v$. First observe that, since we are dealing with the simple case, both u and v fix the same unique chamber $\partial_\infty \mathcal{C}$ in the boundary of X and belong to the same minimal parabolic subgroup $G_\xi = Z_\xi N_\xi$, where $Z_\xi = Z_G(A)$, $A = \exp(\mathfrak{a})$ and $N_\xi = N$. Any conjugator from u to v must map $\partial_\infty \mathcal{C}$ to itself, hence $g \in G_\xi$ as well. We may therefore write g as $g = a'n$ where $a' \in Z_G(A)$ and $n \in N$.

Since $n \in N$ it follows that we may write n as

$$n = \exp \left(\sum_{\lambda \in \Lambda^+} Z_\lambda \right)$$

where $Z_\lambda \in \mathfrak{g}_\lambda$. When we conjugate u by n we get the following:

$$\begin{aligned} nun^{-1} &= \exp \left(\sum_{\lambda \in \Lambda^+} Z_\lambda \right) \exp \left(\sum_{\lambda \in \Pi} Y_\lambda \right) \exp \left(- \sum_{\lambda \in \Lambda^+} Z_\lambda \right) \\ &= \exp \left(\sum_{\lambda \in \Pi} Y_\lambda + \tilde{Y} \right) \end{aligned}$$

where \tilde{Y} is the sum of elements \tilde{Y}_λ from the non-simple positive root-spaces. Let \mathfrak{a} be the maximal abelian subspace of \mathfrak{p} such that $A = \exp(\mathfrak{a})$. The exponential map, when restricted to $Z_{\mathfrak{g}}(\mathfrak{a})$, is surjective onto $Z_G(A)$. So there exists $H' \in Z_{\mathfrak{g}}(\mathfrak{a})$ such that $a' = \exp(H')$. We can decompose $Z_{\mathfrak{g}}(\mathfrak{a})$ into the direct sum (see, for example, [Ebe96, 2.17.10])

$$Z_{\mathfrak{g}}(\mathfrak{a}) = \mathfrak{k} \cap Z_{\mathfrak{g}}(\mathfrak{a}) \oplus \mathfrak{a}.$$

Hence there exists unique $U \in \mathfrak{k} \cap Z_{\mathfrak{g}}(\mathfrak{a})$ and $H \in \mathfrak{a}$ such that $H' = U + H$. Since U and H commute, $a' = \exp(U) \exp(H) = \exp(H) \exp(U)$. Conjugating nun^{-1} by $\exp(H)$ gives us

$$\begin{aligned} \exp(H)nun^{-1}\exp(-H) &= \exp(H) \exp \left(\sum_{\lambda \in \Pi} Y_\lambda + \sum_{\lambda \in \Lambda^+ \setminus \Pi} \tilde{Y}_\lambda \right) \exp(-H) \\ &= \exp \left(\sum_{\lambda \in \Pi} e^{\lambda(H)} Y_\lambda + \sum_{\lambda \in \Lambda^+ \setminus \Pi} e^{\lambda(H)} \tilde{Y}_\lambda \right). \end{aligned}$$

Conjugating this by $\exp(U)$ gives us v as

$$v = \exp \left(\sum_{\lambda \in \Pi} e^{\text{ad}(U)} e^{\lambda(H)} Y_\lambda + \sum_{\lambda \in \Lambda^+ \setminus \Pi} e^{\text{ad}(U)} e^{\lambda(H)} \tilde{Y}_\lambda \right).$$

Notice that, since $U \in Z_{\mathfrak{g}}(\mathfrak{a})$, for each $\lambda \in \Lambda^+ \setminus \Pi$ the term $e^{\text{ad}(U)}e^{\lambda(H)}\tilde{Y}_\lambda$ is in the root-space \mathfrak{g}_λ . But the exponential map gives a bijection between \mathfrak{n} and N . Hence

$$\sum_{\lambda \in \Pi} W_\lambda = \sum_{\lambda \in \Pi} e^{\text{ad}(U)}e^{\lambda(H)}Y_\lambda + \sum_{\lambda \in \Lambda^+ \setminus \Pi} e^{\text{ad}(U)}e^{\lambda(H)}\tilde{Y}_\lambda.$$

It follows that $W_\lambda = e^{\text{ad}(U)}e^{\lambda(H)}Y_\lambda$ for each simple root λ and $0 = e^{\text{ad}(U)}e^{\lambda(H)}\tilde{Y}_\lambda$ when λ is non-simple. Thus $\tilde{Y} = 0$ and in particular

$$nun^{-1} = u.$$

It follows that

$$v = gug^{-1} = a'nun^{-1}a'^{-1} = a'ua'^{-1}.$$

In order to finish the proof we find an element $H_0 \in \mathfrak{a}$ to do the required job. Let $C_\lambda(U) \in \mathbb{R}$ be such that $[U, Y_\lambda] = C_\lambda(U)Y_\lambda$. Then $e^{\text{ad}(U)}Y_\lambda = e^{C_\lambda(U)}Y_\lambda$ and in particular we see that there exists a positive constant $C_\lambda = e^{C_\lambda(U)+\lambda(H)}$ for each simple root λ such that

$$W_\lambda = C_\lambda Y_\lambda.$$

Now we notice that in \mathfrak{a} we have sufficient degrees of freedom to choose $H_0 \in \mathfrak{a}$ such that $\lambda(H_0) = C_\lambda$ for each $\lambda \in \Pi$. Then H_0 is the required element to complete the proof. \square

REMARK: Note that to the existence of the constants $C_\lambda(U)$ required the dimension of each simple root-space in \mathfrak{g} to be equal to 1. So Proposition 3.3.12 requires \mathfrak{g} to be split.

Let u, v be unipotent elements contained in the same maximal unipotent subgroup N of G , both of which have all simple entries non-zero. By Algorithm A we can construct g_1 and g_2 in N such that all non-simple entries in $u' = g_1ug_1^{-1}$ and $v' = g_2vg_2^{-1}$ are zero. By Proposition 3.3.12 there exists $g_3 \in A$ such that $g_3u'g_3^{-1} = v'$. Put $g = g_2^{-1}g_3g_1$. Then

$$gug^{-1} = v.$$

With this process we can find a short conjugator for u and v .

Theorem 3.3.13. *Let u, v be conjugate unipotent elements in $N \cap \Gamma$ whose simple entries are all non-zero. Then there exists $g \in G$ such that $gug^{-1} = v$ and which satisfies:*

$$d_G(1, g) \leq L(d_G(1, u) + d_G(1, v))$$

where L depends on the root-system Λ associated to G and $N \cap \Gamma$.

Proof. Recall that $g = g_2^{-1}g_3g_1$ with g_2 and g_1 as in Algorithm A. By Proposition 3.3.11

$$d_G(1, g_1) + d_G(1, g_2) \leq K(d_G(1, u) + d_G(1, v))$$

where K depends on Λ , c_0 and δ . All we need to do now is obtain a linear upper bound for the size of g_3 . By Proposition 3.3.12 this is member of A , equal to $\exp(H)$ for some $H \in \mathfrak{a}$, which satisfies the following for each simple root λ :

$$e^{\lambda(H)} = \frac{\|W_\lambda\|}{\|Y_\lambda\|} \quad (3.7)$$

where Y_λ is the λ -entry of u and W_λ is the λ -entry of v . The size $d_G(1, g_3)$ is given by the norm of H , which is equal to the Killing form

$$B(H, H) = \text{Trace}(\text{ad}(H)^2) = \sum_{\lambda \in \Lambda} \lambda(H)^2.$$

Since every root in Λ can be expressed as an integer linear combination of simple roots, it follows that there exists a constant S_Λ such that when we take the sum over only the simple roots, rather than all positive roots, we get:

$$\sum_{\lambda \in \Pi} \lambda(H)^2 \leq \|H\|^2 = B(H, H) \leq S_\Lambda \sum_{\lambda \in \Pi} \lambda(H)^2. \quad (3.8)$$

By combining (3.7) and (3.8) we get

$$\begin{aligned} d_G(1, g_3) &= \|H\| \\ &\leq S_\Lambda \sum_{\lambda \in \Pi} \lambda(H)^2 \\ &= S_\Lambda \sum_{\lambda \in \Pi} (\ln\|W_\lambda\| - \ln\|Y_\lambda\|)^2 \\ &= S_\Lambda \sum_{\lambda \in \Pi} (\ln\|W_\lambda\|)^2 + (\ln\|Y_\lambda\|)^2 - 2 \ln\|W_\lambda\| \ln\|Y_\lambda\| \\ &\leq S_\Lambda \sum_{\lambda \in \Pi} (\ln d_G(1, v))^2 + (\ln d_G(1, u))^2 - 2 \ln(\delta)^2 \end{aligned}$$

This is therefore sufficient to conclude that the size of g , for sufficiently large u, v , is bounded above by a linear function of $d_G(1, u) + d_G(1, v)$, the coefficient of which will depend on K, δ and S_Λ . This completes the proof. \square

Appendix A

Tables and Figures for Lemma 3.3.9

Table A.1: The simple case for root systems of type E_8

Height	Order	λ	μ or μ_1	μ_2 (if needed)
2	1	$\lambda_1 + \lambda_2$	λ_2	
	2	$\lambda_2 + \lambda_4$	λ_4	
	3	$\lambda_3 + \lambda_4$	λ_3	
	4	$\lambda_4 + \lambda_5$	λ_5	
	5	$\lambda_5 + \lambda_6$	λ_6	
	6	$\lambda_6 + \lambda_7$	λ_7	
	7	$\lambda_7 + \lambda_8$	λ_8	
3	4	$\lambda_1 + \lambda_2 + \lambda_4$	$\lambda_1 + \lambda_2$	
	1	$\lambda_2 + \lambda_3 + \lambda_4$	$\lambda_3 + \lambda_4$	
	3	$\lambda_2 + \lambda_4 + \lambda_5$	λ_2	λ_5
	2	$\lambda_3 + \lambda_4 + \lambda_5$	$\lambda_1 + \lambda_2$	
	5	$\lambda_4 + \lambda_5 + \lambda_6$	$\lambda_5 + \lambda_6$	
	6	$\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_6 + \lambda_7$	
	7	$\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_7 + \lambda_8$	
4	1	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$	$\lambda_1 + \lambda_2 + \lambda_4$	
	3	$\lambda_1 + \lambda_2 + \lambda_4 + \lambda_5$	$\lambda_1 + \lambda_2$	λ_4
	2	$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$	$\lambda_2 + \lambda_4 + \lambda_5$	
	5	$\lambda_2 + \lambda_4 + \lambda_5 + \lambda_6$	$\lambda_2 + \lambda_4$	λ_6
	4	$\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6$	$\lambda_4 + \lambda_5 + \lambda_6$	
	6	$\lambda_4 + \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_5 + \lambda_6 + \lambda_7$	
	7	$\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_6 + \lambda_7 + \lambda_8$	
5	3	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$	
	1	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$	$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$	

Table A.1: The simple case for root systems of type E_8

Height	Order	λ	μ or μ_1	μ_2 (if needed)
	4	$\lambda_1 + \lambda_2 + \lambda_4$ $+ \lambda_5 + \lambda_6$	$\lambda_1 + \lambda_2$	$\lambda_5 + \lambda_6$
	2	$\lambda_2 + \lambda_3 + \lambda_4$ $+ \lambda_5 + \lambda_6$	$\lambda_2 + \lambda_4 + \lambda_5 + \lambda_6$	
	6	$\lambda_2 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7$	$\lambda_2 + \lambda_4$	$\lambda_6 + \lambda_7$
	5	$\lambda_3 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7$	$\lambda_4 + \lambda_5 + \lambda_6 + \lambda_7$	
	7	$\lambda_4 + \lambda_5 + \lambda_6$ $+ \lambda_7 + \lambda_8$	$\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
6	3	$\lambda_1 + \lambda_2 + \lambda_3$ $+ 2\lambda_4 + \lambda_5$	$\lambda_1 + \lambda_2 + \lambda_3$ $+ \lambda_4 + \lambda_5$	
	5	$\lambda_1 + \lambda_2 + \lambda_3$ $+ \lambda_4 + \lambda_5 + \lambda_6$	$\lambda_1 + \lambda_2 + \lambda_3$	$\lambda_5 + \lambda_6$
	2	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6$	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$	
	4	$\lambda_1 + \lambda_2 + \lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + \lambda_2 + \lambda_4$ $+ \lambda_5 + \lambda_6$	
	1	$\lambda_2 + \lambda_3 + \lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_2 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7$	
	7	$\lambda_2 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_4 + \lambda_5$	$\lambda_7 + \lambda_8$
	6	$\lambda_3 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_4 + \lambda_5 + \lambda_6$ $+ \lambda_7 + \lambda_8$	
7	1	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+ 2\lambda_4 + \lambda_5$	$\lambda_1 + \lambda_2 + \lambda_3$ $+ 2\lambda_4 + \lambda_5$	
	3	$\lambda_1 + \lambda_2 + \lambda_3$ $+ 2\lambda_4 + \lambda_5 + \lambda_6$	$\lambda_1 + \lambda_2 + \lambda_3$ $+ \lambda_4 + \lambda_5 + \lambda_6$	
	2	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+ 2\lambda_5 + \lambda_6$	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6$	
	4	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_2 + \lambda_3 + \lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	
	6	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$	λ_7
	5	$\lambda_1 + \lambda_2 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	
	7	$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_3 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	
8	5	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+ 2\lambda_4 + \lambda_5 + \lambda_6$	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+ 2\lambda_4 + \lambda_5$	

Table A.1: The simple case for root systems of type E_8

Height	Order	λ	μ or μ_1	μ_2 (if needed)
	1	$\lambda_1 + \lambda_2 + \lambda_3$ $+2\lambda_4 + 2\lambda_5 + \lambda_6$	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6$	
	3	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	
	2	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	
	6	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	
9	4	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + 2\lambda_5 + \lambda_6$	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + \lambda_5 + \lambda_6$	
	5	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + \lambda_5$	λ_7
	3	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	
	6	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	1	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + \lambda_7$	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	
	2	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	
10	1	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6$	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + 2\lambda_5 + \lambda_6$	
	2	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	
	3	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	6	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7$	λ_6
	4	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+ \lambda_6 + \lambda_7 + \lambda_8$	
	5	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + \lambda_7$	
11	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6$	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6$	
	2	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	
	3	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	

Table A.1: The simple case for root systems of type E_8

Height	Order	λ	μ or μ_1	μ_2 (if needed)
	6	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+ \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	5	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + \lambda_7 + \lambda_8$	
12	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6$	
	2	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	
	3	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	5	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + 2\lambda_7 + \lambda_8$	
13	2	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6 + \lambda_7$	
	3	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	1	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	
	4	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	5	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	
14	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 3\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + 2\lambda_6 + \lambda_7$	
	2	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	3	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	
15	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+4\lambda_4 + 3\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 3\lambda_5 + 2\lambda_6 + \lambda_7$	
	2	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	λ_7

Table A.1: The simple case for root systems of type E_8

Height	Order	λ	μ or μ_1	μ_2 (if needed)
	1	$\lambda_1 + 3\lambda_2 + 3\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
23	1	$2\lambda_1 + 4\lambda_2 + 2\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 3\lambda_2 + 2\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
	2	$2\lambda_1 + 3\lambda_2 + 3\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 3\lambda_2 + 3\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
24	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 4\lambda_2 + 2\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
25	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
26	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
27	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 4\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
28	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 4\lambda_6 + 3\lambda_7 + \lambda_8$	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 4\lambda_6 + 2\lambda_7 + \lambda_8$	
29	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 4\lambda_6 + 3\lambda_7 + 2\lambda_8$	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 4\lambda_6 + 3\lambda_7 + \lambda_8$	

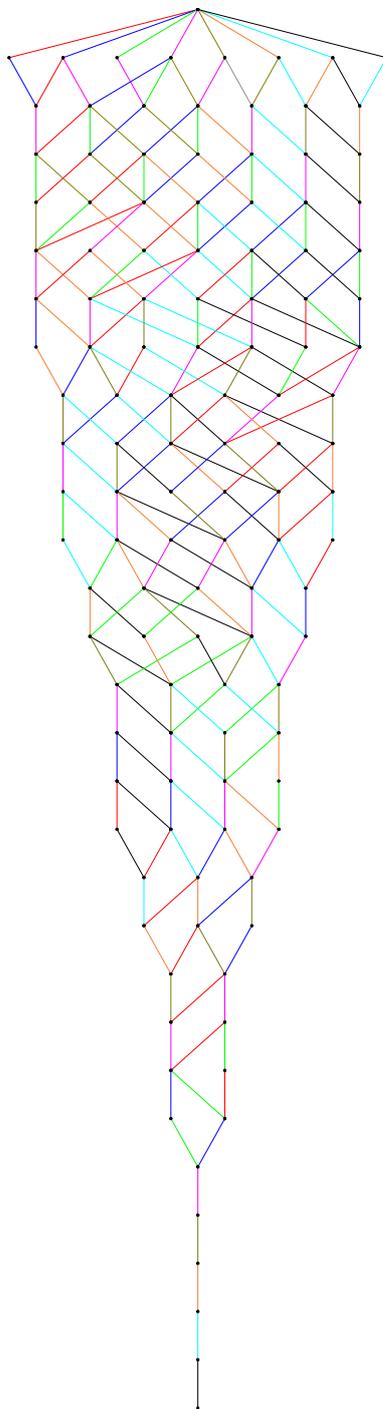


Figure A.1: A graphical depiction of the positive roots in E_8 . The vertices correspond to positive roots (the top vertex is 0), while the edges correspond to addition of a simple root, when reading downwards. Each root has its own colour.

Table A.2: The simple case for root systems of type F_4

Height	Order	λ	μ or μ_1	μ_2 (if needed)
2	1	$\lambda_1 + \lambda_2$	λ_2	
	2	$\lambda_2 + \lambda_3$	λ_3	
	3	$\lambda_3 + \lambda_4$	λ_4	
3	2	$\lambda_1 + \lambda_2 + \lambda_3$	$\lambda_1 + \lambda_2$	
	1	$\lambda_2 + 2\lambda_3$	$\lambda_2 + \lambda_3$	
	3	$\lambda_2 + \lambda_3 + \lambda_4$	$\lambda_3 + \lambda_4$	
4	1	$\lambda_1 + \lambda_2 + 2\lambda_3$	$\lambda_1 + \lambda_2 + \lambda_3$	
	3	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$	$\lambda_1 + \lambda_2$	λ_4
	2	$\lambda_2 + 2\lambda_3 + \lambda_4$	$\lambda_2 + \lambda_3 + \lambda_4$	
5	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3$	$\lambda_1 + \lambda_2 + 2\lambda_3$	
	3	$\lambda_1 + \lambda_2 + 2\lambda_3 + \lambda_4$	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$	
	2	$\lambda_2 + 2\lambda_3 + 2\lambda_4$	$\lambda_2 + 2\lambda_3 + \lambda_4$	
6	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + \lambda_4$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$	
	2	$\lambda_1 + \lambda_2 + 2\lambda_3 + 2\lambda_4$	$\lambda_2 + 2\lambda_3 + 2\lambda_4$	
7	1	$\lambda_1 + 2\lambda_2 + 3\lambda_3 + \lambda_4$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + \lambda_4$	
	2	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 2\lambda_4$	$\lambda_1 + \lambda_2 + 2\lambda_3 + 2\lambda_4$	
8	1	$\lambda_1 + 2\lambda_2 + 3\lambda_3 + 2\lambda_4$	$\lambda_1 + 2\lambda_2 + 3\lambda_3 + \lambda_4$	
9	1	$\lambda_1 + 3\lambda_2 + 3\lambda_3 + 2\lambda_4$	$\lambda_1 + 2\lambda_2 + 3\lambda_3 + 2\lambda_4$	
10	1	$\lambda_1 + 3\lambda_2 + 4\lambda_3 + 2\lambda_4$	$\lambda_1 + 2\lambda_2 + 3\lambda_3 + 2\lambda_4$	
11	1	$2\lambda_1 + 3\lambda_2 + 4\lambda_3 + 2\lambda_4$	$\lambda_1 + 3\lambda_2 + 4\lambda_3 + 2\lambda_4$	

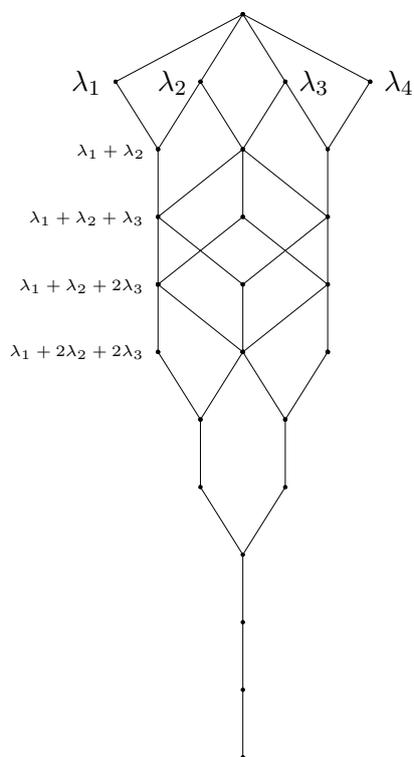


Figure A.2: A graphical depiction of the positive roots in F_4 . The horizontal levels correspond to heights. The elements of height one are labelled, and one more in each height up to 5, but the rest are not. When you move down a height, following an edge corresponds to adding a simple root.

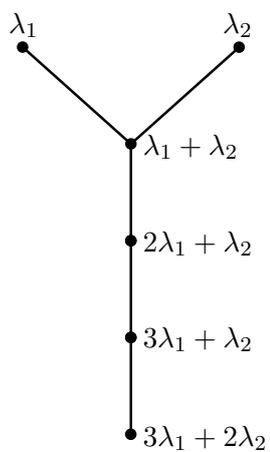


Figure A.3: A graphical depiction of the positive roots in G_2 .

Bibliography

- [AGS06] G. N. Arzhantseva, V. S. Guba, and M. V. Sapir. Metrics on diagram groups and uniform embeddings in a Hilbert space. *Comment. Math. Helv.*, 81(4):911–929, 2006.
- [Aus11] Tim Austin. Amenable groups with very poor compression into Lebesgue spaces. *Duke Math. J.*, 159(2):187–222, 2011.
- [Bau72] G Baumslag. A finitely presented metabelian group with a free abelian derived group of infinite rank. *Proc. Amer. Math. Soc.*, 35:61–62, 1972.
- [Bau74] G Baumslag. Finitely presented metabelian groups. In *Proceedings of the Second International Conference on the Theory of Groups (Australian Nat. Univ., Canberra, 1973)*, volume 372 of *Lecture Notes in Math.*, pages 65–74. Springer, Berlin, 1974.
- [BD11] Jason Behrstock and Cornelia Drutu. Divergence, thick groups, and short conjugators. *arXiv:1110.5005v1 [math.GT]*, 2011.
- [BG96] M. R. Bridson and S. M. Gersten. The optimal isoperimetric inequality for torus bundles over the circle. *Quart. J. Math. Oxford Ser. (2)*, 47(185):1–23, 1996.
- [BGS85] Werner Ballmann, Mikhael Gromov, and Viktor Schroeder. *Manifolds of nonpositive curvature*, volume 61 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1985.
- [BH99] Martin R. Bridson and André Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

- [BMV10] O. Bogopolski, A. Martino, and E. Ventura. Orbit decidability and the conjugacy problem for some extensions of groups. *Trans. Amer. Math. Soc.*, 362(4):2003–2036, 2010.
- [BNW08] Laurent Bartholdi, Markus Neuhauser, and Wolfgang Woess. Horocyclic products of trees. *J. Eur. Math. Soc. (JEMS)*, 10(3):771–816, 2008.
- [Boo59] William W. Boone. The word problem. *Ann. of Math. (2)*, 70:207–265, 1959.
- [BP94] M. R. Bridson and Ch. Pittet. Isoperimetric inequalities for the fundamental groups of torus bundles over the circle. *Geom. Dedicata*, 49(2):203–219, 1994.
- [BS78] Robert Bieri and Ralph Strebel. Almost finitely presented soluble groups. *Comment. Math. Helv.*, 53(2):258–278, 1978.
- [CF63] Richard H. Crowell and Ralph H. Fox. *Introduction to knot theory*. Based upon lectures given at Haverford College under the Philips Lecture Program. Ginn and Co., Boston, Mass., 1963.
- [CFL58] K.-T. Chen, R. H. Fox, and R. C. Lyndon. Free differential calculus. IV. The quotient groups of the lower central series. *Ann. of Math. (2)*, 68:81–95, 1958.
- [CGW09] John Crisp, Eddy Godelle, and Bert Wiest. The conjugacy problem in subgroups of right-angled artin groups. *Journal of Topology*, 2(3):442–460, 2009.
- [CK10] Jeff Cheeger and Bruce Kleiner. Metric differentiation, monotonicity and maps to L^1 . *Invent. Math.*, 182(2):335–370, 2010.
- [CKW94] D. I. Cartwright, V. A. Kaĭmanovich, and W. Woess. Random walks on the affine group of local fields and of homogeneous trees. *Ann. Inst. Fourier (Grenoble)*, 44(4):1243–1288, 1994.
- [CM77] Donald J. Collins and Charles F. Miller, III. The conjugacy problem and subgroups of finite index. *Proc. London Math. Soc. (3)*, 34(3):535–556, 1977.

- [CW12] Matthieu Calvez and Bert Wiest. A fast solution to the conjugacy problem in the 4-strand braid group. *arXiv:1204.6507*, 2012.
- [dC06] Yves de Cornulier. Finitely presented wreath products and double coset decompositions. *Geom. Dedicata*, 122:89–108, 2006.
- [Deh12] M. Dehn. Transformation der Kurven auf zweiseitigen Flächen. *Math. Ann.*, 72(3):413–421, 1912.
- [DLS93] Carl Droms, Jacques Lewin, and Herman Servatius. The length of elements in free solvable groups. *Proc. Amer. Math. Soc.*, 119(1):27–33, 1993.
- [DO11] Tara C. Davis and Alexander Yu. Olshanskii. Subgroup distortion in wreath products of cyclic groups. *J. Pure Appl. Algebra*, 215(12):2987–3004, 2011.
- [Ebe96] Patrick B. Eberlein. *Geometry of nonpositively curved manifolds*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1996.
- [ECH⁺92] David B. A. Epstein, James W. Cannon, Derek F. Holt, Silvio V. F. Levy, Michael S. Paterson, and William P. Thurston. *Word processing in groups*. Jones and Bartlett Publishers, Boston, MA, 1992.
- [Fox53] Ralph H. Fox. Free differential calculus. I. Derivation in the free group ring. *Ann. of Math. (2)*, 57:547–560, 1953.
- [Fox54] Ralph H. Fox. Free differential calculus. II. The isomorphism problem of groups. *Ann. of Math. (2)*, 59:196–210, 1954.
- [Fox56] Ralph H. Fox. Free differential calculus. III. Subgroups. *Ann. of Math. (2)*, 64:407–419, 1956.
- [Fox60] Ralph H. Fox. Free differential calculus. V. The Alexander matrices re-examined. *Ann. of Math. (2)*, 71:408–422, 1960.
- [GK04] Erik Guentner and Jerome Kaminker. Exactness and uniform embeddability of discrete groups. *J. London Math. Soc. (2)*, 70(3):703–718, 2004.
- [Gro87] M. Gromov. Hyperbolic groups. In *Essays in group theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, pages 75–263. Springer, New York, 1987.

- [Gro93] M. Gromov. Asymptotic invariants of infinite groups. In *Geometric group theory, Vol. 2 (Sussex, 1991)*, volume 182 of *London Math. Soc. Lecture Note Ser.*, pages 1–295. Cambridge Univ. Press, Cambridge, 1993.
- [Gru67] K. W. Gruenberg. *Some cohomological topics in group theory*. Queen Mary College Mathematics Notes. Queen Mary College, London, 1967.
- [Har81] O. G. Harlampovič. A finitely presented solvable group with unsolvable word problem. *Izv. Akad. Nauk SSSR Ser. Mat.*, 45(4):852–873, 928, 1981.
- [Hel63] S. Helgason. Duality and Radon transform for symmetric spaces. *Bull. Amer. Math. Soc.*, 69:782–788, 1963.
- [Hel01] Sigurdur Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 34 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Corrected reprint of the 1978 original.
- [Hir73] Friedrich E. P. Hirzebruch. Hilbert modular surfaces. *Enseignement Math. (2)*, 19:183–281, 1973.
- [HKM10] Thomas J. Haines, Michael Kapovich, and John J. Millson. Ideal triangles in Euclidean buildings and branching to Levi subgroups. *arXiv:1011.6636v1 [math.RT]*, 2010.
- [JOR10] Ronghui Ji, Crichton Ogle, and Bobby Ramsey. Relatively hyperbolic groups, rapid decay algebras and a generalization of the Bass conjecture. *J. Noncommut. Geom.*, 4(1):83–124, 2010. With an appendix by Ogle.
- [KL97] Bruce Kleiner and Bernhard Leeb. Rigidity of quasi-isometries for symmetric spaces and Euclidean buildings. *Inst. Hautes Études Sci. Publ. Math.*, (86):115–197 (1998), 1997.
- [KR66] M. I. Kargapolov and V. N. Remeslennikov. The conjugacy problem for free solvable groups. *Algebra i Logika Sem.*, 5(6):15–25, 1966.
- [KT04] Linus Kramer and Katrin Tent. Asymptotic cones and ultrapowers of Lie groups. *Bull. Symbolic Logic*, 10(2):175–185, 2004.

- [KY06] Gennadi Kasparov and Guoliang Yu. The coarse geometric Novikov conjecture and uniform convexity. *Adv. Math.*, 206(1):1–56, 2006.
- [Leu04] Enrico Leuzinger. Tits geometry, arithmetic groups, and the proof of a conjecture of Siegel. *J. Lie Theory*, 14(2):317–338, 2004.
- [Li10] Sean Li. Compression bounds for wreath products. *Proc. Amer. Math. Soc.*, 138(8):2701–2714, 2010.
- [LMR00] Alexander Lubotzky, Shahar Mozes, and M. S. Raghunathan. The word and Riemannian metrics on lattices of semisimple groups. *Inst. Hautes Études Sci. Publ. Math.*, (91):5–53 (2001), 2000.
- [Mag32] W. Magnus. Das Identitätsproblem für Gruppen mit einer definierenden Relation. *Math. Ann.*, 106(1):295–307, 1932.
- [Mag39] Wilhelm Magnus. On a theorem of Marshall Hall. *Ann. of Math. (2)*, 40:764–768, 1939.
- [Mal58] A. I. Mal'cev. On homomorphisms onto finite groups. *Uchen. Zap. Ivanovskogo Gos. Ped. Inst.*, 18(5):49–60, 1958. English transl: On homomorphisms onto finite groups, *Amer. Math. Soc. Translations*, Series 2, 119: 67–79, 1983.
- [Mat66] Jane Matthews. The conjugacy problem in wreath products and free metabelian groups. *Trans. Amer. Math. Soc.*, 121:329–339, 1966.
- [McK43] J. C. C. McKinsey. The decision problem for some classes of sentences without quantifiers. *J. Symbolic Logic*, 8:61–76, 1943.
- [MM00] H. A. Masur and Y. N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.*, 10(4):902–974, 2000.
- [Mos66] A. Włodzimierz Mostowski. On the decidability of some problems in special classes of groups. *Fund. Math.*, 59:123–135, 1966.
- [Nos82] G. A. Noskov. On conjugacy in metabelian groups. *Mat. Zametki*, 31(4):495–507, 653, 1982.

- [Nov55] P. S. Novikov. *Ob algoritmičeskoj nerazrešimosti problemy toždestva slov v teorii grupp*. Trudy Mat. Inst. im. Steklov. no. 44. Izdat. Akad. Nauk SSSR, Moscow, 1955.
- [NP11] Assaf Naor and Yuval Peres. L_p compression, traveling salesmen, and stable walks. *Duke Math. J.*, 157(1):53–108, 2011.
- [Ols] Alexander Yu. Olshanskii. Private communication.
- [Osi01] D. V. Osin. Subgroup distortions in nilpotent groups. *Comm. Algebra*, 29(12):5439–5463, 2001.
- [OV90] A. L. Onishchik and È. B. Vinberg. *Lie groups and algebraic groups*. Springer Series in Soviet Mathematics. Springer-Verlag, Berlin, 1990. Translated from the Russian and with a preface by D. A. Leites.
- [Par00] Anne Parreau. *Dégénérescences de sous-groupes discrets de groupes de Lie semisimples et actions de groupes sur des immeubles affines*. PhD thesis, Univ. Orsay, 2000.
- [Pré06] Jean-Philippe Préaux. Conjugacy problem in groups of oriented geometrizable 3-manifolds. *Topology*, 45(1):171–208, 2006.
- [Pré12] Jean-Philippe Préaux. Conjugacy problem in groups of non-orientable 3-manifolds. *arXiv:1202.4148v2 [math.GR]*, 2012.
- [Rag72] M. S. Raghunathan. *Discrete subgroups of Lie groups*. Springer-Verlag, New York, 1972. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68.
- [RS70] V. N. Remeslennikov and V. G. Sokolov. Certain properties of the Magnus imbedding. *Algebra i Logika*, 9:566–578, 1970. English transl: Some properties of a Magnus embedding, *Algebra and Logic*, 9: 67-79, 1970.
- [Sal11] Andrew Sale. Short conjugators in solvable groups. *arXiv:1112.2721v1 [math.GR]*, 2011.
- [Sal12] Andrew Sale. On the Magnus embedding and the conjugacy length function of wreath products and free solvable groups. *arXiv:1202.5343v1*, 2012.

- [San40] I. N. Sanov. Solution of Burnside's problem for exponent 4. *Leningrad State Univ. Annals [Uchenye Zapiski] Math. Ser.*, 10:166–170, 1940.
- [Sch72] Wolfgang M. Schmidt. Norm form equations. *Ann. of Math. (2)*, 96:526–551, 1972.
- [Shi63] Hideo Shimizu. On discontinuous groups operating on the product of the upper half planes. *Ann. of Math. (2)*, 77:33–71, 1963.
- [ST10] Yehuda Shalom and Terence Tao. A finitary version of Gromov's polynomial growth theorem. *Geom. Funct. Anal.*, 20(6):1502–1547, 2010.
- [ST12] Melanie Stein and Jennifer Taback. Metric properties of Diestel-Leader groups. *arXiv:1202.4199v2 [math.GR]*, 2012.
- [Tao11] Jing Tao. Linearly bounded conjugator property for mapping class groups. *arXiv:1106.2341v2 [math.GT]*, 2011.
- [Var74] V. S. Varadarajan. *Lie groups, Lie algebras, and their representations*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1974. Prentice-Hall Series in Modern Analysis.
- [Vas11] Svetla Vassileva. Polynomial time conjugacy in wreath products and free solvable groups. *Groups Complex. Cryptol.*, 3(1):105–120, 2011.
- [vdG88] Gerard van der Geer. *Hilbert modular surfaces*, volume 16 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1988.
- [Vin94] È. B. Vinberg, editor. *Lie groups and Lie algebras, III*, volume 41 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 1994. Structure of Lie groups and Lie algebras, A translation of it Current problems in mathematics. Fundamental directions. Vol. 41 (Russian), Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1990 [MR1056485 (91b:22001)], Translation by V. Minachin [V. V. Minakhin], Translation edited by A. L. Onishchik and È. B. Vinberg.
- [Woe05] Wolfgang Woess. Lamplighters, Diestel-Leader graphs, random walks, and harmonic functions. *Combin. Probab. Comput.*, 14(3):415–433, 2005.

- [Yu00] Guoliang Yu. The coarse Baum-Connes conjecture for spaces which admit a uniform embedding into Hilbert space. *Invent. Math.*, 139(1):201–240, 2000.