Abstract commensurability and the Gupta–Sidki group

Alejandra Garrido
Mathematical Institute, University of Oxford
garridoangul@maths.ox.ac.uk

28th July 2014

Abstract

We study the subgroup structure of the infinite torsion \( p \)-groups defined by Gupta and Sidki in 1983. In particular, following results of Grigorchuk and Wilson for the first Grigorchuk group, we show that all infinite finitely generated subgroups of the Gupta–Sidki 3-group \( G \) are abstractly commensurable with \( G \) or \( G \times G \). As a consequence, we show that \( G \) is subgroup separable and from this it follows that its membership problem is solvable.

Along the way, we obtain a characterization of finite subgroups of \( G \) and establish an analogue for the Grigorchuk group.

1 Introduction

Groups of automorphisms of regular rooted trees have received considerable attention in the last few decades. Their interest derives from the striking properties of some of the first examples studied, the Grigorchuk group \( [6] \) and the Gupta–Sidki \( p \)-groups \( [10] \). These groups are easily understood examples of infinite finitely generated torsion groups, answering the General Burnside Problem. Furthermore, the Grigorchuk group was the first group to be shown to be of intermediate word growth and to be amenable but not elementary amenable (see \( [7] \)). Amenability of the Gupta–Sidki \( p \)-groups (among many other examples) was proved in \( [3] \). These and other striking results prompt one to ask what other unusual properties these groups may have.

In \( [8] \), the authors establish another notable result about the Grigorchuk group, namely that all its infinite finitely generated subgroups are (abstractly) commensurable with the group itself. Recall that two groups are \( (\text{abstractly}) \) \textit{commensurable} if they have isomorphic subgroups of finite index. This notion translates into geometric terms as “having a common finite-sheeted covering”: two spaces which have a common finite-sheeted covering have commensurable fundamental groups. For this reason it is an important concept in geometric group theory. It also appears in the study of lattices in semisimple Lie groups, in profinite groups and other areas of group theory. Having only one commensurability class of infinite finitely generated subgroups is a very strong restriction on subgroup structure and examples where it is known to hold are scarce.
Following a similar general strategy to that in [8], we will show that an analogous result holds for the Gupta–Sidki 3-group:

**Theorem 1.** Every infinite finitely generated subgroup of the Gupta–Sidki 3-group $G$ is commensurable with $G$ or $G \times G$.

This raises two questions. The first is whether the commensurability classes are actually distinct; by contrast, the Grigorchuk group is commensurable with its direct square. This is the motivation for the work carried out in [5]. In that paper, more general results on the structure of subgroups of branch groups yield that the classes are indeed distinct, for many examples of groups acting on $p$-regular trees where $p$ is an odd prime.

The second question concerns the restriction to $p = 3$ in Theorem 1. It seems likely that this restriction is unnecessary. However, our proof relies heavily on a delicate length reduction argument that is only available for $p = 3$.

Our first main theorem allows us to prove another result about the subgroup structure of $G$, also parallel to a result in [5].

**Theorem 2.** The Gupta–Sidki 3-group is subgroup separable and hence has solvable membership problem.

Recall that a group is **subgroup separable** (or LERF) if each of its finitely generated subgroups is an intersection of subgroups of finite index. This condition is strong and is only known to hold in very special cases such as free groups, surface groups, polycyclic groups and some 3-manifold groups (see [12, 18, 14, 13]). It is related to the membership problem (or generalized word problem). The membership problem for a finitely generated group $H$ is solvable if there is an algorithm which given a finitely generated subgroup $K \leq H$ and an element $h \in H$ determines whether or not $h \in K$. It is an easy exercise to show that if a group is subgroup separable and has solvable (ordinary) word problem then it has solvable membership problem. The Gupta–Sidki $p$-groups, the Grigorchuk group and, more generally, groups of “spinal type” have solvable word problem by a fast algorithm first described in [7] (see also [2]).

The proofs of our main theorems both rely on an auxiliary result, Theorem 6, on finitely generated subgroups of the Gupta–Sidki 3-group. It is worth mentioning that all our results hold for the general case where $p$ is any odd prime, except for the length reduction argument in this Theorem 6. For this reason, all definitions and preliminary results are stated for the general case in section 2 and we only focus on the case $p = 3$ for the proof of Theorem 6 in section 4. We depart from preliminary results in section 3, where we discuss maximal subgroups. In [16, 17] it was shown that all maximal proper subgroups of the Grigorchuk and Gupta–Sidki groups have finite index. We will show in Theorem 3 that this maximal subgroups property passes to all finitely generated subgroups of the Gupta–Sidki 3-group, as a consequence of Theorem 1. The proofs of Theorem 1 and Theorem 2 are presented in the final section 5. The arguments generalize easily to the case $p > 3$ assuming that the analogue of Theorem 6 holds. In this final section we also establish the following characterization of finite subgroups of the 3-group, using the analysis carried out in section 4.
Theorem 3. Let $H$ be a finitely generated subgroup of $G$. Then $H$ is finite if and only if no vertex section of $H$ is equal to $G$.

Minor changes in the proof of this theorem and the detailed analysis carried out in [8] yield an identical characterization of the finite subgroups of the Grigorchuk group.

Theorem 4. Let $H$ be a finitely generated subgroup of the Grigorchuk group $\Gamma$. Then $H$ is finite if and only if no vertex section of $H$ is equal to $\Gamma$.

2 Definitions and preliminaries

We begin by defining the trees on which our groups act, the automorphism groups of these trees, and some of their subgroups which will be used in our proofs.

For an integer $d \geq 2$, we may think of the vertices of the $d$-regular rooted tree $T$ as finite words over the alphabet $\{0, \ldots, d-1\}$. We think of the empty word $\varepsilon$ as the root. The words $u, v$ are joined by an edge if $v = uw$ (or $u = vw$) for some $w$ in the alphabet.

We can impose a metric on $T$ by assigning unit length to each edge. Then vertex $v$ will be at distance $n$ from vertex $u$ if the unique path joining them consists of $n$ edges. The distance of a vertex $v$ from the root is the level of $v$. The set of vertices of level $n$ is called the $n$th layer of $T$ and is denoted by $L_n$.

For a vertex $v \in L_n$, the subtree consisting of vertices of level $m \geq n$ separated from the root by $v$ is the subtree rooted at $v$ and it is denoted by $T_v$. Since $T$ is regular, for every vertex $v$ there is an obvious isomorphism from $T$ to $T_v$ taking $u$ to $vu$. We will identify all $T_v$ with $T$ in the rest of the paper.

An automorphism of a rooted tree $T$ is a permutation of the vertices that preserves the adjacency relation. We denote the group of all automorphisms of $T$ by Aut $T$. For any vertex $v$ of $T$ write $v^g$ for the image of $v$ under $g \in$ Aut $T$. We write

$$\text{St}(v) := \{g \in \text{Aut} T \mid v^g = v\}$$

for the stabilizer of $v$. The subgroup

$$\text{St}(n) := \bigcap_{v \in L_n} \text{St}(v)$$

is the $n$th level stabilizer. For any group $\Gamma$ acting on $T$ we denote by $\text{St}_\Gamma(v)$ and $\text{St}_\Gamma(n)$ the intersection of $\Gamma$ with the above subgroups. The level stabilizers $\text{St}_\Gamma(n)$ have finite index in $\Gamma$. We say that $\Gamma$ has the congruence subgroup property if every finite index subgroup of $\Gamma$ contains some $\text{St}_\Gamma(n)$.

For every $x \in \text{St}(v)$ there is a unique automorphism $x_v \in$ Aut $T$ which is simply the restriction of $x$ to the subtree $T_v(= T)$. Hence for every group $\Gamma$ acting on $T$ and every vertex $v$ of $T$ this restriction yields a homomorphism

$$\varphi_v : \text{St}_\Gamma(v) \rightarrow \text{Aut} T, \quad x \mapsto x_v.$$
The image of this homomorphism is denoted by $\Gamma_v$ and called the vertex section of $\Gamma$ at $v$ (some authors call it an upper companion group). We may also refer to $x_v = \varphi_v(x)$ as the vertex section of $x$ at $v$. Observe that if $v = uw$ then $\varphi_v(\text{St}_\Gamma(v)) = \varphi_w(\text{St}_{\Gamma_u}(w)) = \varphi_w \circ \varphi_u(\text{St}_\Gamma(uw))$.

Notice that although the image $\varphi_v(\text{St}_\Gamma(v))$ is a subgroup of $\text{Aut} T$ it may not be a subgroup of $G$. We say that a group $\Gamma$ acting on $T$ is fractal if $\varphi_v(\text{St}_\Gamma(v)) = \Gamma$ for every vertex $v$ of $T$.

For every $n$, define

$$\psi_n : \text{St}_\Gamma(n) \to \prod_{v \in \mathcal{L}_n} \psi_v(\text{St}_\Gamma(v)) \leq \prod_{v \in \mathcal{L}_n} \text{Aut} T_n$$

$$x \mapsto (x_v)_{v \in \mathcal{L}_n}.$$ 

This is an embedding, so we may identify elements of $\text{St}_\Gamma(n)$ with their image under $\psi_n$. For the first level we usually omit the subscript and just write $\psi$. In fact, we often (following the custom in the literature) omit the $\psi$ altogether for elements of the first level stabilizer.

We can now introduce the family of Gupta–Sidki $p$-groups. Let $p > 2$ be a prime and let $T$ be the $p$-regular rooted tree. The Gupta–Sidki $p$-group $G = \langle a, b \rangle$ is the subgroup of $\text{Aut} T$ generated by two automorphisms $a$ and $b$. The automorphism $a$ cyclically permutes the first layer as $(1 \ 2 \ \ldots \ p)$. The element $b = (a, a^{-1}, 1, \ldots, 1, b)$ is recursively defined by its action on subtrees rooted at the first layer: it acts on $T_0$ as $a$ acts on $T$, as $a^{-1}$ on $T_1$, as $b$ on $T_{p-1}$ and trivially on the other subtrees. The figure below shows the action of $b$ on $T$ for $p = 3$:

![Diagram of the action of $b$ on the $p$-regular rooted tree](image)

We establish some preliminary results and properties of the Gupta–Sidki $p$-groups that will be useful later on.

**Notation.** From now on, $G$ will usually denote the Gupta–Sidki $p$-group unless otherwise stated, and we will omit the subscript in $\text{St}_G(n)$, $\text{St}_G(v)$. The derived subgroup of $G$ will be denoted by $G'$ and the direct product of $i$ copies of $G$ will be written $G^{(\times i)}$. 

4
It is immediate from the definition of $G$ as a subgroup of $\text{Aut} T$ that $G$ is residually finite (as $\bigcap_{n=1}^{\infty} \text{St}_G(n) = 1$). Gupta and Sidki already proved that $G$ is just infinite, that is, $G$ is infinite but all its proper quotients are finite ([11]).

For $i = 1, \ldots, p - 1$, we write
\[ b_i := b^{a^i} = (1, \ldots, b, a, a^{-1}, \ldots, 1) \]
where $b$ is in the $i$th co-ordinate and $b_0 := b$. Let $B$ denote $\langle b^G \rangle$, the normal closure of $b$. Then it is easy to see that $B = \langle b^G \rangle$ is equal to $\langle b, b_1, \ldots, b_{p-1} \rangle = \text{St}(1)$. Note also that $b$ has order $p$ (see [10]).

We include the proof of the following to illustrate the use of projection arguments since these will play an important role later on.

**Proposition 2.1 ([10]).** For every $n$, the $n$th level stabilizer $\text{St}(n)$ is a subdirect product of $p^n$ copies of $G$. Therefore $G$ is infinite and fractal.

**Proof.** We proceed by induction. For $n = 1$ the map $\psi : \text{St}(1) \to G^{(\times p)}$ is a homomorphism. We claim that $\varphi_i(\text{St}(1)) = G$ for $i \in \{0, \ldots, p-1\}$. This is easily seen by examining the images of generators of $B$:

\[
\psi(b) = (a, a^{-1}, 1, \ldots, b), \\
\psi(b_1) = (b, a, a^{-1}, 1, \ldots, 1), \\
\vdots \\
\psi(b_{p-1}) = (a^{-1}, 1, \ldots, 1, b, a).
\]

Suppose that the result is true for $n \geq 1$. Then a similar argument as for $\text{St}(1)$ shows that $\text{St}(n+1)$ is a subdirect product of $p$ copies of $\text{St}(n)$, each of them a subdirect product of $p^n$ copies of $G$. This immediately shows that $G$ is fractal and infinite. \qed

The following are easy generalizations to $p > 3$ of results proved in [19] for $p = 3$.

**Proposition 2.2.** We have

(i) $G/G' = \langle aG', bG' \rangle \cong C_p \times C_p$;

(ii) $\psi(B') = (G')^{(\times p)}$;

(iii) $G/B' \cong C_p \wr C_p$.

**Proof.** The first item is a straightforward verification. For the second, we have

\[
\psi([b_0 b_1, b_1^{-1} b_{p-1}]) \\
= [(a, a^{-1}, 1, \ldots, b)(b, a, a^{-1}, \ldots, 1), (b^{-1}, a^{-1}, a, \ldots, 1)(a^{-1}, 1, \ldots, b, a)] \\
= ([ab, b^{-1} a^{-1}], [1, a^{-1}], [a^{-1}, a], \ldots, [1, b], [b, a]) \\
= (1, \ldots, 1, [b, a])
\]
and it is easy to see that, for \( i = 1, \ldots, p - 1 \),
\[
\psi([b_0 b_1, b_1^{-1} b_{p-1}]^a) = \psi([b_i b_{i+1}, b_{i+1}^{-1} b_{p+i-1}]) = (1, \ldots, [b, a], \ldots, 1)
\]

where \([b, a]\) is in the \( i \)th coordinate.

From the above we obtain
\[
G / B' = G / (G')^{(x_p)} = \langle a \rangle B / (G')^{(x_p)} \cong C_p / C_p.
\]

We are now able to show a commensurability property.

**Proposition 2.3.** For every odd prime \( p \), if \( i \equiv j \mod p - 1 \) then \( G^{(x_i)} \) is commensurable with \( G^{(x_j)} \).

**Proof.** For fixed \( i \in \{0, \ldots, p - 1\} \), we show by induction on \( n \) that \( G^{(x_i)} \) is commensurable with \( G^{(x_{n(p-1)+i})} \). For the base case, we deduce from parts [iii] and [iii] of [Proposition 2.2] that \( G^{(x_{p-1+i})} = G^{(x_p)} \times G^{(x_{i-1})} \) is commensurable with \( G \times G^{(x_{i-1})} = G^{(x_i)} \). Now suppose the claim holds for \( n \). Then \( G^{(x_{(n+1)(p-1)+i})} = G^{(x_{n(p-1)+i})} \times G^{(x_{(x_{i-1}+1})}} \) is commensurable with \( G^{(x_{i-1})} \times G^{(x_{(p-1)})} \), which is in turn commensurable with \( G^{(x_{i-1})} \times G = G^{(x_i)} \). Hence \( G^{(x_{(n+1)(p-1)+i})} \) is commensurable with \( G^{(x_i)} \) and our claim follows by induction.

The following will be very useful.

**Proposition 2.4.** An element \( g \in G \) is in \( B \) if and only if there exist \( n_0, \ldots, n_{p-1} \in \{0, \ldots, p-1\} \) and \( c_0, \ldots, c_{p-1} \in G' \) such that
\[
g = (a^{-n_{p-1}+n_0} c_0, a^{-n_{p-1}+n_1} c_1, \ldots, a^{-n_{p-1}+n_{p-1}} c_{p-1}).
\]
This representation is unique.

**Proof.** The ‘if’ direction is obvious. If \( g \in B \) then \( g \) is a product of \( b_0, \ldots, b_{p-1} \) and
\[
B / B' = \{b_0^{r_0} b_1^{r_1} \cdots b_{p-1}^{r_{p-1}} \mid r_i \in \{0, \ldots, p-1\}\}.
\]

Thus for some \( n_i \in \{0, \ldots, p-1\} \), \( c \in B' \) we have
\[
g = b_0^{n_0} b_1^{n_1} \cdots b_{p-1}^{n_{p-1}} c
= (a^{n_0} b_1 a^{-n_{p-1}} d_0, a^{-n_0+n_1} b_2 d_1, \ldots, b^{n_0} a^{-n_{p-1}+n_{p-1}} d_{p-1})
= (a^{n_0} a^{-n_{p-1}+n_0} c_0, a^{-n_0+n_1} b_2 d_1, \ldots, b^{-n_{p-1}+n_{p-1}} a^{n_0} c_{p-1})
\]

where \( c_i, d_i \in G' \) (since \( B' = (G')^{(x_p)} \)).

The \( n_i \) are uniquely determined and therefore so are the \( c_i \).

**Lemma 2.5.** The following assertions hold:
(i) $b^{(1)} = (b, b, \ldots, b) \in G'$;

(ii) $G' \geq \text{St}(2)$.

Proof. For the first item note that $bb_1 \cdots b_{p-1} \equiv b^{(1)} \mod B'$ since $bb_1 \cdots b_{p-1} = (ab^{-1}, b, \ldots, b)$.

Therefore it suffices to show that $bb_1 \cdots b_{p-1} \in G'$. To see this, observe that for $i = 0, \ldots, p - 1$ we have $b_i = b[b, a^i]$ so that

$$bb_1 \cdots b_{p-1} = b^2[b, a]b^{-2}b^3[b, a^2]b^{-3} \cdots b^{-1}[b, a^{-2}]b[b, a^{-1}] \in G',$$

as required.

For the second item, let $g = (g_0, \ldots, g_{p-1}) \in \text{St}(2)$, so $g_i \in \text{St}(1)$ for each $i$. Proposition 2.4 implies that $g = (b^n c_0, \ldots, b^n c_{p-1}) \equiv (b^{(1)})^n \mod B'$ for some $n$. Thus $g \in G'$ by the previous part. \hfill \Box

The next result is [1, Proposition 8.4] (for $p = 3$), but the proof given there is not quite clear to us.

**Proposition 2.6.** The Gupta–Sidki $p$-group $G$ has the congruence subgroup property.

Proof. By [1, Proposition 3.8], it suffices to show that $G''$ contains some level stabilizer $\text{St}(m)$. The case $p > 3$ follows from [9, Lemma 2], which shows that $G'' \geq G' \times \cdots \times G'$, the direct product of $p^2$ copies of $G'$. Thus, by Lemma 2.5 we have $G'' \geq \text{St}(2) \times \cdots \times \text{St}(2) \geq \text{St}(4)$.

For the case $p = 3$ we must prove an analogous version of [9, Lemma 2]. Let $C = [G', G]$ be the third term of the lower central series of $G$. Since $G'$ contains $B' = G' \times G' \times G'$ and the elements $[a, b]^{a^{-1}} = (a, ab, b^{-1}a)$, $b^{(1)} = (b, b, b)$, we have

$$(x, a], 1, 1) \in G''$$

and $([x, b], 1, 1) \in G''$ for any $x \in G'$. Therefore $G'' \geq C \times \{ 1, 1 \}$ as $C = \{ [x, a], [x, b] \mid x \in G' \}^G$ and $G'' \leq \text{St}(1)$. Conjugating by suitable powers of $a$ we obtain that $G'' \geq C \times C \times C$. Now, $[b^{-1}, a], b_1 b_2] = (1, [a, b], 1) \in C$ and from this we conclude that $C \geq G' \times G' \times G'$. Thus $G'' \geq \text{St}(4)$ as above. \hfill \Box

This property will be important for the proofs of Theorem 1 and Theorem 5 but it is also obviously useful for the study of finite quotients of $G$. In this direction, we point out that in [9] the authors give an explicit formula for the indices $|\Gamma : \text{St}_G(n)|$ not just for the Gupta–Sidki $p$-group, but a more general class of groups $\Gamma$ which act on $p$-regular rooted trees ($GGS$ groups). They also prove, using different methods, some of the properties stated in the previous lemmas for arbitrary GGS groups.

**Lemma 2.7.** Let $g = (g_0, \ldots, g_{p-1}) \in B$. If $g_i G' \in \langle g_0 \rangle G'$ for every $i$ then $g_0 G' = b^r G'$ for some $r$. 

7
Proof. It suffices to prove the result when \( g_0 G' \in \langle ab^k \rangle G' \) for some \( k \). Writing \( y = ab^k \), we obtain

\[
gB' = (g_0, \ldots, g_{p-1})B' \\
= (y^{r_0}, \ldots, y^{r_{p-1}})B' \\
= (a^{-n_{p-1}+n_0}b^{r_1}, a^{-n_0+n_{p-1}}b^{r_2}, \ldots, a^{-n_{p-2}+n_{p-1}}b^{r_0})B'
\]

from Proposition 2.4. This gives the equations

\[
\begin{align*}
r_0 &= n_0 - n_{p-1}, & kr_0 &= n_1, \\
r_1 &= n_1 - n_0, & kr_1 &= n_2, \\
\vdots &= \vdots \\
r_{p-1} &= n_{p-1} - n_{p-2}, & kr_{p-1} &= n_0
\end{align*}
\]

which can be rewritten as

\[
\begin{align*}
r_0 &= kr_{p-1} - kr_{p-2}, \\
r_1 &= kr_0 - kr_{p-1}, \\
\vdots &= \vdots \\
r_{p-1} &= kr_{p-2} - kr_{p-3}.
\end{align*}
\]

We can then perform Gaussian elimination on the augmented matrix corresponding to this system to obtain

\[
\begin{pmatrix}
1 & 0 & \ldots & 0 & k & -k & 0 \\
-k & 1 & 0 & \ldots & 0 & k & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & \ldots & 0 & k & -k & 1 & 0
\end{pmatrix}
\leftrightarrow
\begin{pmatrix}
1 & 0 & \ldots & 0 & k & -k & 0 \\
0 & 1 & 0 & \ldots & k^2 & k - k^2 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & \ldots & 0 & 0 & 0 & 1 & 0
\end{pmatrix}.
\]

Thus \( r_0 = r_1 = \cdots = r_{p-1} = 0 \) is the unique solution of this system, forcing \( g_0 \in G' \), as required.

We can strengthen this to the following lemma which will be essential in the proof of Theorem 6.

**Lemma 2.8.** Let \( H \) be a subgroup of \( G \) which is not contained in \( St(1) \). Then either all first level vertex sections \( H_u \) of \( H \) are equal to \( G \), or they are all contained in \( St(1) \) so that \( St_H(1) = St_H(2) \).

**Proof.** We first show that all first level vertex sections of \( H \) are conjugate in \( G \). Since \( H \) is not contained in \( St(1) \) there exists \( s \in H \) with \( s = (s_0, \ldots, s_{p-1}) \in St(1) \). Thus, for every \( h = (h_0, \ldots, h_{p-1}) \in St_H(1) \) we have \( h^{as} = (h_0^{s_0}, h_1^{s_1}, \ldots, h_{p-1}^{s_{p-1}}) \in St_H(1) \). From this we see that \( H_i = H_i^{s_i} \) for \( i = 0, \ldots, p-1 \) and our claim follows.

Hence we may assume that no first level vertex section of \( H \) is equal to \( G \). We examine the image of \( H \) modulo \( B' \) (or equivalently, the images of its first level vertex
sections modulo \( G' \). Since all these vertex sections are conjugate in \( G \), they must have the same images modulo \( G' \). In particular, for each \( h \in \text{St}_H(1) \), the hypothesis of Lemma 2.7 holds and the result follows.

The proof of Theorem 1 relies on a word length reduction argument. Instead of the usual word length for finitely generated groups, we use a length function which only takes into account the number of conjugates of \( b \).

Let \( g \in G \) and let \( a^{s_0}b^{r_1}a^{s_1} \cdots b^{r_m}a^{s_m} \) be a shortest (in the usual sense) word in \( \{ a, b \} \) representing \( g \). We can rewrite this word as

\[ b^{r_1} \cdots b^{r_m}a^v \]

where \( b_{i_j} \neq b_{i_{j+1}} \) for each \( j \) and \( v \in \{0, \ldots, p-1\} \).

Define the length \( l(g) \) of \( g \) to be \( m \). Thus \( l(g) \) is the number of conjugates of \( b \) in a shortest word in \( a, b \) representing \( g \).

The following easy lemma shows how the length of elements of \( G \) is reduced as we project down levels of the tree. Notice that this holds for any odd prime \( p \).

**Lemma 2.9.** Let \( g \in \text{St}(1) \) and suppose that \( l(g) = m \). Then

(i) \( l(\varphi_k(g)) \leq \frac{1}{2}(m+1) \) and

(ii) if \( g \in \text{St}(2) \) then \( l(\varphi_j(\varphi_k(g))) \leq \frac{1}{4}(m+3) \)

for \( k, j \in \{0, \ldots, p-1\} \).

**Proof.** Since \( g \in \text{St}(1) \), we can write it in the form \( g = b_{i_1}^{r_1} \cdots b_{i_m}^{r_m} \) for some \( i_j, r_j \in \{0, \ldots, p-1\} \). Now, the image \( \varphi_k(g) \) for each \( k \in \{0, \ldots, p-1\} \) is, at worst, of one of the following forms:

1. \( a^{r_1}b^{r_2} \cdots a^{r_m} \), so \( l(\varphi_k(g)) \leq \frac{m-1}{2} \);
2. \( a^{r_1}b^{r_2} \cdots a^{r_{m-1}}b^{r_m} \), so \( l(\varphi_k(g)) \leq \frac{m}{2} \);
3. \( b^{r_1}a^{r_2} \cdots b^{r_{m-1}}a^{r_m} \), so \( l(\varphi_k(g)) \leq \frac{m}{2} \);
4. \( b^{r_1}a^{r_2} \cdots b^{r_m} \), so \( l(\varphi_k(g)) \leq \frac{m+1}{2} \).

This proves the first item.

It now follows that for any \( g \in \text{St}(2) \) we have

\[ l(\varphi_j(\varphi_k(g))) \leq \frac{1}{2} \left( \frac{m+1}{2} + 1 \right) = \frac{1}{4}(m+3), \]

for every \( j, k \in \{0, \ldots, p-1\} \).
3 Maximal subgroups

In this section we establish some results about groups whose maximal subgroups all have finite index. Once Theorem 1 is proved, these results will show that, for each finitely generated subgroup of the Gupta–Sidki 3-group, all maximal subgroups have finite index. The results and proofs are analogous to those in [8].

Notation. We will write $H \leq_f \Gamma$ and $H \leq_s \Gamma_1 \times \cdots \times \Gamma_n$ to mean, respectively, that $H$ is a finite index subgroup of $\Gamma$ and that $H$ is a subdirect product of $\Gamma_1 \times \cdots \times \Gamma_n$.

Lemma 3.1. Let $\Gamma$ be an infinite finitely generated group and $H$ a subgroup of finite index in $\Gamma$. If $\Gamma$ has a maximal subgroup $M$ of infinite index then $H$ has a maximal subgroup of infinite index containing $H \cap M$.

Proof. We first show that any proper subgroup of $H$ containing $H \cap M$ must be of infinite index in $H$. Suppose for a contradiction that there is a proper finite index subgroup $L$ of $H$ containing $H \cap M$. Then $K$, the normal core of $L$ in $\Gamma$, is of finite index in $\Gamma$ and therefore $KM \leq \Gamma$ must be of finite index too. Now, $M$ is a maximal subgroup of $\Gamma$ contained in $KM$, so either $M = KM$ or $KM = \Gamma$. If the former holds, then $M$ is of finite index in $\Gamma$, a contradiction. If the latter is true, we obtain $H = K(M \cap H) \leq L$, contradicting the assumption that $L$ is a proper subgroup of $H$.

Since $H$ is finitely generated, every proper subgroup is contained in a maximal subgroup (this can be shown without using Zorn’s Lemma, see [15]) and, by the above, the maximal subgroup containing $H \cap M$ must be of infinite index in $H$. \qed

Recall that a chief factor of a group $\Gamma$ is a minimal normal subgroup of a quotient group of $\Gamma$.

Lemma 3.2 ([8], Lemma 3). Let $\Gamma_1, \cdots, \Gamma_n$ be groups with the properties that all chief factors are finite and all maximal subgroups have finite index. If $\Delta \leq_s \Gamma_1 \times \cdots \times \Gamma_n$ then all chief factors of $\Delta$ are finite and all maximal subgroups of $\Delta$ have finite index.

Theorem 5. Let $\Gamma := G^{(x,k)}$ be the direct product of $k$ copies of $G$. If $H$ is a group commensurable with $\Gamma$ then all maximal subgroups of $H$ have finite index in $\Gamma$.

Proof. By definition of commensurability, there exist $K \leq_f H$ and $J \leq_f \Gamma$ with $K$ isomorphic to $J$. For $i = 1, \ldots, k$, let $G_i$ denote the $i$th direct factor of $\Gamma$. Then for every $i$ the subgroup $J_i := J \cap G_i$ has finite index in $G_i$. As $G$ has the congruence subgroup property, for each $i$ there is some $n_i$ with $1 \times \cdots \times \text{St}(n_i) \times \cdots \times 1 \leq J_i$ and so $S := \text{St}(n_1) \times \cdots \times \text{St}(n_k) \leq_f J$.

Now, $G$ is residually finite and just infinite and so all of its chief factors are finite. Hence, by the main result in [14], all of its maximal subgroups have finite index. Furthermore, we saw in Proposition 2.1 that $\text{St}(n)$ is a subdirect product of $p^n$ copies of $G$. Thus $S$ satisfies the assumptions of Lemma 3.2 and all of its maximal subgroups are of finite index. Lemma 3.1 then implies that all maximal subgroups of $J$ must have finite index in $J$. The result now follows on applying Lemma 3.1 to $K \cong J$ and $H$. \qed
4 The case $p = 3$: a key theorem

**Notation.** From now on we restrict our attention to the case $p = 3$, so $G$ will denote the Gupta–Sidki 3-group. Recall that $H_v$ denotes the vertex section of a subgroup $H \leq \text{Aut} T$ at vertex $v$; that is, $H_v = \varphi_v(\text{St}_H(v)) = \varphi_{u_n} \circ \cdots \circ \varphi_{u_1}(\text{St}_H(v))$ where $v = u_1 \ldots u_n$ is considered as a string of letters $u_i \in \{0, \ldots, p - 1\}$.

**Theorem 6.** Let $X$ be a family of subgroups of $G$ satisfying

(I) $1 \in X$, $G \in X$;

(II) if $H \in X$ then $L \in X$ for all $L \leq G$ such that $H \leq L$;

(III) if $H$ is a finitely generated subgroup of $\text{St}(1)$ and all first level vertex sections of $H$ are in $X$ then $H \in X$.

Then all finitely generated subgroups of $G$ are in $X$.

**Proof.** Note that if $X$ satisfies properties (I)–(III) then so does the subfamily $\{ H \mid H^g \in X \text{ for all } g \in G \}$. We may therefore replace $X$ by this subfamily and assume that if $H \in X$ so is every $G$-conjugate of $H$.

Suppose for a contradiction that there are finitely generated subgroups of $G$ which are not in $X$. Choose among them some subgroup $H$ generated by a finite set $S$ such that $D = \max \{ l(s) \mid s \in S \}$ is as small as possible.

If $H \leq \text{St}(1)$ then by (III) at least one of the first level vertex sections of $H$ is not in $X$ and the generating set of this vertex section has elements of length at most $\frac{1}{2}(D + 1) < D$ by Lemma 2.9, contradicting the choice of $H$.

Therefore $H$ is not contained in $\text{St}(1)$. We will show that there exists some $v \in L_2$ such that the vertex section $H_v$ is not in $X$ and has a generating set consisting of elements of length less than $D$.

Since $\text{St}_H(1)$ is a finitely generated subgroup of $\text{St}(1)$, (III) implies that not all first level vertex sections of $\text{St}_H(1)$ are in $X$. However, if one of them is in $X$ then, as all first level vertex sections of $H$ are conjugate in $G$, they must all be in $X$. Thus no first level vertex section of $H$ is in $X$; in particular, none of them is equal to $G$, and Lemma 2.8 asserts that they are all contained in $\text{St}(1)$. For each $k \in \{0, 1, 2\}$, property (III) again implies that one of $H_{k0}, H_{k1}, H_{k2}$ is not in $X$. We claim that for some $k$ every such vertex section is generated by elements of length less than $D$.

Pick some element $t \in S \setminus \text{St}(1)$. Then, as $\text{St}_H(1)$ has index 3 in $H$, the set $T := \{ 1, t, t^{-1} \}$ is a Schreier transversal to $\text{St}_H(1)$ in $H$. Consequently, $\text{St}_H(1) = \text{St}_H(2)$ is generated by

$$X = \{ t_1 s t_2^{-1} \mid t_i \in T, s \in S, t_1 s t_2^{-1} \in \text{St}_H(1) \}.$$  

The elements of this set have length at most $3D$, so the second level vertex sections of $H$ are generated by elements of length at most $(3D + 3)/4$, by Lemma 2.9. Our claim follows if $D > 3$. For $D = 3$ and $D = 2$ the claim follows from Lemma 4.3 and Lemma 4.2 respectively, while Lemma 4.1 shows that if $D = 1$ then $H \in X$. \hfill \Box
Lemma 4.1. Let $H$ be a subgroup of $G$ with a finite generating set $S$ consisting of elements of length at most 1. Then $H \in \mathcal{X}$.

Proof. Any $s \in S$ must be of the form $a^k$, $b_j^q$, $b_i^r a^k$ where $k, r \in \{1, 2\}$ and $i \in \{0, 1, 2\}$. If $S$ consists of only one element then $H$ is finite and therefore in $\mathcal{X}$ by (I) and (II). Thus $S$ must contain at least two elements and they clearly cannot all be of the form $a^k$.

If $S$ contains an element of the form $a^k$ and an element of any other form then $H = G \in \mathcal{X}$. Suppose that all elements of $S$ are of the form $b_j^q$. Then either $H = \langle b_0, b_1, b_2 \rangle = \text{St}(1)$, so $H \in \mathcal{X}$ by (I) and Proposition 2.1 or $H = \langle b_i^r, b_j^q \rangle < \text{St}(1)$ for $i \neq j \in \{0, 1, 2\}$ and $r, q \in \{1, 2\}$, so two of the first level vertex sections of $H$ are $G$ and the other one is $(a)$; therefore $H \in \mathcal{X}$ by (III).

Suppose that $S$ contains an element of the form $b_i^r a$ (we may assume that the power of $a$ is 1 as $(b_i^r a)^{-1} = b_{i-1}^{-1} a^2$). If there is some $b_j^q \in S$ such that $j = i$ then $b_j^q = a^q$ in the $i$th coordinate; but $(b_i^r a)^3 = b_i^r b_{i-1}^q b_{i+1}^q = (b_i^r, b_i^q)^{a_i}$ has $b^r$ in the $i$th coordinate, so the vertex section $H_i$ of $H$ at vertex $i$ is $G \in \mathcal{X}$. Since $H \not\subset \text{St}(1)$, the first level vertex sections of $H$ are conjugate in $G$, hence all first level vertex sections of $H$ are in $\mathcal{X}$ and $H \in \mathcal{X}$.

Finally, suppose that all elements of $S$ are of the form $b_i^r a$, so $S$ contains elements $b_i^r a, b_j^q a$ and $b_i^r a (b_j^q a)^{-1} \in H$. If $i = j$ then $a \in H$ and $H = G \in \mathcal{X}$. If $i \neq j$ then

\[ b_i^r a (b_j^q a)^{-1} = (a^r, a^{-1}, b^s)^{a_i} (a^q, a^q, b^{-q})^{a_j} \]

has $b^r a^q$ in the $i$th coordinate while $(b_i^r a)^3 \in H$ has $b^r$ in the $i$th coordinate. Thus $H_i = G$ and $H \in \mathcal{X}$ by the argument in the previous paragraph.

Lemma 4.2. Let $H \not\subset \text{St}(1)$ be a subgroup of $G$ with a finite generating set $S$ consisting of elements of length at most 2 and such that $H_u \subset \text{St}(1)$ for all $u \in \mathcal{L}_1$. Then $H_u$ is generated by elements of length at most 1 for all $u \in \mathcal{L}_2$.

Proof. First note that no generator of $H$ is in $\text{St}(1)$ since the only possibilities are elements of the form $b_i^r$ and $b_i^r b_j^q$, which are not in $\text{St}(2)$.

If $S$ has an element $t$ of length less than 2 then, as in the proof of Theorem 6, the set $T = \{1, t, t^{-1}\}$ is a transversal of $\text{St}_H(1)$ in $H$ and $\text{St}_H(1)$ is generated by the set $X = \{t_1 s t_2^{-1} \mid t_i \in T, s \in S, t_1 s t_2^{-1} \in \text{St}_H(1)\}$. The elements of $X$ have length at most 4, so by Lemma 2.9 all second level vertex sections of $H$ will be generated by elements of length at most 1.

Suppose then that all elements of $S$ have length 2, that is, they are all of the form $b_i^r b_j^q a$. It suffices to consider only elements of this form as $(b_i^r b_j^q a)^{-1} = b_{i+1}^{-1} b_{i-1}^r a^2$. Pick some $t = b_{i_1}^{r_1} b_{j_2}^{q_2} a \in S$ to form the transversal $T$ so that $\text{St}_H(1)$ is generated by $X$ as above. Then every element of $X$ is of the form $t^{-1} s, s t^{-1}, t^3$ or $t s t$ where $s = b_{i_1}^{q_1} b_{j_2}^{q_2} a \in S$.

We show that all possible combinations of $i_1, i_2, j_1, j_2, r_1, r_2, q_1, q_2$ give rise to generators of second level vertex sections of $H$ of length at most 1.
The forms $t^{-1}s = b_{i_2}^{-r_2}b_{i_1}^{-r_1}b_{i_0}^{q_1}b_{i_2}^{q_2}$ and $st^{-1} = b_{i_1}^{q_1}b_{i_2}^{-r_2}b_{i_1}^{-r_1}$ yield elements of length at most 4. Thus, by Lemma 2.9, their second level vertex sections have length at most 1.

Since $i_1 \neq i_2$, an element of the form $t^3 = b_{i_1}^{r_1}b_{i_2}^{r_2}b_{i_1}^{q_1}b_{i_2}^{q_2}$ will have at most two separate instances of each of $b_0, b_1, b_2$. Hence its first level vertex sections will have length at most 2 and so the second level vertex sections have length at most 1.

More care is required for the form $tst = b_{i_1}^{r_1}b_{i_2}^{r_2}b_{i_1}^{q_1}b_{i_2}^{q_2}$, easy combinatorial arguments (using that $i_1 \neq i_2$ and $j_1 \neq j_2$) show that the first level vertex sections of an element of this form cannot have length greater than 3 and that the only way they can have length 3 is if $i_2 = i_1 - 1$ and either $j_1 = i_1 + 1$ or $j_2 = i_1 + 1$. For ease of exposition, assume without loss of generality that $i_1 = 0$. Then

$$tyt = b_{i_0}^{r_1}b_{i_2}^{r_2}b_{i_0}^{q_1}b_{i_2}^{q_2} = b_{i_0}^{r_1}b_{i_2}^{r_2}b_{i_0}^{q_1}b_{i_2}^{q_2}, \quad j_1 - 1 = i_1$$

The vertex sections at vertices 0 and 1 have length at most 2, while the one at vertex 2 looks like

$$b_{i_1}^{r_1}a^{q_1}a^{-r_1}b_{i_2}^{r_2}, \quad j_1 \neq i_1 = j_2;$$

$$b_{i_0}^{r_1}a^{q_1}a^{-r_1}b_{i_2}^{r_2}, \quad j_1 \neq i_1, j_2 \neq i_1 - 1;$$

$$b_{i_0}^{r_1}a^{q_1}a^{-r_1}b_{i_2}^{r_2}, \quad j_2 \neq i_1 = j_1;$$

$$b_{i_0}^{r_1}a^{q_1}a^{-r_1}b_{i_2}^{r_2}, \quad j_2 \neq i_1, j_1 \neq i_1 - 1.$$

In the first and third cases we have $\varphi_2(t^{-1}s) = b^{-r_2}a^{q_1}a^{-r_1}b^{r_2}$ (as $H_2 \leq S(1)$ by assumption); while, in the second and fourth cases we have $\varphi_2(st^{-1}) = a^{q_1}a^{q_2}a^{-r_2}b^{-r_1} = b^{-r_1}$. Thus, in all cases the vertex section of $H$ at vertex 22 is $G$, which is indeed generated by elements of length at most 1.

**Lemma 4.3.** Let $H \neq St(1)$ be a subgroup of $G$ with a finite generating set $S$ consisting of elements of length at most 3 and such that $H_\alpha \leq S(1)$ for all $\alpha \in \mathcal{L}_G$. Then there exists $v \in \mathcal{L}_2$ such that $H_v$ is generated by elements of length less than 3.

**Proof.** If there exists $t \in S \setminus S(1)$ with $l(t) \leq 2$ then, by the same argument as in the proof of Lemma 4.2, the generating set with $H(t)$ consists of elements of length at most 7, the second level vertex sections of which have length strictly less than 3.

If all generators in $S \setminus S(1)$ have length 3, pick one of them, $t$, to form the transversal $T$. Assume that $t$ has the form $b_{i_1}^{r_1}b_{i_2}^{r_2}b_{i_3}^{r_3}$ for $i_1, i_2, i_3 \in \{0, 1, 2\}$, $i_2 \neq i_1, i_2 \neq i_3$ and $r_1, r_2, r_3 = \pm 1$. It suffices to consider elements of this form since $t^{-1} = b_{i_3}^{r_3}b_{i_2}^{-r_2}b_{i_1}^{-r_1}a$. We may pick $k \in \{0, 1, 2\}$ such that neither $\varphi_k(b_{i_3}+1)$ nor $\varphi_k(b_{i_3})$ is a power of $b$. Then the elements of $X$ have length no more than 9 and the choice of $k$ ensures that $\varphi_k(x) = a^{\pm 1} \varphi_k(bz)$ for each $x \in X$ where $l(bz) \leq 8$. Hence the vertex sections $H_{v_j}$ for $j \in \{0, 1, 2\}$ are generated by elements of length at most $(8 + 3)/4 < 3$, as required. 

\[\]
5 The case $p = 3$: proof of main theorems

Using the length reduction arguments and results in the previous section, we easily obtain a characterization of the finite subgroups of $G$.

**Theorem 3.** Let $H$ be a finitely generated subgroup of $G$. Then $H$ is finite if and only if no vertex section of $H$ is equal to $G$.

*Proof.* For the non-trivial implication, we make the crucial observation that for each $v \in T$, every vertex section of $H_v$ is a vertex section of $H$. Let $H$ be generated by a finite set $S$. We proceed by induction on $D$, the maximum length of elements in $S$. If $D = 1$, then by the proof of [Lemma 4.1] either $H$ is finite or $H_u = G$ for every $u \in L_1$.

Assume that the theorem holds whenever $D \leq n$ with $n \geq 1$. For $D = n + 1$ we consider the cases $H \leq \text{St}(1)$ and $H \not\leq \text{St}(1)$ separately. If $H \leq \text{St}(1)$ then each first level vertex section $H_u$ of $H$ is generated by elements of length at most $(D + 1)/2 = (n + 2)/2 < D$, so that $H_u$ is finite by inductive hypothesis. Thus $H$ itself must be finite as the map $\psi : H \to H_0 \times H_1 \times H_2$ is an injective homomorphism.

If $H \not\leq \text{St}(1)$ there are three cases to consider depending on the vertex section $H_0$ at vertex 0. If $H_0 \leq \text{St}(1)$ then, by Case 3 in the proof of [Theorem 6], $H_0$ and

An identical characterization of finite subgroups of the Grigorchuk group $\Gamma = \langle a, b, c, d \rangle$ is obtained using the same methods as in the above proof and the analysis carried out in [8, Theorem 3]. For the reader’s convenience, we sketch a proof of the non-trivial implication. We use the same length function as in [8], namely, the usual word length for finitely generated groups.

**Theorem 4.** Let $H \leq \Gamma$ be finitely generated. Then $H$ is finite if and only if no vertex section of $H$ is equal to $\Gamma$.

*Proof.* Let $H$ be generated by a finite set $S$ such that $1 \in S$ and $S^{-1} = S$ and let $D$ be the maximum length of elements of $S$. We induct on $D$.

If $D = 1$ then either $H$ is finite or $H = \Gamma$, a contradiction.

Assume that the theorem holds whenever $D \leq n$ with $n \geq 1$. For $D = n + 1$ we consider the cases $H \leq \text{St}(1)$ and $H \not\leq \text{St}(1)$ separately. If $H \leq \text{St}(1)$ then each first level vertex section $H_u$ of $H$ is generated by elements of length at most $(D + 1)/2 = (n + 2)/2 < D$, by [8, Lemma 7], so that $H_u$ is finite by inductive hypothesis. Thus $H$ itself must be finite as the map $\psi : H \to H_0 \times H_1 \times H_2$ is injective.

If $H \not\leq \text{St}(1)$ there are three cases to consider depending on the vertex section $H_0$ at vertex 0. If $H_0 \leq \text{St}(1)$ then, by Case 3 in the proof of [8, Theorem 3], $H_0$ and
Proof. We begin with the easy cases. If all subgroups so in Lemma 5.1. If \( G \) is commensurable with \( G \) with finite index. Thus \( L \) is finite, making \( 00 \) is finite. As \( H \) acts transitively on the second layer of the tree, there exists \( h = (a(s_0, s_1), h_1) \in H \) with \( s_0, s_1 \in \Gamma \) swapping the vertices 00 and 01. For any \( g = ((g_{00}, g_{01}), g_1) \in \text{St}_H(00) \) we have \( g^h = ((g_{01}^s, g_{00}^s), g_1^h) \), whence \( H_{00}^s = H_{01} \). Hence \( H_{01} \) is finite, making \( H_0 \) and therefore \( H \) finite.

If \( H_0 \Gamma' = \langle ad \rangle \Gamma' \) or \( \langle a, d \rangle \Gamma' \), then \( H_{000} \leq \text{St}_\Gamma(1) \) by [8] Lemma 6]. Case 2 of the proof of [8] Theorem 3] shows that \( H_{000} \) and \( H_{001} \) are generated by elements of length less than \( D \) so that \( H_{00} \) is finite. As \( H \) acts transitively on the second layer of the tree, there exists \( h = (a(s_0, s_1), h_1) \in H \) with \( s_0, s_1 \in \Gamma \) swapping the vertices 00 and 01. For any \( g = ((g_{00}, g_{01}), g_1) \in \text{St}_H(00) \) we have \( g^h = ((g_{01}^s, g_{00}^s), g_1^h) \), whence \( H_{00}^s = H_{01} \). Hence \( H_{01} \) is finite, making \( H_0 \) and therefore \( H \) finite.

If \( H_0 \Gamma' = \langle ac \rangle \Gamma' \) or \( \langle a, c \rangle \Gamma' \), then \( H_{0000} \leq \text{St}_\Gamma(1) \) by [8] Lemma 6]. Case 1 of the proof of [8] Theorem 3] shows that \( H_{0000} \) and \( H_{0001} \) are generated by elements of length less than \( D \) so that \( H_{000} \) is finite. Since \( H \) acts transitively on the third layer, there is an element \( h \in H \) swapping the vertices 000 and 001. Then \( H_{0000} \) and \( H_{0001} \) are conjugate in \( \Gamma \), by an argument similar to the one above. Thus \( H \) is finite by the arguments in the previous case and the theorem follows by induction.

We now move on to the proof of the two main results.

Theorem 1. Every infinite finitely generated subgroup of the Gupta–Sidki 3-group \( G \) is commensurable with \( G \) or the direct square \( G \times G \).

To prove this theorem it suffices to show that properties (I)–(III) in Theorem 6 hold for the class \( \mathcal{C} \) of subgroups of \( G \) which are finite, commensurable with \( G \) or commensurable with \( G \times G \). Clearly, (I) holds as the trivial subgroup and \( G \) are in this class. It is also easy to see that (II) is satisfied by \( \mathcal{C} \) if \( H \in \mathcal{C} \) is commensurable with \( G \) or \( G \times G \) and \( J \leq_f H \) is a subgroup isomorphic to a finite index subgroup of \( G \) or \( G \times G \) then, for any \( L \leq G \) containing \( H \) as a finite index subgroup, \( J \) is also contained in \( L \) with finite index. Thus \( L \in \mathcal{C} \). If \( H \in \mathcal{C} \) is finite then any \( L \) containing \( H \) with finite index must also be finite, so \( L \in \mathcal{C} \) too. Thus it only remains to show that \( \mathcal{C} \) satisfies

(III) If \( H \) is a finitely generated subgroup of \( \text{St}(1) \) and all first level vertex sections of \( H \) are in \( \mathcal{C} \) then \( H \in \mathcal{C} \).

This is what we will prove in the next lemma, using similar ideas to those in [8]. To simplify notation, we will write \( \mathcal{G} \) to mean \( \text{‘G or } G \times G \text{’} \).

Lemma 5.1. If \( H \leq_s H_1 \times H_2 \times H_3 \) where each direct factor \( H_i \) is in \( \mathcal{C} \) then \( H \in \mathcal{C} \).

Proof. We begin with the easy cases. If all subgroups \( H_i \) are finite then \( H \) is finite so in \( \mathcal{C} \). If two of the \( H_i \) are finite and the other is commensurable with \( \mathcal{G} \) then \( H \) is commensurable with \( \mathcal{G} \) and is therefore in \( \mathcal{C} \).

If only one of the direct factors, say \( H_3 \), is finite and the other two are commensurable with \( \mathcal{G} \), we reduce to the case where \( H_1 \leq_f \mathcal{G} \) and \( H_2 \) is a subdirect product of finitely many copies of \( G \) as follows. Let \( L_1 \leq_f H_1 \) such that \( L_1 \cong L_1 \leq_f \mathcal{G} \) and define \( L := H \cap (L_1 \times H_2 \times H_3) \). Then we may replace \( H, H_1, H_2 \) and \( H_3 \) by \( L, L_1 \) and the projections of \( L \) in \( H_2 \) and \( H_3 \). This way we can assume that \( H_1 \cong L_1 \leq_f \mathcal{G} \) and then we may as well ignore the isomorphism and take \( H_1 \) to be a subgroup of finite index in \( \mathcal{G} \).
Now, by Proposition 2.6, $G$ has the congruence subgroup property (for every $N \le_f G$ there exists $n$ such that $\text{St}(n) \le G$) and by Proposition 2.1 for every $n$ we have $\text{St}(n) \le G$. A corresponding statement is also true for $G \times G$: if $N \le G \times G$, then $N_1 := N \cap (G \times 1) \le G \times 1$ and $N_2 := N \cap (1 \times G) \le 1 \times G$. So there exist $n_1, n_2$ such that $\text{St}(n_1) \times 1 \le N_1$ and $1 \times \text{St}(n_2) \le N_2$; therefore $\text{St}(m) \times \text{St}(m) \le N$ where $m = \max\{n_1, n_2\}$. Thus, since $H_2$ is commensurable with $G$, it contains a subgroup $M_2$ of finite index such that $M_2 \cong M_2 \le G$ and $M_2 \le G^{(\times n)}$ for some $n$. So, by a similar procedure as with $H_1$, we may take $M_2$ instead of $H_2$ and obtain

$$H \le_h H_1 \times G_1 \times \cdots \times G_r \times H_3$$

where $H_1 \le G$, $G_i \cong G$ and $H_3$ is finite. Also, take $r$ to be as small as possible.

Denote by $A$ the whole direct product $A := H_1 \times G_1 \times \cdots \times G_r \times H_3$ and for each $i$ write $G_i := (1, \ldots, G_i, 1) \times H_3$. Denote by $B$ the subgroup $B := G_1 \cdots G_r H_3$ of $A$. Now, $H$ projects onto every $G_i$ and $H_3$ but the map

$$H \to H_1 \times G_1 \times \cdots \times G_{i-1} \times G_{i+1} \times \cdots \times G_r \times H_3$$

is not injective so it has a non-trivial kernel $K_i \le H$ which is also normal in $G_i$. Then, since $G$ is just infinite, every $K_i$ has finite index in $G_i$. Take $K_{r+1}$ to be the kernel of the projection map from $H$ to the direct product of all factors of $A$ except $H_3$, so $K_{r+1}$ has finite index in $H_3$. Thus the product of all these kernels, $C := K_1 \cdots K_{r+1}$, is a normal subgroup of $B$ with finite index. Clearly, $A = BH$ and $C \le H$.

Let us examine the cosets of $H$ in $A$. For every $a \in A$, we have $a = bh$ for some $b \in B, h \in H$. So $(aH) \cap B = (bhH) \cap B = b(H \cap B)$. But since $C \le H \cap B$ and $C \le B$, we must have $H \cap B \le_f B$. So since every coset of $H$ in $A$ corresponds to a coset of $H \cap B$ in $B$ (by the calculation above), we must have finitely many of them i.e., $H \le f A$. Therefore $H$ is isomorphic to a subgroup of finite index in a direct product of finitely many copies of $G$ and a finite group, and so it is commensurable with $G \times H_3$ by Proposition 2.3. Hence $H$ is commensurable with $G$.

We will use this same procedure for the case where all $H_i$ are commensurable with $G$.

By the same arguments as for $H_1$ above, we may assume that $H_1$ and $H_2$ are subgroups of finite index in $G$ and by the same argument as for $H_2$ above, we may take $H_3$ to be a subdirect product of finitely many copies of $G$. Thus we have

$$H \le_h H_1 \times H_2 \times G_1 \times \cdots \times G_r$$

where $G_i \cong G$, $H_i \le G$ and $r$ is as small as possible. By defining $A := H_1 \times H_2 \times G_1 \times \cdots \times G_r$, $B := G_1 \cdots G_r$ and $C := K_1 \cdots K_r$ (where $K_i$ is the kernel of the projection map from $H$ to the direct product of all factors of $A$ except $G_i$), we may argue as above to obtain that $H$ has finite index in $A$. Hence $H$ is isomorphic to a subgroup of finite index in a product of finitely many copies of $G$, making it commensurable with $G$ (by Proposition 2.3).
Theorem 2. The Gupta–Sidki 3-group \( G \) is subgroup separable and so the generalized word problem for \( G \) is solvable.

To prove this theorem it suffices to show that the conditions of Theorem 6 hold for the class \( S \) of finitely generated subgroups of \( G \) all of whose subgroups of finite index are closed with respect to the profinite topology on \( G \).

Clearly, \( S \) satisfies (I). To see that it also satisfies (II) let \( H \leq L \) for some \( H \) in \( S \); thus \( L \) is finitely generated. For any \( K \leq H \), we have \( K \cap H \leq H \) so \( K \cap H \) is closed in \( G \) by assumption. But \( K \cap H \) also has finite index in \( K \), hence each of its finitely many cosets is also closed in \( G \) and therefore so is \( K \), their union.

Before we can show that \( S \) also satisfies (III) we need the following lemma.

Lemma 5.2. (i) Suppose that \( H_0 \) is a group all of whose quotient groups are residually finite and that each of the groups \( G_1, \ldots, G_n \) either is finite or is residually finite, just infinite and not virtually abelian. Let \( H \leq_n H_0 \times G_1 \times \cdots \times G_n \). Then every quotient of \( H \) is residually finite.

(ii) If \( H \) is abstractly commensurable with \( G \) or \( G \times G \) then every quotient of \( H \) is residually finite.

Proof. This is essentially Lemma 12 in [8]. The first part is identical and the proof of the second only requires small modifications. Suppose that \( K \leq H \); we want to show that \( K \) is an intersection of subgroups of finite index in \( H \). Since \( H \) is commensurable with \( G \) or \( G \times G \) and they both have the congruence subgroup property, there is some normal subgroup \( N \) of finite index in \( H \) which is a subdirect product of finitely many copies of \( G \). By the same argument as in the proof of (II) it suffices to show that \( K \cap N \) is closed in \( N \) with respect to the profinite topology on \( N \). This follows from the first part of the lemma, as it implies that \( N/(K \cap N) \) is residually finite. 

We may now proceed to show that \( S \) also satisfies (III).

Lemma 5.3. Let \( H \) be a finitely generated subgroup of \( \text{St}(1) \) such that its first level vertex sections \( H_0, H_1, H_2 \) are in \( S \). Then \( H \) is in \( S \).

Proof. Let \( K \leq H \) and \( \bar{K} := \psi(K) \). We will show that \( \bar{K} \) is closed in \( G \times G \times G \), so that \( K \) is closed in \( \text{St}(1) \leq G \) and hence in \( G \). By assumption, \( H_i \) is closed in \( G \) for \( i = 0, 1, 2 \), so that their direct product is closed in \( G \times G \times G \). Notice that since each first level vertex section \( K_i \) of \( K \) is of finite index in the corresponding \( H_i \), it suffices to show that \( \bar{K} \) is closed in \( G \times G \times G \).

Let \( L_0 \times 1 \times 1 := (H_0 \times 1 \times 1) \cap \bar{H} \). Then \( L_0 \) is normal in \( H_0 \), which is commensurable with \( G \) or \( G \times G \) by Theorem 1. Thus, by Lemma 5.2, \( H_0/L_0 \) is residually finite and we may find a collection \( (N_j)_{j \in J} \) of subgroups of finite index in \( H_0 \) whose intersection is \( L_0 \). It is not hard to show that \( \bar{H} \) is the intersection of the subgroups \( (N_j \times 1 \times 1)\bar{H} \) and so it will be enough to prove that each of them is closed in \( G \times G \times G \). Since \( K_j \) is of finite index in \( H_0 \), we have that \( (K_j \times 1 \times 1)\bar{H} \) is of finite index in \( H_0 \times H_1 \times H_2 \). Further, there exist subgroups \( M_0, M_1, M_2 \) of finite index in \( H_0, H_1, H_2 \) such that their
direct product is contained in \((K_j \times 1 \times 1)\bar{H}\). Thus \(M_0 \times M_1 \times M_2\) is closed in \(G \times G \times G\) and hence so is \((K_j \times 1 \times 1)\bar{H}\), as required.

Acknowledgements

I would like to thank my supervisor, John S. Wilson, for suggesting the problem and for his continued guidance and support.

This work was carried out as part of my DPhil studies, financially supported by Fundación La Caixa (Spain).

References

[1] L. Bartholdi and R. I. Grigorchuk. On parabolic subgroups and Hecke algebras of some fractal groups. Serdica Math. J., 28(1):47–90, 2002.

[2] L. Bartholdi, R. I. Grigorchuk, and Z. Šunič. Branch groups. In Handbook of algebra, Vol. 3, pages 989–1112. North-Holland, Amsterdam, 2003.

[3] L. Bartholdi, V. A. Kaimanovich, and V. V. Nekrashevych. On amenability of automata groups. Duke Math. J., 154(3):575–598, 2010.

[4] G. A. Fernández-Alcober and A. Zugadi-Reizabal. GGS-groups: order of congruence quotients and Hausdorff dimension. Trans. Amer. Math. Soc., 366(4):1993–2017, 2014.

[5] A. Garrido and J. S. Wilson. On subgroups of finite index in branch groups. Journal of Algebra, 397(0):32 – 38, 2014.

[6] R. I. Grigorchuk. On Burnside’s problem on periodic groups. Funktsional. Anal. i Prilozhen., 14(1):53–54, 1980.

[7] R. I. Grigorchuk. Degrees of growth of finitely generated groups, and the theory of invariant means. Mathematics of the USSR-Izvestiya, 25(2):259, 1985.

[8] R. I. Grigorchuk and J. S. Wilson. A structural property concerning abstract commensurability of subgroups. J. London Math. Soc. (2), 68(3):671–682, 2003.

[9] R. I. Grigorchuk and J. S. Wilson. A minimality property of certain branch groups. In Groups: topological, combinatorial and arithmetic aspects, volume 311 of London Math. Soc. Lecture Note Ser., pages 297–305. Cambridge Univ. Press, Cambridge, 2004.

[10] N. Gupta and S. Sidki. On the Burnside problem for periodic groups. Math. Z., 182(3):385–388, 1983.

[11] N. Gupta and S. Sidki. Some infinite \(p\)-groups. Algebra i Logika, 22(5):584–589, 1983.
[12] M. Hall, Jr. Coset representations in free groups. *Trans. Amer. Math. Soc.*, 67:421–432, 1949.

[13] D. D. Long and A. W. Reid. *Surface Subgroups and Subgroup Separability in 3-manifold Topology*. IMPA, 2005.

[14] A. I. Mal’tsev. On homomorphisms onto finite groups. Transl., Ser. 2, Am. Math. Soc. 119, 67-79 (1983); translation from Uch. Zap. Ivanov. Gos. Pedagog Inst. 18, 49-60 (1958), 1958.

[15] B. H. Neumann. Some remarks on infinite groups. *Journal of the London Mathematical Society*, s1-12(2):120-127, 1937.

[16] E. L. Pervova. Everywhere dense subgroups of a group of tree automorphisms. *Tr. Mat. Inst. Steklova*, 231(Din. Sist., Avtom. i Beskon. Gruppy):356–367, 2000.

[17] E. L. Pervova. Maximal subgroups of some non locally finite \( p \)-groups. *Internat. J. Algebra Comput.*, 15(5-6):1129–1150, 2005.

[18] P. Scott. Subgroups of surface groups are almost geometric. *J. London Math. Soc. (2)*, 17(3):555–565, 1978.

[19] S. Sidki. On a 2-generated infinite 3-group: subgroups and automorphisms. *J. Algebra*, 110(1):24–55, 1987.