SHORT CONJUGATORS IN SOLVABLE GROUPS

ANDREW SALE

Abstract. We give an upper bound on the size of short conjugators in certain solvable groups. Diestel–Leader graphs, which are a horocyclic product of trees, are discussed briefly and used to study the lamplighter groups. The other solvable groups we look at can be recognised in a similar vein, as groups which act on a horocyclic product of well known spaces. These include the Baumslag–Solitar groups BS(1, q) and semidirect products $\mathbb{Z}_n \rtimes \mathbb{Z}_k$. Results can also be applied to the conjugacy of parabolic elements in Hilbert modular groups and to elements in 3–manifold groups.

1. Introduction

In 1911 Max Dehn [Deh11] set out three decision problems for finitely presented groups: the word problem, the conjugacy problem and the isomorphism problem. Over the last century these have been well-studied and have motivated developments in combinatorial group theory and, more recently, geometric group theory. For the word problem we require a finitely presented group $G = \langle A | R \rangle$, and ask if there is an algorithm which, on input a word $w$ on the generators $A$ and their inverses, determines whether or not $w$ is equal to the identity in the group $G$. The word problem is a special case of the conjugacy problem, which asks whether there is an algorithm which determines when two given words on a chosen set of generators represent conjugate elements in the group. These questions may also be asked of recursively presented groups, and furthermore the conjugacy problem can be extended to ask whether one can find, in some sense, a short conjugator between two given conjugate elements of a group.

The groups of interest in this paper are all 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, but excludes the lamplighter groups 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 [Most66] 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. Recent work of Vassileva [Vas11] based on Matthews’ paper has yielded a polynomial upper bound for the computational complexity of an algorithm for the conjugacy problem in suitably chosen wreath products, and hence also in free solvable groups.

In Section 2 we look at solvable groups which are the semidirect product of two finitely generated torsion-free abelian groups. These are useful in the study of conjugacy in Hilbert modular groups (see for example [vdG88] or [Hir73]), an example of non-uniform lattices in higher rank Lie groups, as their parabolic subgroups are of the form $\mathbb{Z}^n \rtimes \mathbb{Z}^{n-1}$. In particular our results show that a pair of conjugate elements $u, v$ which lie in the same parabolic subgroup of a Hilbert modular group will have a conjugator whose length is bounded exponentially by the lengths of $u$ and $v$, while if $u$ and $v$ are also unipotent then this bound is in fact linear. Theorem 2.1 also completes the picture for conjugacy length in fundamental groups of prime 3–manifolds, adding to the results of Behrstock and Drutu [BD11] for non-geometric prime 3–manifolds and Ji, Ogle and Ramsay [JOR10] for nilmanifolds.

Using a result of Bartholdi, Neuhauser and Woess [BNW08], in Section 3 we exploit the geometry of Diestel–Leader graphs, which are horocyclic products of two trees, to study conjugacy in lamplighter groups $\mathbb{Z}_q \wr \mathbb{Z}$, showing that they enjoy a linear upper bound on minimal conjugator length. We can compare the geometry of lamplighter groups with the geometry of $\mathbb{Z}_2 \rtimes \mathbb{Z}$ by observing that the latter group acts by isometries on the horocyclic product of two copies of the hyperbolic plane, rather than two copies of a tree.

The solvable Baumslag–Solitar groups have a geometry that lies somewhere in between that of lamplighter groups and that of $\mathbb{Z}_2 \rtimes \mathbb{Z}$, and similar ideas can be used to study conjugacy in the solvable groups BS(1, $q$). In Section 4 we use the space described by Farb and Mosher (see for example [Mos99], [FM98], [FM99]) which is quasi-isometric to BS(1, $q$), and is essentially the horocyclic product of a tree and the hyperbolic plane. Once again, the method analogous to that used for the lamplighter group and $\mathbb{Z}_2 \rtimes \mathbb{Z}$ yields a linear upper bound on conjugator length.

The main results of this paper are the following:

**Theorem 1.** Let $\Gamma$ be one of the following groups:
- $\mathbb{Z}^n \rtimes \varphi \mathbb{Z}$, where $\varphi$ is a matrix in $\text{SL}_n(\mathbb{Z})$ whose eigenvalues are all positive real numbers;
- a lamplighter group $\mathbb{Z}_q \wr \mathbb{Z}$;
- a solvable Baumslag–Solitar group BS(1, $q$).

Then there exists a constant $K$ such that two elements $u, v \in \Gamma$ are conjugate if and only if there exists some conjugator $\gamma \in \Gamma$ such that

$$|\gamma| \leq K(|u| + |v|).$$

The results for the lamplighter and Baumslag–Solitar groups are effective. That is, we calculate the value of the constant $K$ in these cases. The polycyclic case can be extended to obtain an exponential upper bound in the more general case:

**Theorem 2.** Let $\Gamma = \mathbb{Z}^n \rtimes \varphi \mathbb{Z}^k$, where the image of $\varphi : \mathbb{Z}^k \hookrightarrow \text{SL}_n(\mathbb{Z})$ consists of semisimple elements. When $k \geq 2$ there exists a positive constant $A$ such that whenever $(u, v)$ and $(w, z)$ are conjugate elements in $\Gamma$ there exists a conjugator $(x, y) \in \Gamma$ such that:

$$d_\Gamma((1, \mathbf{x}, \mathbf{y})) \leq A^{d_\Gamma((1, (u, v)) + d_\Gamma((1, (w, z))))},$$

Acknowledgements: The author would like to thank Cornelia Drutu for the valuable conversations on this topic. He would also like to express gratitude to
Ashot Minasyan for clarifying the situation regarding the solubility of the conjugacy problem for the lamplighter group.

2. Short Conjugators for Lattices in SOL

Remeslennikov proved in 1969 [Rem69] that a polycyclic group is conjugacy separable, and hence, by the aforementioned theorem of Mal’cev [Mal58] and Mostowski [Mos60], has solvable conjugacy problem. The algorithm of Mal’cev and Mostowski is terribly inefficient and sheds absolutely no light on the length of conjugators. However, a more practical solution to the conjugacy problem for a particular family of polycyclic groups — lattices in SOL — is given by Préaux [Pré06]. We give here a bound on the length of the shortest conjugator between two conjugate elements in such a lattice, and later extend the method to solvable groups of the form \( \mathbb{Z}^n \rtimes \varphi \mathbb{Z}^k \).

**Theorem 2.1.** Let \( \Gamma = \mathbb{Z}^2 \rtimes \varphi \mathbb{Z} \), where \( \varphi \) is a hyperbolic matrix in \( \text{SL}_2(\mathbb{Z}) \) with eigenvalue \( \lambda > 1 \). There exists a constant \( K > 0 \), independent of the choice of \( \varphi \), and a constant \( B = B(\varphi) \) such that two elements \( (u, v) \) and \( (w, z) \) of \( \Gamma \) satisfying

\[
B \leq d_\Gamma(1, (u, v)) + d_\Gamma(1, (w, z))
\]

are conjugate if and only if there exists a conjugator \( (x, y) \) in \( \Gamma \) for \( (u, v) \) and \( (w, z) \) such that

\[
d_\Gamma(1, (x, y)) \leq K \max \left\{ \log \lambda - \frac{1}{\log \lambda}, \left( d_\Gamma(1, (u, v)) + d_\Gamma(1, (w, z)) \right) \right\}.
\]

**Proof.** Suppose \( (u, v)(x, y) = (x, y)(w, z) \). By direct calculation we obtain the following two equations:

\[
\begin{align*}
(1) & \quad u + \varphi^y(x) = x + \varphi^y(w) \\
(2) & \quad v + y = y + z
\end{align*}
\]

In particular, first notice that equation (2) immediately implies that \( v = z \). We now consider the two cases when this is zero and when it is non-zero.

**Case 1:** \( v = z = 0 \)

In this case, equation (1) becomes \( u = \varphi^y(w) \). If we write \( u \) and \( v \) in coordinates with respect to a basis of eigenvectors for \( \varphi \) as \( u = (u_1, u_2) \) and \( w = (w_1, w_2) \) then this gives us \( u_1 = \lambda^y w_1 \) and \( u_2 = \lambda^{-y} w_2 \). Hence, in particular, we can write \( y \) as:

\[
y = \frac{\log |u_1| - \log |w_1|}{\log \lambda}.
\]

In order to complete this case we need to find an upper and lower bound on \( |u_1| \) and \( |w_1| \). We do this by applying Roth's Theorem.

In \( \mathbb{R}^2 \), let \( E_1 \) be the eigenspace corresponding to eigenvalue \( \lambda \) and \( E_2 \) the eigenspace corresponding to \( \lambda^{-1} \). The value \( |u_1| \) measures the distance from \( u \) to the eigenspace \( E_2 \) in a direction parallel to \( E_1 \). Let \( (1, \alpha) \) span \( E_2 \). Note that \( \alpha \) is an algebraic number since it lies in the quadratic number field \( \mathbb{Q}(\lambda) \). We will first measure the shortest distance from \( u \) to \( E_2 \).

Let \( \theta \) be the angle between the vectors \((0, 1)\) and \((1, \alpha)\). With respect to the canonical basis, suppose \( u = (p, q) \). Then \( d(u, E_2) = d(u, (p, pa)) \sin \theta \). But \( d(u, (p, pa)) = |q - pa| \). By Roth's Theorem, for every \( \varepsilon > 0 \) there exists a positive constant \( C = C(\alpha, \varepsilon) \) such that

\[
|q - pa| > \frac{C}{|p|^{1+\varepsilon}}.
\]
Meanwhile, note that \( \sin \theta = (\alpha^2 + 1)^{-1} \), so we get in the end a lower bound:

\[
|u_1| > \frac{C_0}{||u||^{1+\varepsilon}}
\]

where the constant \( C_0 \) depends on \( \varepsilon \) and on \( \Gamma \).

Now we need an upper bound. Let \( \psi \) be the angle between the eigenspaces \( E_1 \) and \( E_2 \), chosen so that \( 0 < \psi \leq \pi/2 \). Then a simple geometric argument gives us that \( ||u|| \geq |u_1| \sin \psi \).

Putting these bounds together gives us an upper bound for \( |\log |u_1|| \) as the greater of \( \log ||u|| - \log \sin \psi \) and \( (1 + \varepsilon) \log ||u|| - \log C_0 \). Let \( B \) be the constant \( \max\{ -\log \sin \psi, -\log C_0 \} \). Then we have:

\[
|y| \leq \frac{(1+\varepsilon)(\log ||u|| + \log ||w||) + 2B}{\log \lambda}.
\]

In conclusion, since the first component of \( Z^2 \times_{\varphi} Z \) is exponentially distorted when \( \varphi \) is hyperbolic, it follows that \( (0, y) \) is a conjugator for \( (u, 0) \) and \( (w, 0) \), and, when \( B \leq d_\Gamma(1, (u, 0)) + d_\Gamma(1, (w, 0)) \), we also have

\[
dr(1, (0, y)) \leq \frac{K_1}{\log \lambda}(dr(1, (u, 0)) + dr(1, (w, 0)))
\]

for some constant \( K_1 \) which is independent of the choice of \( u, w \) and \( \varphi \).

**Case 2: \( v = z \neq 0 \)**

We are given that there exists some conjugator \( (x', y') \) in \( \Gamma \) for \( (u, v) \) and \( (u, z) \).

Since \( (u, v)^m \), for any integer \( m \), is in the centraliser of \( (u, v) \), we can premultiply \( (x', y') \) by an appropriate power of \( (u, v) \) to obtain a conjugator \( (x, y) \) in \( \Gamma \) for which \( 0 \leq y < |v| \). From equation (4) we see that \( x \) satisfies

\[
(\text{Id} - \varphi^y)x = u - \varphi^y w.
\]

Since \( 1 \) is not an eigenvalue of \( \varphi^y \) it follows that \( \det(\text{Id} - \varphi^y) \neq 0 \) and thus we obtain a unique solution for \( x \). This expression also gives us an upper bound on the size of \( x \), which will be \( (\lambda^{|v|} + 1)(||u|| + \lambda^{|v|}||w||) \). Because of the exponential distortion, we are interested in finding an upper bound for \( \log ||x|| \). The above upper bound on \( x \) gives:

\[
\log ||x|| \leq \log ||u|| + \log ||w|| + 2|v|\log \lambda + 3\log 2.
\]

Hence we can find a constant \( K_2 \) such that

\[
dr(1, (x, y)) \leq K_2 \log \lambda(d_\Gamma(1, (u, v)) + d_\Gamma(1, (w, z))),
\]

and, since the choice of \( \varphi \) does not effect the distortion constants, \( K_2 \) does not depend on the choice of \( (u, v), (w, z) \) or \( \varphi \).

Let \( M \) be a prime 3-manifold with fundamental group \( G \). Recent work of Behrstock and Drutu [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 \( K(||u| + |v||)^2 \). Combining Theorem 2.1, the corresponding result of Ji, Ogle and Ramsey for nilmanifolds [JOR10, §2.1], and the result of Behrstock and Drutu we have the following:

**Theorem 2.2.** 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

\[
|g| \leq K(||u| + |v||)^2.
\]

By extending the methods of Theorem 2.1 we are able to obtain an upper bound for conjugator length in a larger collection of solvable groups.
Theorem 2.3. Let $\Gamma = Z^n \rtimes \varphi Z^k$, where the image of $\varphi : \mathbb{Z}^k \to SL_n(\mathbb{Z})$ consists of semisimple elements. Then there exist positive constants $A, B$ such that whenever $(u, v)$ and $(w, z)$ are conjugate elements in $\Gamma$ there exists a conjugator $(x, y) \in \Gamma$ such that:

1. if $k = 1$ then $d_T(1, (x, y)) \leq B(d_T(1, (u, v)) + d_T(1, (w, z)))$;
2. if $k > 1$ then $d_T(1, (x, y)) \leq A^d_T(1, (u, v)) + d_T(1, (w, z))$.

Proof. Suppose $(u, v)(x, y) = (x, y)(w, z)$. As in the proof of Theorem 2.1 by direct calculation we obtain the following two equations:

$$\begin{align*}
u + \varphi(v)(x) &= x + \varphi(y)(w) \\
v + y &= y + z
\end{align*}$$

As before, notice that equation (4) immediately implies that $v = z$. We now consider the two cases when this is zero and when it is non-zero.

Case 1: $v = z = 0$

In this case, equation (4) becomes $u = \varphi(y)(w)$. We apply a similar method to that used in the equivalent case in Theorem 2.1 but extra steps are required.

Suppose $\varphi(Z^k)$ is generated by matrices $\varphi_1, \ldots, \varphi_k$, so that if $y = (y_1, \ldots, y_k)$ then

$$\varphi(y) = \varphi_1^{y_1} \cdots \varphi_k^{y_k}.$$  

Fix a basis of eigenvectors of the matrices in $T$. With respect to this basis, let $u, v$ be represented with coordinates $(u_1, \ldots, u_n), (w_1, \ldots, w_n)$ respectively. Suppose the eigenvalues of $\varphi_i$ are $\lambda_{j,i}$ for $j = 1, \ldots, n$ and $i = 1, \ldots, k$. Then, from (3) 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 $Ly = d$, where $L$ is the $n \times k$ matrix with $(r, s)$–entry equal to $\log |\lambda_{r,s}|$ and $d$ is the vector with $j$th entry equal to $\log |u_j| - \log |w_j|$. Since the matrices $\varphi_1, \ldots, \varphi_k$ generate a copy of $Z^k$, 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 $L'y = d'$. By Cramer’s Rule, for each $i = 1, \ldots, k$ we have

$$y_i = \frac{\det(L^{(i)})}{\det(L')}$$  

where $L^{(i)}$ is obtained from $L'$ by replacing the $i$th column with $d'$. Hence $|y_i|$ is bounded by a linear expression in the terms $|\log(|u_j|)| + |\log(|w_j|)|$, for $j = 1, \ldots, k$, and the coefficients are determined by the choice of $\varphi$. Therefore to obtain an upper bound for each $|y_i|$ we need to obtain a lower and upper bound on the values of $\log |u_j|$ and $\log |w_j|$.

Let $E_1, \ldots, E_n$ be the eigenspaces for $T$. Then $|u_1|$ corresponds to the distance from $u$ to the hyperplane $E_1 \oplus \cdots \oplus E_n$ 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 $E_2 \oplus \cdots \oplus E_n$. This lower bound follows from the subspace theorem of Schmidt [Sch72]. Since $E_2, \ldots, E_n$ are eigenspaces for $\varphi(y) \in SL_n(\mathbb{Z})$, there exists algebraic numbers $\alpha_1, \ldots, \alpha_n$ such that $E_2 \oplus \cdots \oplus E_n = \{ x \in \mathbb{R}^n \mid x \cdot (\alpha_1, \ldots, \alpha_n) = 0 \}$. Then, for $u \in \mathbb{Z}^n$, we have

$$d(u, E_2 \oplus \cdots \oplus E_n) = \frac{|u \cdot (\alpha_1, \ldots, \alpha_n)|}{|| (\alpha_1, \ldots, \alpha_n) ||}.$$
Thus, by the subspace theorem, for every $\varepsilon > 0$ there exists a positive constant $C$ such that for every $u \in \mathbb{Z}^n$ we have the following bound on the distance to the hyperplane:

$$d(u, E_2 \oplus \ldots \oplus E_n) \geq \frac{C}{\|u\|^{n-1+\varepsilon}}.$$  

In particular this gives us a lower bound on $|u_1|$ and hence

$$\log |u_1| \leq (n - 1 + \varepsilon) \log \|u\| - \log C.$$

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 $\ell$, determined by $n$, $\varepsilon$ and $\varphi$, such that

$$\|\log (|u_1|)\| \leq \ell \log \|u\|.$$

To conclude this case, 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 $\varphi$ and independent of $u, w$. Furthermore, because the first coordinate in $\Gamma$ is exponentially distorted and the second coordinate undistorted, we indeed have a linear upper bound on conjugator length:

$$d_{\Gamma}(1, (0, y)) \leq B(d_{\Gamma}(1, (u, 0)) + d_{\Gamma}(1, (w, 0)))$$

for some $B > 0$ independent of $u, w$.  

**Case 2:** $v = z \neq 0$

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

$$x = x_1 + x_2, \ u = u_1 + u_2, \ w = w_1 + w_2$$

where $x_1, u_1, w_1 \in E_1$ and $x_2, u_2, w_2 \in V$. Since $\varphi(y)$ commutes with $\varphi(v)$ it follows that this decomposition of $\mathbb{R}^n$ is preserved by $\varphi(y)$. Thus equation (3) gives rise to a pair of equations:

$$u_1 = \varphi(y)w_1$$
$$\ (Id - \varphi(y))x_2 = u_2 - \varphi(y)w_2.$$  

We can solve the former using Case 1. Solving the latter is equivalent to solving equation (3) in the situation where $\varphi(v)$ has no eigenvalues equal to 1. We deal with this in the following.

Rewrite equation (3) as $(Id - \varphi(v)x = u - \varphi(y)w$. Observe that, since all the coefficients in this system are integers and $\det(Id - \varphi(v)) \neq 0$, the entries of $x$ will be bounded by a polynomial in the entries of $Id - \varphi(v)$ and $u - \varphi(y)w$. So if we find a conjugator $(x, y)$ with $y$ bounded then a bound on $x$ will also follow.

We will look at the projection onto $\mathbb{Z}^k$ of the centraliser $Z_\Gamma(u, v)$. Let $(x', y')$ be any element in $\Gamma$ conjugating $(u, v)$ and $(w, z)$.  

**Case 2a:** $k = 1$

When $k = 1$, $Z_\Gamma(u, v)$ projects onto a cyclic subgroup of $\mathbb{Z}$ which contains $v$. In particular there exists $(a, b) \in Z_\Gamma(u, v)$ so that $(x, y) := (a, b)(x', y')$ satisfies $0 \leq y < |v|$. Since 1 is not an eigenvalue of $\varphi(v)$ it follows that $Id - \varphi(v)$ is invertible, so $x = (Id - \varphi(v))^{-1}(u - \varphi(y)w))$ and there is a polynomial $P$ with coefficients determined by $\varphi$ and degree at most $n$, such that $\|x\| \leq P(e^{\|w\|}, \|u\|, e^{\|w\|} \|w\|).$
Hence, using the exponential distortion of the first coordinate in $\Gamma$,
\[
d_{\Gamma}(1, (x, y)) \preceq \log \|x\| + |y| \\
\preceq \log (P(e^{\|v\|}, \|u\|, e^{\|y\|}\|w\|)) + |v| \\
\preceq |v| + \log \|u\| + \log \|w\| \\
\preceq d_{\Gamma}(1, (u, v)) + d_{\Gamma}(1, (w, z)).
\]

**Case 2b:** $k > 1$

The aim here is as before, to find a conjugator $(x, y)$ in which the size of $y$ is controlled. We will show that we can control each coordinate of $y$, bounding it by an exponential in $\|v\|$.

First note that $(a, b)$ is in the centraliser $Z_{\Gamma}(u, v)$ if and only if
\[
a = (Id - \varphi(v))^{-1}(Id - \varphi(b))u \in \mathbb{Z}^n.
\]

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 $(Id - \varphi(v))^{-1}(Id - \varphi(mb))u \in \mathbb{Z}^n$.

Let $L := (Id - \varphi(v))\mathbb{Z}^n$. Denote by $d$ the determinant of $Id - \varphi(v)$. Then $d$ is the index of $L$ in $\mathbb{Z}^n$. Since $\varphi(b)$ commutes with $\varphi(v)$, and hence $Id - \varphi(v)$, it follows that $\varphi(b)L = L$. Let $\bar{u}$ be the image of $u$ in $\mathbb{Z}^n/L$. Then there exists some $t \leq d$ such that $\varphi(tb)\bar{u} = \bar{u}$. In particular $(Id - \varphi(tb))u \in L$.

In the above, if we let $b$ be one of the canonical generators of $\mathbb{Z}^n$, then we see that we can control each coordinate and find $(x, y) = (a, b)(x', y')$ for some $(a, b) \in Z_{\Gamma}(u, v)$ such that $0 \leq y_i < rd$. Since the determinant of $Id - \varphi(v)$ is exponential in $\|v\|$, we have shown that there exists a conjugator $(x, y)$ for $(u, v)$ and $(w, z)$ such that $\|y\| \preceq e^{\|v\|}$. Thus, in conclusion:
\[
d_{\Gamma}(1, (x, y)) \preceq \|y\| + \log \|x\| \\
\preceq e^{|v|} + \log (e^{|y|}\|w\| + \|u\|) \\
\preceq e^{|v|} + \log \|u\| + \log \|w\| \\
\preceq \exp (d_{\Gamma}(1, (u, v)) + d_{\Gamma}(1, (w, z))).
\]

The exponential bound arises because of the way the projection of $Z_{\Gamma}(u, v)$ onto the $\mathbb{Z}^k$-component lies inside $\mathbb{Z}^k$. Indeed, the first part of Case 2b deals with exactly this issue: how we control the shape of this quotient. In particular, one may ask the following:

**Question.** Can one find a pair of conjugate elements in $\Gamma$ whose shortest conjugator is exponential in the sum of the lengths of the two given elements?

We can now apply this result 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 $\varphi$ 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 $\varphi$ to choose from. Hence, by Theorem 2A 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 2.4.** Let $\Gamma = \text{SL}_2(\mathcal{O}_K)$ 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 $\Gamma$, such that a pair of elements $u, v$ in the same
parabolic subgroup of $\Gamma$ are conjugate in $\Gamma$ 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 $\Gamma$, then this upper bound is linear.

Proof. Since $u, v$ are in the same parabolic subgroup of $\Gamma$ then Theorem 2.3 gives the first conclusion. The second conclusion, for unipotent elements, follows from Case 1 (when $v = z = 0$) in the proof of Theorem 2.3. □

3. The Lamplighter Groups

The groups studied in Section 2 were all polycyclic, we now move on to study some solvable groups which lack this property. The first such groups we look at will be the lamplighter groups $\mathbb{Z}_q \wr \mathbb{Z}$. These are not polycyclic and are not finitely presented, but are recursively presented and conjugacy separable.

Bartholdi, Neuhauser and Woess [BNW08] have shown that the Cayley graph of a lamplighter group, with respect to a certain set of generators, is a Diestel–Leader graph, denoted $DL_2(q)$. We make use of this to understand the size of elements in the lamplighter groups, and hence show that they have linear conjugacy length.

3.1. 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 tree and $\omega$ a boundary point of $T$. For any vertex $x$ in $T$ there is a unique geodesic ray emerging from $x$ that is asymptotic to $\omega$. Given a pair of vertices, $x, y$, the corresponding rays will coincide from some vertex $x \bowtie y$ onwards. Using the terminology of [BNW08], $x \bowtie y$ 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 $b$ on the vertices of $T$ as
\[ b(y) = d_T(y, x_0 \bowtie y) - d_T(x_0, x_0 \bowtie y). \]

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

Given a collection $T_1, \ldots, T_n$ of trees together with a chosen collection of respective Busemann functions $b_1, \ldots, b_n$, we define the horocyclic product to be
\[ \prod_{i=1}^n bT_i = \left\{ (y_1, \ldots, y_n) \in T_1 \times \ldots \times T_n \mid \sum_{i=1}^n b_i(y_i) = 0 \right\}. \]

The Diestel–Leader graph $DL(q_1, \ldots, q_d)$ is the horocyclic product of trees $T_{q_i}$, where $T_q$ is the $q + 1$ regular tree. When $q_1 = q_2 = \ldots = q_d = q$, the corresponding Diestel-Leader graph is also denoted by $DL_d(q)$.

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 $\omega$, with the zero level-set passing through $x_0$. The level-sets of Busemann functions on manifolds 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 $DL_2(q)$, 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
\[ \gamma = \begin{pmatrix} \rho^0 & f \\ 0 & 1 \end{pmatrix}, \]
are indeed pairs of horocycles as defined above, one horocycle in each tree.
3.2. Lamplighter Groups and Diestel–Leader Graphs.

**Theorem 3.1** (Bartholdi–Neuhauser–Woess [BNW08, (3.14)]). Let \( \Gamma = \mathbb{Z}_q \wr \mathbb{Z} \) be the lamplighter group. We can represent \( \Gamma \) 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 \( \Gamma \) 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 \( T_q \) 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 \( q^{-v_0(f)} \) (compare with the \( \mathbb{Q}_p \)-adic integers). We use this to define an ultrametric on \( \mathbb{Z}_q((t)) \). See [CKW94, §4] for a more complete picture.

Fix a boundary point \( \omega \in \partial_\infty T_q \). From each vertex \( x \) in the tree there is a unique geodesic ray in the equivalence class \( \omega \). 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 \( T_q \). Suppose the vertex \( x \) is in the \( k \)-th horocycle with respect to \( \omega \). 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 can number the edges in such a way that the geodesic ray from any vertex in \( T_q \) asymptotic to \( \omega \) passes through finitely many edges which are labelled with non-zero elements from \( \mathbb{Z}_q \). We assign to \( x \) the Laurent series
\[
f_x = \sum_{j=-\infty}^{k} a_j t^j.
\]

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 \( \mathbb{Z}_q((t)) \). We identify this to a vertex in the tree as follows: the radius \( q^{-n} \) of the tree 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 \( T_q \) for which \( f_x = f' \), where \( f_x \) is as in (5), and identify this vertex with \( B(f, q^{-n}) \).

![Figure 1](image)

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

The above identification describes a map

\[
\mathcal{V}(T_q) \to \mathcal{B} = \{ B(f, q^{-n}) : f \in \mathbb{Z}_q((t)), n \in \mathbb{Z} \}
\]

where \( \mathcal{V}(T_q) \) is the vertex set of \( T_q \). 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 \( \Gamma_2(\mathbb{Z}_q) \), consisting of matrices

\[
\gamma = \begin{pmatrix} t^s & P \\ 0 & 1 \end{pmatrix}
\]

where \( P \in \mathbb{Z}_q[t^{-1}, t] \) and \( n \in \mathbb{Z} \), acts on \( \mathcal{B} \) as

\[
\gamma \cdot B(f, q^{-n}) = B(P + t^s f, q^{-n-s}).
\]

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 \( \mathbb{Z}_q((t^{-1})) \) — this is possible 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 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 be 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 $\text{DL}_2(q)$ 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, 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 $\text{DL}_2(q)$ 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.

![Figure 2. Moving in $\text{DL}_2(q)$ from $(v_1, v_2)$ to $(w_1, w_2)$: the dotted lines demonstrate the routes each bead needs to take](image)

3.3. Short Conjugators in the Lamplighter Groups. Before we discuss conjugacy, and in particular the issue of finding short conjugators, we need to have an understanding of the size of elements in the lamplighter groups. This means we need to be able to estimate their word lengths, and for this we use the generating set from Theorem 3.1 and the geometry of the Cayley graph $\text{DL}_2(q)$ as described above.

**Lemma 3.2.** Let $\gamma = \begin{pmatrix} t^n & f \\ 0 & 1 \end{pmatrix}$. Then:

1. $d_T(1, \gamma) \geq |n|$;
2. $d_T(1, \gamma) \geq \max\{v_0^+(f), 0\} + \max\{-v_0(f), 0\} \geq v_0^+(f) - v_0(f)$;
3. if $n = 0$ then $d_T(1, \gamma) = 2\max\{v_0^+(f) + 1, 0\} - 2\min\{v_0(f), 0\}$;
4. if $f = 0$ then $d_T(1, \gamma) = |n|$;
5. $d_T(1, \gamma) \leq |n| + 2(v_0^+(f) - v_0(f))$.

**Proof.** For each case we look in $\text{DL}_2(q)$ at a geodesic from $x = (B(0, q^0), B^-(0, q^{-1}))$ to $\gamma x = y = (B(f, q^{-n}), B^-(f, q^{n-1}))$. 
Observe that such a geodesic, when projected onto the first tree, needs to travel from the 0–th horocycle to the n–th horocycle, which are a distance |n| apart, thus giving (1).

For (2), first suppose \( f \notin B(0, q^0) \) and \( f \notin B^-(0, q^{-1}) \). The former is equivalent to saying \( v_0(f) < 0 \) while the latter is equivalent to \( v_0^*(f) \geq 0 \). Consider any geodesic \( c \) from \( x \) to \( y \); project \( c \) onto the first tree to obtain a path \( c_1 \) and project onto the second tree to obtain a path \( c_2 \). The first path, \( c_1 \), must travel from \( B(0, q^0) \) to \( B(P, q^{-k}) \). In order to do so it must pass through the vertex \( B(0, q^{-v_0(f)}) \), in particular it must pass through the \( v_0(f) \)–th horocycle of the first tree. Similarly \( c_2 \) must pass through the vertex \( B^-(0, q^{-v_0(f)}) \), which is contained in the \((-v_0^*(f)-1)\)–th horocycle of the second tree. Since \( c_1 \) and \( c_2 \) are projections from a geodesic in \( \text{DL}_2(q) \), we deduce that \( c_1 \) must pass through the \((v_0^*(f)+1)\)–th horocycle. So \( c_1 \) starts in the 0–th horocycle, must travel up to the \( v_0(f) \)–th horocycle and must travel down to the \((v_0^*(f)+1)\)–th horocycle. In particular, the length of \( c_1 \), and hence the length of \( c \), is bounded below by \( |v_0(f)+1|+|v_0(f)| = v_0^*(f)+1-v_0(f) \).

Suppose, on the other hand, that \( f \in B(0, q^0) \). Then \( 0 \leq v_0(f) \leq v_0^*(f) \). In particular, \( c_2 \) must still pass through \( B^-(0, q^{-v_0(f)}) \), so any geodesic \( c \) from \( x \) to \( y \) will have length at least \( v_0^*(f) = \max\{v_0(f), 0\} + \max\{-v_0(f), 0\} \). If instead \( f \in B^-(0, q^{-1}) \), then the argument is similar.

If \( n = 0 \) then our path must start and end in the same horocycle. By considering the projections \( c_1 \) and \( c_2 \) in the case when \( f \notin B(0, q^0) \) or \( f \notin B^-(0, q^{-1}) \), by a similar argument to above we see that any geodesic \( c \) will have length equal to \( 2(v_0(f)+1)-2v_0(f) \). Also \( v_0^*(f) \geq 0 \) and \( v_0(f) < 0 \), hence (3) holds in this case. If \( f \in B(0, q^0) \), then \( B(0, q^0) = B(f, q^0) \), so in the first tree we do not need to move. In the second tree however we do need to move, and we need to move up to the \((-v_0^*(f)-1)\)–th horocycle and back down to the 0–th horocycle, giving distance \( 2(v_0^*(f)+1) \). Since \( f \in B(0, q^0) \) we know that \( v_0^*(f) \geq 0 \), so (3) holds again. Similarly, if \( f \in B^-(0, q^{-1}) \) then any geodesic \( c \) will have length \( -2v_0(f) \) and (3) holds.

If instead we consider the case when \( f = 0 \) then the geodesic \( c \) will pass through all points \((B(0, q^m), B^-(0, q^{-m-1}))\) for \( m = 0, \ldots, n \). Hence it has length \(|n|\) and (4) holds.

Finally, (5) follows from writing \( \gamma \) as the product
\[
\gamma = \left( \begin{array}{cc} t^0 & f \\ 0 & 1 \end{array} \right) \left( \begin{array}{cc} t^n & 0 \\ 0 & 1 \end{array} \right)
\]
and applying the triangle inequality, using also (4) and (3).  

Let \( \gamma \in \Gamma \). In the following we write \(|\gamma|\) to be \( d_\Gamma(1, \gamma) \), where \( d_\Gamma \) is the word metric on the lamplighter group \( \Gamma \) given by the generating set in Theorem 3.1.

**Theorem 3.3.** Let \( u, v \) be a pair of elements in the lamplighter group \( \Gamma = \mathbb{Z}_q \wr \mathbb{Z} \). Then \( u, v \) are conjugate in \( \Gamma \) if and only if there exists a conjugator \( \gamma \) such that
\[
|\gamma| \leq 3(|u| + |v|).
\]

**Proof.** We use the representation of \( \Gamma \) from Theorem 3.1. Suppose
\[
u = \left( \begin{array}{cc} t^r & P \\ 0 & 1 \end{array} \right), \quad \gamma = \left( \begin{array}{cc} t^n & f \\ 0 & 1 \end{array} \right).
\]
Then, by direct calculation, \( u\gamma = \gamma v \) if and only if the following equations hold:
\[
\begin{align}
(6) \quad s + n &= n + r \\
(7) \quad P + tvf &= f + tvQ
\end{align}
\]
Clearly, equation (6) implies \( r = s \). We split the proof into the two cases when this is zero or non-zero.

**Case 1:** \( r = s = 0 \)

In this case, equation (7) is equivalent to \( P = t^nQ \). We may take \( f = 0 \) and denote \( P \) by

\[
P = \sum_{k=K_0}^{K_1} a_k t^k.
\]

Then it follows that for \( Q \) we have the expression

\[
Q = \sum_{k=K_0}^{K_1} a_k t^{k-n}.
\]

From Lemma 3.2 parts (3) and (4):

\[
|\gamma| = |n| \\
|u| = 2 \max\{K_1 + 1, 0\} - 2 \min\{K_0, 0\} \\
|v| = 2 \max\{K_1 - n + 1, 0\} - 2 \min\{K_0 - n, 0\}
\]

We now look at all the cases, depending on where \( n \) and 0 lie relative to \( K_0 \) and \( K_1 \).

**Case 1a:** \( n \leq K_0 \leq K_1 \)

If \( n \geq 0 \) then, since \( K_1 \geq n \),

\[
|u| + |v| = 4K_1 - 2n + 4 \geq 2n = 2|\gamma|.
\]

Now suppose \( n < 0 \). If \( K_1 \geq 0 \) then

\[
|u| + |v| = 2K_1 - 2n + 2 \geq -2n = 2|\gamma|.
\]

If \( K_1 < 0 \) then, since \( K_1 - K_0 \geq 0 \),

\[
|u| + |v| = 2(K_1 - K_0) - 2n + 2 \geq -2n = 2|\gamma|.
\]

**Case 1b:** \( K_0 \leq n \leq K_1 \)

First suppose \( 0 \leq K_0 \). Then

\[
|u| + |v| = 2(K_1 + 1) + 2(K_1 - K_0 + 1) \geq 2K_1 \geq 2n = 2|\gamma|.
\]

Next suppose \( K_0 \leq 0 \leq K_1 \). Since in this case \( |n| \leq \max\{-K_0, K_1\} \leq K_1 - K_0 \) it follows that

\[
|u| + |v| = 4K_1 - 4K_0 + 4 \geq 4|n|.
\]

Finally suppose \( K_1 \leq 0 \). Then

\[
|u| + |v| = -2K_0 + 2(K_1 - K_0 + 1) \geq -2K_0 \geq -2n = 2|\gamma|.
\]

**Case 1c:** \( K_0 \leq K_1 \leq n, 0 \)

This is analogous to Case 1a. Similar calculations yield \( |\gamma| \leq \frac{1}{2}(|u| + |v|) \).

**Case 2:** \( r = s \neq 0 \)

The aim is to first bound \( n \), then using equation (7) we can subsequently bound \( v_0(f) \) and \( v_0^{-1}(f) \) as well. Firstly, by switching to look at \( u^{-1} \) and \( v^{-1} \) if necessary, we may assume \( s > 0 \). Furthermore, we can find a conjugator \( \gamma \) for \( u, v \) with \( 0 \leq n < s \) by premultiplying any given conjugator by an appropriate power of \( u \) until we reach such a \( \gamma \).
Suppose we have the following:

\[ f = \sum_{P_0} \lambda_k t^k, \quad u = \sum_{P_0} a_k t^k, \quad v = \sum_{Q_0} b_k t^k. \]

Using equation (7), for each \( k \in \mathbb{Z} \) we get:

\[ \lambda_k - \lambda_{k-s} = a_k - b_{k-n}. \]

In particular, if \( k > \max \{ P_1, Q_1 + n \} \) or \( k < \min \{ P_0, Q_0 + n \} \) then \( a_k = b_{k-n} = 0 \), so \( \lambda_k = \lambda_{k-s} \). If \( \lambda_k \neq 0 \) for some \( k < \min \{ P_0, Q_0 + n \} \), then this means \( \lambda_k - m s \neq 0 \) for each positive integer \( m \), contradicting the fact that \( f \in \mathbb{Z}[t^{-1}, t] \). Alternatively, if \( \lambda_k \neq 0 \) for some \( k > \max \{ P_1 - s, Q_1 + n - s \} \) then \( \lambda_{k+m} \neq 0 \) for each positive integer \( m \), again giving a contradiction. Hence:

\[ v_0(f) \geq \min \{ P_0, Q_0 + n \} = \min \{ v_0(P), v_0(Q) + n \} \]

and

\[ v_0^{-}(f) \leq \max \{ P_1 - s, Q_1 + n - s \} = \max \{ v_0^{-}(P) - s, v_0^{-}(Q) + n - s \}. \]

Using Lemma 3.2 [6] we get

\[ |\gamma| \leq |u| + 2(v_0^{-}(f) - v_0(f)). \]

Making use of Lemma 3.2 [2] and that \( s > 0 \) we observe that

\[ v_0^{-}(f) - v_0(f) \leq \max \{ |u|, v_0^{-}(P) - v_0(Q), v_0^{-}(Q) - v_0(P) + n, |v| \} \]

\[ \leq \max \{ |u|, v_0^{-}(P) + \max \{ -v_0(P), 0 \} + \max \{ v_0^{-}(Q), 0 \} - v_0(Q), v_0^{-}(P) + v_0^{-}(Q) - v_0(P) + \max \{ -v_0(Q), 0 \} + n, |v| \} \]

\[ \leq |u| + |v| + n \]

Hence, putting the pieces together, we obtain

\[ |\gamma| \leq 3(|u| + |v|). \]

\[ \square \]

4. Solvable Baumslag–Solitar Groups

We now move on to look at solvable groups whose geometry lies, in some sense, between the geometry of the lamplighter groups and lattices in SOL. The solvable Baumslag–Solitar groups are, like the lamplighter groups, not polycyclic, but on the other hand, like lattices in SOL, are finitely presented. They are given by the presentation

\[ BS(1, q) = \langle a, b \mid aba^{-1} = b^q \rangle. \]

Since they are finitely presented metabelian groups, a result of Noskov [Nos82] tells us that they have solvable conjugacy problem. The geometry of Baumslag–Solitar groups is well understood and has been studied in particular by Farb and Mosher [Mos99, FM98, FM99]. They describe a 2–complex \( X \) on which \( BS(1, q) \) acts cocompactly and properly discontinuously by isometries. Hence, rather than studying words on \( \{ a, b \} \), we can use the quasi-isometry between \( X \) and \( BS(1, q) \), equipped with the word metric. The space \( X \) was originally discussed in [ECH+92].

In Section 3.1 we described the notion of a horocyclic product of graphs. We can extend this notion beyond graphs, for example taking a non-positively curved manifold \( M \) with any Busemann function (in the usual sense). If we take \( M \) to be the upper-half plane model of the hyperbolic plane \( \mathbb{H} \) with the Busemann function corresponding to the boundary point \( \infty \), then the horocyclic product of \( \mathbb{H} \) with a \((q + 1)-\)valent tree will give us \( X \). We should take a little extra care, however, as we should rescale the metric on \( \mathbb{H} \) so that the distance between the horospheres \( \{ i + x \mid x \in \mathbb{R} \} \) and \( \{ q^i i + x \mid x \in \mathbb{R} \} \) is \( r \). This way they will line up as we want.
Figure 3. “Treebolic” space, the horocyclic product of $H$ and the tree $T_q$. For each geodesic in $T_q$ asymptotic to $\omega$ we can see a copy of $H$, such as in the shaded region.

them to in the horocyclic product. Note that we will always assume $\mathbb{H}$ has this rescaled metric. The term “treebolic” has been given to spaces such as $X$, a name given by Saloff-Coste [BSCSW11].

We can recognise $BS(1,q)$ as the group of matrices

$$\gamma = \begin{pmatrix} q^n & f \\ 0 & 1 \end{pmatrix}; \quad n \in \mathbb{Z}, \quad f = \sum_{m=-\infty}^{\infty} \alpha_m q^m,$$

where $f$ is a $q$-adic number, with $\alpha_m \in \{0, \ldots, q-1\}$ and non-zero for at most finitely many negative $m$. Denote by $v_0(f)$ the minimal integer $m$ so that $\alpha_m \neq 0$. The element $\gamma$ will act on the tree in a similar manner to the action of the lamp-lighter group. We identify vertices of the tree with closed balls $B(P,q^{-r})$ in the $q$–adic numbers $\mathbb{Q}_q$, where $P \in \mathbb{Q}_q$ and $r \in \mathbb{Z}$, and $\gamma$ acts on $B(P,q^{-r})$ by sending it to $B(f+q^n P, q^{-n-r})$. A general point in $X$ can be recognised as $(B(P,q^{-r})i+y)$ for $y \in \mathbb{R}$, and $\gamma$ sends this to the point

$$B(f+q^n P, q^{-n-r}), \gamma \cdot (q^{-r}i+y)$$

where $\gamma$ acts on $\mathbb{H}$ in the usual way.

The following Lemma provides some useful relationships between distances in $\mathbb{H}$ and the valuation $v_0(f)$ of the $q$–adic number $f$.

**Lemma 4.1.** Let $f$ be a $q$–adic number. Then

$$(\log q - \log \sqrt{2}) \max\{v_0(f), 0\} \leq d_{\mathbb{H}}(i, i+f) \leq 2 \log(1+f).$$

**Proof.** Using the identity $d_{\mathbb{H}}(i, i+f) = \cosh^{-1}(1 + \frac{f^2}{2})$ we first observe that

$$\log(1+f) - \log \sqrt{2} \leq d_{\mathbb{H}}(i, i+f) \leq 2 \log(1+f).$$

Then, since $f = \sum \alpha_m q^m$, notice that $f \geq q^{v_0(f)}$. If $v_0(f)$ is positive then we use $\log(1+f) \geq \log q^{v_0(f)}$. When $v_0(f)$ is non-positive we instead use $\log(1+f) \geq 0$. Hence

$$\max\{v_0(f) \log q, 0\} \leq \log(1+f)$$

and the result follows. \qed
Since BS(1, q) is quasi-isometric to $X$, instead of looking at word length we may consider the distance elements in the group move points in $X$.

**Lemma 4.2.** Let $\gamma = \left( \begin{array}{cc} q^n & f \\ 0 & 1 \end{array} \right)$ and fix the basepoint $x = (B(0, q^0), i)$ in $X$. Then:

1. $d_X(x, \gamma x) \geq |n|$;
2. $d_X(x, \gamma x) \geq \frac{1}{2} (d_\mathbb{B}(i, \gamma \cdot i) + \max\{-v_0(f), 0\})$;
3. $d_X(x, \gamma x) \geq \max\{-v_0(f), 0\} \geq -v_0(f)$;
4. $d_X(x, \gamma x) \geq d_\mathbb{B}(i, \gamma \cdot i)$;
5. Let $A = \min\{\frac{1}{2} (\log q - \log \sqrt{2}), 1\}$, if $n = 0$ then
   
   $A |v_0(f)| \leq d_X(x, \gamma x) \leq d_\mathbb{B}(i, \gamma \cdot i) + 2 \max\{-v_0(f), 0\}$;
6. If $f = 0$ then $d_X(x, \gamma x) = |n|$;
7. $d_X(x, \gamma x) \leq |n| + d_\mathbb{B}(i, \gamma \cdot i) + 2 = v_0(f)$.

**Proof.** We look in $X$ at a geodesic $c$ from $x$ to $\gamma \cdot x = (B(f, q^{-n}), \gamma \cdot i)$. Note that projecting onto either $\mathbb{T}_q$ or $\mathbb{H}$ is distance reducing. Projecting onto the tree gives us a path from the $0$-th horocycle to the $n$-th horocycle, hence the length of $c$ must be bounded below by $|n|$, giving [1]

Let $c_1$ be the projection of $c$ onto the tree, and $c_2$ the projection onto $\mathbb{H}$. If $v_0(f) < 0$ then $c_1$ must pass through $B(0, q^{-v_0(f)})$, giving [3]. Meanwhile, $c_2$ will clearly have to have length at least $d_\mathbb{B}(i, \gamma \cdot i)$, giving [3]. As a corollary of these two we get [2].

Now suppose $n = 0$. For the lower bound we consider again the projections $c_1$ and $c_2$. If $v_0(f) < 0$ then $c_1$ must travel up to the $v_0(f)$-th horocycle and then back down to the $0$-th horocycle, thus the length of $c_1$ is bounded below by $2 \max\{-v_0(f), 0\}$. Meanwhile $c_2$ is a path from $i$ to $\gamma \cdot i$, so its length is bounded below by $d_\mathbb{B}(i, \gamma \cdot i)$. Combining these, and also applying Lemma [5], give a lower bound:

$$d_X(x, \gamma x) \geq \frac{1}{2} (\log q - \log \sqrt{2}) \max\{v_0(f), 0\} + \max\{-v_0(f), 0\}.$$

Then taking $A = \min\{\frac{1}{2} (\log q - \log \sqrt{2}), 1\}$ gives the lower bound in [5]. The upper bound comes from considering instead a path $c'$ which consists of a geodesic in $X$ from $x$ to $(B(f, q^0), i)$, then a geodesic in $\mathbb{H}$ from $i$ to $\gamma \cdot i$ (see Figure 4). This path has length $2 \max\{-v_0(f), 0\} + d_\mathbb{B}(i, \gamma \cdot i)$.

**Figure 4.** Two paths from $x$ to $y$, when $n = 0$ and $v_0(f) < 0$. The solid paths shows the geodesic in $X$, while the dotted path is that which is used to obtain the upper bound for $n = 0$ and $v_0(f) < 0$. The solid paths shows the geodesic in $X$, while the dotted path is that which is used to obtain the upper bound for [5].
If instead \( f = 0 \), then we obtain (6) by observing that the geodesic will pass through the points \((B(0, q^k), q^k \cdot i)\) for each \( k = 0, \ldots, n \), and each adjacent pair of such points are distance 1 apart in \( X \).

Finally, (7) is obtained by writing \( \gamma \) as the product:

\[
\gamma = \begin{pmatrix} q^0 & f \\ 0 & 1 \end{pmatrix} \begin{pmatrix} q^n & 0 \\ 0 & 1 \end{pmatrix}
\]

then applying the triangle inequality, using also (5) and (6) \( \Box \).

With the basepoint \( x = (B(0, q^0), i) \) as above, for \( \gamma \in \text{BS}(1, q) \) let \( d_X(x, \gamma x) \) be denoted by \( |\gamma| \).

**Theorem 4.3.** Let \( q \) be an integer greater than or equal to 2 and let \( u, v \) be a pair of elements in the Baumslag–Solitar group \( \text{BS}(1, q) \). Then \( u, v \) are conjugate in \( \text{BS}(1, q) \) if and only if there exists a conjugator \( \gamma \) such that

\[
|\gamma| \leq \frac{2}{\log \sqrt{2}}(|u| + |v|).
\]

**Proof.** Suppose

\[
u = \begin{pmatrix} q^s & P \\ 0 & 1 \end{pmatrix}, \quad v = \begin{pmatrix} q^r & Q \\ 0 & 1 \end{pmatrix}, \quad \gamma = \begin{pmatrix} q^n & f \\ 0 & 1 \end{pmatrix} \]

By direct calculation, \( u\gamma = \gamma v \) if and only if both the following equations hold:

\[
s + n = n + r
\]

\[
P + q^n f = f + t^n Q
\]

As before, we divide the proof into two cases, depending on whether \( r = s \) is zero or not.

**Case 1:** \( r = s = 0 \)

In this case, from equation (9) we get \( P = q^n Q \) and hence \( v_0(P) - n = v_0(Q) \). Lemma 4.2 (5) tells us that

\[
|u| + |v| \geq A |v_0(P)| + A |v_0(P) - n|.
\]

We can put \( f = 0 \) and we therefore have \( |\gamma| = |n| \) by Lemma 4.2 (6). For \( u \) and \( v \), we have the following four possibilities for a lower bound on \( |u| + |v| \):

\[
\begin{array}{ll}
(11) & 2A v_0(P) - An \quad \text{if} \quad v_0(P) \geq 0 \text{ and } v_0(P) \geq n; \\
(12) & An \quad \text{if} \quad v_0(P) \geq 0 \text{ and } v_0(P) \leq n; \\
(13) & -2A v_0(P) + An \quad \text{if} \quad v_0(P) \leq 0 \text{ and } v_0(P) \leq n; \\
(14) & -An \quad \text{if} \quad v_0(P) \leq 0 \text{ and } v_0(P) \geq n.
\end{array}
\]

Cases (12) and (14) are straight-forward since the lower bound we have obtained is equal to \( A |\gamma| \) in both cases.

For case (11), if \( n \leq 0 \) then we use \( 2A v_0(P) - An \geq -An = A |\gamma| \). If \( n \geq 0 \) then we use the fact that in this case \( v_0(P) - n \geq 0 \) to deduce

\[
2A v_0(P) - An \geq A v_0(P) \geq An = A |\gamma|.
\]

Case (13) is analogous to case (11).

Since \( \frac{1}{A} \leq \frac{2}{\log \sqrt{2}} \), we obtain

\[
|\gamma| \leq \frac{2}{\log \sqrt{2}}(|u| + |v|).
\]

**Case 2:** \( r = s \neq 0 \)

By taking inverses of \( u, v \) if necessary, we may assume that \( r > 0 \). We can play the same trick of premultiplying \( \gamma \) by an appropriate power of \( u \) to ensure that \( 0 \leq n < r \). By Lemma 4.2 (7)

\[
|\gamma| \leq n + d_{12}(i, i + f) - 2v_0(f).
\]
By a similar argument as in Theorem 3.3, we can show that
\[ v_0(f) \geq \min\{v_0(P), v_0(Q) + n\} . \]

From (9), we have
\[ f = P - q^n Q \leq q^n - q^r P \leq Q . \]

Using this and the triangle inequality we obtain:
\[ d_{\mathbb{E}}(i, i + f) \leq d_{\mathbb{E}}(i, q^r i + f) + d_{\mathbb{E}}(i, q^r i) \]
\[ = d_{\mathbb{E}}(i, q^r i + f) + d_{\mathbb{E}}(i, q^r i) \]
\[ = d_{\mathbb{E}}(i, q^r i + r) . \]

Hence \(|\gamma| \leq 2r + d_{\mathbb{E}}(i, vi) + 2 \max\{-v_0(P), -v_0(Q) - n\} , \) and since \(n \geq 0\) this leads to:
\[ |\gamma| \leq 2r + d_{\mathbb{E}}(i, vi) + 2 \max\{-v_0(P), -v_0(Q)\} \]
\[ \leq 2r + d_X(x, vx) + 2 \max\{d_X(x, ux), d_X(x, vx)\} \]
\[ \leq 4(|u| + |v|) \]

using Lemma 4.2 (3) and (4). \(\square\)

References

[BD11] Jason Behrstock and Cornelia Drutu. Divergence, thick groups, and short conjugators. arXiv:1110.5005v1 [math.GT], 2011.

[BNW08] Laurent Bartholdi, Markus Neuhauser, and Wolfgang Woess. Horocyclic products of trees. J. Eur. Math. Soc. (JEMS), 10(3):771–816, 2008.

[BCSW11] Alexander Bendikov, Laurent Saloff-Coste, Maura Salvatori, and Wolfgang Woess. The heat semigroup and Brownian motion on strip complexes. Adv. Math., 226(1):992–1055, 2011.

[CKW94] D. I. Cartwright, V. A. Ka˘ımanovich, and W. Woess. Random walks on the affine group of local fields and of homogenous trees. Ann. Inst. Fourier (Grenoble), 44(4):1243–1288, 1994.

[Deh11] M. Dehn. Über unendliche diskontinuierliche Gruppen. Math. Ann., 71(4):116–144, 1911.

[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.

[FM98] Benson Farb and Lee Mosher. A rigidity theorem for the solvable Baumslag-Solitar groups. Invent. Math., 131(2):419–451, 1998. With an appendix by Daryl Cooper.

[FM99] Benson Farb and Lee Mosher. Quasi-isometric rigidity for the solvable Baumslag-Solitar groups. II. Invent. Math., 137(3):613–649, 1999.

[Har81] O. G. Harlampović. A finitely presented solvable group with unsolvable word problem. Izv. Akad. Nauk SSSR Ser. Mat., 45(4):852–873, 928, 1981.

[Hir73] Friedrich E. P. Hirzebruch. Hilbert modular surfaces. Enseignement Math. (2), 19:183–281, 1973.

[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.

[Mat58] 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.

[Mos66] A. Wlodzimierz Mostowski. On the decidability of some problems in special classes of groups. Fund. Math., 59:123–135, 1966.

[Mos99] Lee Mosher. Indiscrete representations, laminations, and tilings. In Geometric group theory down under (Canberra, 1996), pages 225–259. de Gruyter, Berlin, 1999.

[Nos82] B. A. Noskov. On conjugacy in metabelian groups. Mat. Zametki, 31(4):495–507, 653, 1982.
Jean-Philippe Préaux. Conjugacy problem in groups of oriented geometrizable 3-manifolds. *Topology*, 45(1):171–208, 2006.

V. N. Remeslennikov. Conjugacy in polycyclic groups. *Algebra i Logika*, 8:712–725, 1969.

Wolfgang M. Schmidt. Norm form equations. *Ann. of Math. (2)*, 96:526–551, 1972.

Hideo Shimizu. On discontinuous groups operating on the product of the upper half planes. *Ann. of Math. (2)*, 77:33–71, 1963.

Svetla Vassileva. Polynomial time conjugacy in wreath products and free solvable groups. *Groups Complex. Cryptol.*, 3(1):105–120, 2011.

Gerard van der Geest. *Hilbert modular surfaces*, volume 16 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1988.

Wolfgang Woess. Lamplighters, Diestel-Leader graphs, random walks, and harmonic functions. *Combin. Probab. Comput.*, 14(3):415–433, 2005.