Aspects of branch groups

Alejandra Garrido
New College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Hilary Term 2015

Para Aita y Ama.

Acknowledgements

I am indebted to my supervisor, John S. Wilson, for his dedication, guidance and encouragement, and for teaching me so many things. I am grateful to Peter M. Neumann, for being a wonderful teacher and always being willing to help. I am also grateful to Benjamin Klopsch, for being a conscientious examiner and providing good feedback.

I thank my mathematical friends, for keeping me from doing maths by talking about other maths. I thank my non-mathematical friends, for keeping me from doing maths by talking about things other than maths. I am especially thankful for the Catweazle Club, a hidden gem in East Oxford.

New College and the Mathematical Institute have provided wonderful environments in which to live and work. I especially wish to thank the night porters at New College, headed by Howard Weller, without whose help my music would have suffered greatly.

I gratefully acknowledge the generous financial support provided by *Fundación "la Caixa"* and the Mathematical Institute.

Finally, I thank my parents, for everything.

Abstract

This thesis is a study of the subgroup structure of some remarkable groups of automorphisms of rooted trees. It is divided into two parts. The main result of the first part is seemingly of an algorithmic nature, establishing that the Gupta–Sidki 3-group G has solvable membership problem. This follows the approach of Grigorchuk and Wilson who showed the same result for the Grigorchuk group. The proof, however, is not algorithmic, and it moreover shows a striking subgroup property of G: that all its infinite finitely generated subgroups are abstractly commensurable with either G or $G \times G$. This is then used to show that G is subgroup separable which, together with some nice presentability properties of G, implies that the membership problem is solvable. The proof of the main theorem is also used to show that G satisfies a "strong fractal" property, in that every infinite finitely generated subgroup acts like G on some rooted subtree.

The second part concerns the subgroup structure of branch and weakly branch groups in general. Motivated by a natural question raised in the first part, a necessary condition for direct products of branch groups to be abstractly commensurable is obtained. From this condition it follows that the Gupta–Sidki 3-group is not abstractly commensurable with its direct square. The first main result in the second part states that any (weakly) branch action of a group on a rooted tree is determined by the subgroup structure of the group. This is then applied to answer a question of Bartholdi, Siegenthaler and Zalesskii, showing that the congruence subgroup property for branch and weakly branch groups is independent of the actions on a tree. Finally, the information obtained on subgroups of branch groups is used to examine which groups have an essentially unique branch action and why this holds.

Contents

Notation

$h^g = g^{-1}hg$	
$[g,h] = g^{-1}h^{-1}gh$	
$\mathrm{N_G}(H)$	the normalizer of H in G
$\mathrm{C_G}(H)$	the centralizer of H in G
$H^G = \langle H \rangle^G$	the subgroup generated by all conjugates of H by elements of G, the normal closure of H in G
$G', G'' = G^{(2)}, \dots, G^{(n)}$	first, second, nth derived subgroups of G
$\gamma_n(G)$	the nth term in the lower central series of G
$\lvert G \colon H\rvert$	the index of H in G
$H \le_{\mathrm{f}} G$, $H \trianglelefteq_{\mathrm{f}} G$	H is a subgroup (resp. normal subgroup) of finite index in G
$G \times \overset{n}{\cdots} \times G = G^{\times n}$	the direct product of n copies of G
$H \le_{\mathrm{s}} G_1 \times \cdots \times G_n$	H is a subdirect product of $G_1, \dots, G_n$
$\mathbb{F}_p$	the field of p elements
$\mathrm{Sym}(n)$	the symmetric group on n letters
$(1\ 2\ \dots\ d)$	the cyclic permutation mapping 1 to 2, 2 to 3, ..., d to 1
C_p	the cyclic group of order p
$H \wr G$	the permutational wreath product of H by G
T	a rooted tree (usually level-homogeneous, or regular)
T_v	the subtree rooted at vertex v
V_n	the vertices of T at distance n from the root
$\mathrm{Aut}\,T$	the automorphism group of the tree T
$\mathrm{St}(v)$, $\mathrm{St}(n)$	the stabilizer in $\mathrm{Aut}\,T$ of a vertex v, of the set V_n
$\mathrm{St}_G(v)$, $\mathrm{St}_G(n)$	the stabilizer in $G \le \mathrm{Aut}\,T$ of a vertex v, of the set V_n
$\mathrm{St}_\rho(v)$, $\mathrm{St}_\rho(n)$	the stabilizer in $\rho(G) \le \mathrm{Aut}\,T$ where $\rho : G \hookrightarrow \mathrm{Aut}\,T$ is a (weakly) branch action
$\mathrm{rist}_G(v)$ or $\mathrm{rist}(v)$	the rigid stabilizer (usually in $G \le \mathrm{Aut}\,T$) of a vertex v in T
$\mathrm{rist}_G(n) = \prod_{v \in V_n} \mathrm{rist}_G(v)$	the nth rigid level stabilizer
$\mathrm{rist}_\rho(v)$, $\mathrm{rist}_\rho(n)$	same as above, where $\rho : G \hookrightarrow \mathrm{Aut}\,T$ a (weakly) branch action
g_v	the restriction of g to T_v
φ_v	the homomorphism $\varphi_v : \mathrm{St}_G(v) \to \mathrm{Aut}\,T$, defined by $g \mapsto g_v$
G_v	the vertex section of G at v, that is, the image of φ_v
ψ_n	$\psi_n : \mathrm{St}_G(n) \to \prod_{v \in V_n} \varphi_v(\mathrm{St}_G(n)) \le \prod_{v \in V_n} \mathrm{Aut}\,T$
$\mathcal{L}$	the structure lattice of a branch group (see Definition 5.1.6)
$\mathcal{B}$	the structure graph of a branch group (see Definition 5.2.2)

Chapter 1

Introduction

In the early 1980s Aleshin [2], Grigorchuk [26], and Gupta and Sidki [37], discovered groups of automorphisms of rooted trees with extraordinary properties. They are groups which can be easily seen to be finitely generated, infinite and periodic, answering the General Burnside Problem (although they are not the first answer to this problem, this being Golod–Shafarevich groups). The group proposed by Grigorchuk was the first to be proved to have intermediate word growth [27]. Indeed, to this day, all known groups of intermediate growth are based on this type of construction. The Grigorchuk group is also the first example of an amenable but not elementary amenable group [27].

Soon a host of similar examples appeared, abstracting some of the defining characteristics of the early ones. Attempts to round up these examples into a unified class, allowing a more systematic approach to their study, led to the definition of two types of groups: On one hand, we have self-similar groups (also known as groups generated by automata), reflecting the "fractal" nature of these examples. These will be defined in the next chapter (Definition 2.1.2). On the other hand, we have branch groups, which have a subgroup structure similar to that of the full automorphism group of a rooted tree.

Definition 1.0.1. Let T be a *level-homogeneous* tree; that is, an infinite, rooted, locally finite tree such that all vertices at the same distance from the root have the same valency. A special case of this is the d-regular rooted tree for some $d \geq 2$, where the valency of the root is d and that of all other vertices is $d+1$. The *level* of a vertex is its distance from the root and we denote by V_n the set of vertices of level n. We picture T with the root at the top so that T_v, the *subtree rooted at* $v \in T$, consists of all vertices below v.

Let a group G act faithfully on T and v be a vertex of T. The *rigid stabilizer* $\mathrm{rist}_G(v)$ *of* v is the subgroup of elements of G which fix every vertex outside the subtree rooted at v, and $\mathrm{rist}_G(n) := \langle \mathrm{rist}_G(v) \mid v \in V_n \rangle$ is the nth *rigid level stabilizer.*

We call the faithful action of G on T a *branch action* if the following holds for all $n \geq 0$:

(i) G acts transitively on V_n;

(ii) $\mathrm{rist}_G(n)$ has finite index in G.

We say that G is a *branch group* if there exists a branch action of G on some tree T.

A *weakly branch action* is one where (ii) above is weakened to '$\mathrm{rist}_G(n)$ is infinite'. Correspondingly, a *weakly branch group* is one with a weakly branch action on some T.

The initial examples mentioned at the beginning lie in the intersection of self-similar and branch groups and tools from both areas have enabled many more interesting results to be proved, as we shall see in Chapter 2. One of the recurring themes in these proofs is the use of length reduction arguments. These will be explained in more detail in the next chapter but the rough idea is that one writes an element of a level stabilizer as a word in the generators, and then, by the self-similarity of the group, the restrictions of this element to the subtrees rooted at that level are again elements of the group (the vertex sections), usually of shorter length. Therefore one can proceed by induction.

This simple idea appeared right from the start, and is used to show that the main examples are torsion groups. It is also the reason why branch self-similar groups have good algorithmic properties. For instance, the word problem is solvable by a fast algorithm which exploits precisely this reduction in length of a word representing the identity. More sophisticated versions of this argument were needed to show that many branch groups, including the key examples, have solvable conjugacy problem. The membership problem (determine, given finitely many elements $x, g_1, \ldots, g_n$, whether or not x is in the subgroup generated by $\{g_1, \ldots, g_n\}$) is harder and was only known (until now) to be solvable for the Grigorchuk group, by work of Grigorchuk and Wilson [33]. In Chapter 3, which is based on [23], we will see that the membership problem is also solvable for the Gupta–Sidki 3-group. The difficulty in the case of the membership problem, and the reason why more results about it have not been shown, is that the length reduction has to be carried out for several elements simultaneously (the finitely many generators of a subgroup). This is hard even for the Grigorchuk group where the fact that it is a 2-group is helpful. Chapter 3 shows that, although the difficulties are still there, this can be done for the 3-group example of Gupta and Sidki.

Recall that two groups are *(abstractly) commensurable* if they have isomorphic finite index subgroups. What is really shown in [33] and Chapter 3 is that the Grigorchuk and Gupta–Sidki groups have a striking commensurability property of finitely generated groups. Namely,

Theorem 3.1. *Every infinite finitely generated subgroup of the Gupta–Sidki* 3*-group* G *is abstractly commensurable with* G *or* $G \times G$*.*

Using this, one can then prove the other main theorem of Chapter 3.

Theorem 3.2. *The Gupta–Sidki* 3*-group is subgroup separable.*

A group is *subgroup separable* or *LERF* if all its finitely generated subgroups are intersections of finite index subgroups, i.e., closed in the profinite topology (more on separability properties will be said in Section 2.3). Since the Gupta–Sidki 3-group has a finite endomorphic presentation (see Definition 2.3.1), this last theorem together with Theorem 2.14 implies that the membership problem is also solvable (see Subsection 2.3.1 for more details).

The proof of Theorem 3.1 also allows for the following characterization of finite subgroups of the Gupta–Sidki 3-group. The vertex sections of a subgroup of a self-similar group are defined in the next section.

Theorem 3.5. *Let H be a finitely generated subgroup of the Gupta–Sidki 3-group G. Then either H is finite or one of its vertex sections is equal to G.*

Before the proofs of these theorems, there is in Chapter 2 a short survey of the study of a generalized version of Gupta–Sidki groups: GGS groups (this stands for Grigorchuk–Gupta–Sidki, a term coined by Baumslag in [13]). This will provide some background, some necessary preliminary results for Chapter 3 and will illustrate the key ideas of length reduction used. It is also sprinkled with the occasional question and possibility for future research.

The first theorem above raises a very natural question: are the commensurability classes of G and $G \times G$ distinct? This leads to the second part of the thesis which is devoted to the subgroup structure of branch and weakly branch groups, showing how it is inextricably linked to all possible such actions on rooted trees. This link is obtained via a generalization of the following theorem by Grigorchuk (Theorem 4.1) on the normal subgroups of (weakly) branch groups:

Theorem 4.1 ([28, Theorem 4])**.** *Suppose that G is a weakly branch group acting on a tree T and let $K \lhd G$ with $K \neq 1$. Then K contains the derived subgroup $\mathrm{rist}_G(n)'$ of $\mathrm{rist}_G(n)$ for some integer n.*

The main result of Chapter 4 is

Theorem 4.2. *Suppose that G is a weakly branch group acting on a tree T. Let H be a subgroup of finite index and let $K \lhd H$. Then for all sufficiently large integers n there is a union X of H-orbits in V_n such that $K \cap \mathrm{rist}_G(n)' = \mathrm{rist}_G(X)'$. More precisely,*

$$\mathrm{rist}_G(X)' \leq K \quad \textit{and} \quad K \cap \mathrm{rist}_G(V_n \setminus X) = [K, \mathrm{rist}_G(V_n \setminus X)] = 1.$$

Intuitively, this shows that all subgroups with finitely many conjugates in a weakly branch group correspond to orbits of their normalizers on some layer of the tree.

This has two applications. The first one, and the initial motivation for the study of such subgroups, is to distinguish abstract commensurability classes of weakly branch groups. As a consequence of Theorem 4.2, an invariant of finite index subgroups of weakly branch groups is obtained, in terms of their maximum number of orbits on layers of any tree. This yields a clean solution to the unresolved commensurability problem from the first part, namely, it yields that the Gupta–Sidki 3-group is not commensurable with its direct square. More generally, the following corollary is obtained:

Corollary 4.2.5. *Let p be a prime and T denote the p-regular rooted tree. Let $G_1, \ldots, G_n$ and $H_1, \ldots, H_m$ be weakly branch groups which act as p-groups at each level of T. If $\prod_{i=1}^n G_i$ and $\prod_{j=1}^m H_j$ are abstractly commensurable then $n \equiv m \mod p-1$.*

The material in Chapter 4 is based on a joint paper with J. S. Wilson ([25]), but some arguments have been simplified and generalized to the weakly branch setting in Section 4.1.

The subgroup structure of branch groups has been studied by several authors. In fact, in some cases this was done before the definition of branch groups was even coined!

This is because they appear naturally in the theory of just infinite groups (infinite groups all of whose proper quotients are finite). Every finitely generated group admits an epimorphism to a just infinite group, so that in studying properties of groups that are preserved by passing to quotients, it is natural to study the just infinite case first.

In [65, 66] Wilson introduced the notion of a *structure lattice* of a (not virtually abelian) just infinite group. This is the quotient of the lattice of subnormal subgroups with finitely many conjugates in the group by the relation of having finite index intersection. The structure lattice can be finite or infinite and the latter case yields precisely the just infinite branch groups. A modified version of the structure lattice (where the equivalence relation is a virtually abelian version of having finite index intersection) was used by P. Hardy in [41] to characterize branch groups in purely group-theoretic terms. This is the approach that will be followed in Chapter 5, where a more direct description of the structure lattice is given, obtained from the lattice of subgroups with finitely many conjugates (without the subnormality assumption). This also allows for a simplification of the arguments used to show that the structure lattice is indeed a lattice. Furthermore, Theorem 4.2 provides a characterization of the structure lattice (Proposition 5.1.7): its elements correspond exactly to (unions of) orbits on layers, not necessarily of the branch group, but of its finite index subgroups.

The *structure graph* of a branch group is a subset of the structure lattice which only takes into account subgroups that behave like the rigid stabilizers (*basal subgroups*). It is defined at the end of Chapter 5 and, as we shall see in Chapters 6 and 7, it is a very useful object in our context. Theorem 4.2 also provides a characterization of this graph in terms of the action of the branch group (Proposition 5.2.3): its vertices are exactly blocks of imprimitivity in the action of the branch group on each layer.

The second application of Theorem 4.2, explained in Chapter 6, is to the congruence subgroup property for branch and weakly branch groups. A group acting faithfully on a rooted tree satisfies the *congruence subgroup property* if every finite index subgroup contains some level stabilizer. The *congruence subgroup problem* consists of determining whether or not a given group satisfies the congruence subgroup property and, if it does not, calculating the *congruence kernel*, which encodes by how much the property is not satisfied (this will be explained in more detail in Chapter 6). Theorem 6.1 provides an answer to a question of Bartholdi, Siegenthaler and Zalesskii ([10]), showing that whether or not a branch group has the congruence subgroup property is independent of the branch action on a tree and indeed that the congruence kernel is determined by the group and not the branch action:

Theorem 6.1. *Let G have two branch actions on trees. Then the congruence kernels with respect to these actions coincide.*

In fact, this follows from the more general Theorem 6.3 which states that the congruence topology of a branch group (which will be defined in Chapter 6) is determined by the group; specifically, by its structure graph. There are also analogous results (Theorem 6.2 and Theorem 6.4) for the property that every finite index subgroup of a branch group contains some rigid level stabilizer.

The material in Chapter 5 and the first part of Chapter 6 is contained in my paper [24].

Although it is not clear how one should define the analogue of the structure lattice in the case of weakly branch groups, similar ideas can be applied to show that two weakly branch actions of the same group yield the same congruence topology:

Theorem 6.5. *Let G be a weakly branch action. The congruence kernels of any two weakly branch actions of G coincide.*

The fact that the congruence subgroup property is independent of the branch action of a group makes one wonder whether every branch group has a unique such action. This has a naive negative answer, as any group with a branch action on, say, the binary tree, also has a branch action on the 4-regular tree, obtained by simply ignoring (or *deleting*) every other layer of the binary tree. We might nevertheless consider these actions to be equivalent and ask when every branch action of a group is obtained in such a way. This was studied in [34] where the authors showed that all branch actions are "equivalent" when the structure graph of a branch group is isomorphic to a rooted tree. Chapter 7 continues where [34] left off, providing a proof (Proposition 7.1.1) that there is an essentially unique branch action of a group (in the sense explained above) if and only if its structure graph is a tree, and giving a necessary condition for this to be the case (Proposition 7.1.4). Examples of groups which satisfy these conditions, and groups that do not, are also given.

Part I

GGS groups

Chapter 2

A short survey on GGS groups

The purpose of this chapter is to define a natural generalization of Gupta–Sidki groups, GGS groups, and to use a (necessarily somewhat biased) account of the state of the art in their study as an excuse to put in context the results of the next chapter and illustrate some of the key ideas that will be used there. There is also a new result on the congruence subgroup property for GGS groups (Theorem 2.9).

2.1 Definitions and preliminaries

Let T be an infinite, locally finite, rooted tree. The *level* of a vertex v is the number of edges in the unique path from v to the root and the *nth layer* V_n is the set of all vertices of level n. We say that T is *level-homogeneous* if there exists a sequence $(m_n)_n$ of integers $m_n \geq 2$ such that every $v \in V_n$ has valency $m_n + 1$ (one edge connecting v to its *parent* in V_{n-1} and m_n edges connecting v to its *children* in V_{n+1}). The special case $m_n = d$ for all $n \in \mathbb{N}$ yields the *d-regular* rooted tree. Only level-homogeneous trees will be considered hereafter. For a vertex $v \in V_n$, the subtree consisting of vertices of level $m \geq n$ separated from the root by v is the *subtree rooted at* v, denoted by T_v.

An *automorphism* of a rooted tree T is a permutation of the vertices that preserves the adjacency relation and the root. The group of all automorphisms of T is denoted by $\operatorname{Aut} T$. Denoting by v^g the image of a vertex $v \in T$ under $g \in \operatorname{Aut} T$, the *stabilizer* of v is

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

and the nth *level stabilizer* is

$$\mathrm{St}(n) := \bigcap_{v \in V_n} \mathrm{St}(v).$$

Since T is locally finite, the level stabilizers have finite index in $\operatorname{Aut} T$, and the intersection of all of them is trivial. This means that we can take the level stabilizers as a basis of neighbourhoods of the identity to obtain a profinite topology with respect to which $\operatorname{Aut} T$ is complete. Thus $\operatorname{Aut} T$ is a profinite group: the inverse limit of the quotients by level stabilizers.

For any group G acting on T, denote by $\mathrm{St}_G(v)$ and $\mathrm{St}_G(n)$ the intersection of the image of G in $\operatorname{Aut} T$ with the above subgroups. The level stabilizers $\mathrm{St}_G(n)$ have finite index in the image of G and $\bigcap_{n \in \mathbb{N}} \mathrm{St}_G(n) = 1$ so that, if the action on T is faithful, then G is residually finite.

GGS groups are groups of automorphisms of a regular rooted tree. The vertices of the *d-regular rooted tree* T are labelled by finite words over the alphabet $\{0, \ldots, d-1\}$. The root is labelled by the empty word and a pair of vertices labelled by u and v are joined by an edge if $v = ux$ (or $u = vx$) for some x in the alphabet. Vertices will be identified with their labels from now on. Since T is regular, for every vertex v there is an obvious isomorphism from T to T_v taking u to vu, allowing for the identification of all T_v with T henceforth.

Due to the self-similarity of the regular tree T, its automorphism group $\operatorname{Aut} T$ decomposes as

$$\operatorname{Aut} T \cong (\operatorname{Aut} T \times \cdots \times \operatorname{Aut} T) \rtimes \mathrm{Sym}(d)$$

where $\mathrm{Sym}(d)$ is the symmetric group on the d letters of the alphabet used to define T (thought of as the vertices in V_1), which acts on $\operatorname{Aut} T \times \cdots \times \operatorname{Aut} T$ by permuting the factors. In other words, $\operatorname{Aut} T$ is the permutational wreath product $\operatorname{Aut} T \wr \mathrm{Sym}(d)$. This means that every automorphism g of T can be uniquely decomposed as

$$g = (g_0, \ldots, g_{d-1})\sigma$$

where $g_0, \ldots, g_{d-1} \in \operatorname{Aut} T$ act on the subtrees $T_0, \ldots, T_{d-1}$ and $\sigma \in \mathrm{Sym}(d)$ is a permutation of V_1. We say that the automorphism g is *rooted* if $g = (1, \ldots, 1)\sigma$; that is, g only permutes (rigidly) the subtrees rooted at level 1. Each g_i is the *vertex section* of g at vertex i. This can be extended to all vertices of T: for a given vertex v, the *vertex section* of g at v is the restriction g_v of g to the subtree T_v; that is, the unique automorphism such that $(vw)^g = (v^g)(w^{g_v})$ for all finite words w over the alphabet.

Notice that for $g \in G \leq \operatorname{Aut} T$ the vertex sections of g may not be elements of G.

Definition 2.1.1. A subgroup $G \leq \operatorname{Aut} T$ is *self-similar* if for every $g \in G$ and every vertex v of T, the vertex section g_v is in (the image of) G. More generally, a faithful action of a group G on T is *self-similar* if the image of G in $\operatorname{Aut} T$ is self-similar.

Remark 1. The above explains the definition of the rigid level stabilizers $\operatorname{rist}_G(n)$ and of a branch group (Definition 1.0.1). Consider for the moment a more general level-homogeneous rooted tree T. Each level stabilizer $\operatorname{St}(n)$ in $\operatorname{Aut}(T)$ decomposes as the direct product $\prod_{v \in V_n} \operatorname{Aut}(T_v)$. The elements of $\operatorname{rist}_G(n)$ are precisely those that exhibit this kind of behaviour, in that they are those elements which can be written as a product of automorphisms of G, where each factor acts only on the subtree rooted at a vertex in V_n. The requirement that each $\operatorname{rist}_G(n)$ have finite index in G in the definition of a branch group ensures that level stabilizers in a branch group behave very similarly to those of the full automorphism group $\operatorname{Aut}(T)$.

Back to self-similar groups and regular trees. For any group G acting on T and a vertex v of T the map $\varphi_v : \operatorname{St}_G(v) \to \operatorname{Aut} T$ which takes $g \in \operatorname{St}_G(v)$ to its vertex section g_v is actually a homomorphism whose image G_v is called the *vertex section of G at v* (some authors call it an *upper companion group* or *projection* at a vertex). Observe that if $v = uw$ then $\varphi_v = \varphi_w \circ \varphi_u$. For each n, the map

$$\psi_n : \operatorname{St}_G(n) \to \prod_{v \in V_n} \varphi_v(\operatorname{St}_G(n)) \leq \prod_{v \in V_n} \operatorname{Aut} T$$

is also a homomorphism; and an injective one at that. We may therefore identify elements of $\operatorname{St}_G(n)$ with their images under ψ_n, which we have already done above for $n = 1$. Following the custom in the literature, we will continue to omit the ψ for elements of the first level stabilizer.

Definition 2.1.2. A subgroup G of $\operatorname{Aut} T$ is *fractal* if $\varphi_v(\operatorname{St}_G(v)) = G$ for every vertex v of T. More generally, a faithful action of G on T is *fractal* if the image of G in $\operatorname{Aut} T$ is fractal.

The theory of self-similar actions/groups is a rich one, with many connections with dynamical systems and fractal geometry. The book [47] is an excellent source for interested readers.

In a similar vein, we can define (weakly) regular branch groups.

Definition 2.1.3. Given a fractal subgroup G of $\operatorname{Aut} T$, if there exists a finite index subgroup $K \leq G$ such that

$$\psi(K) \geq K \times \cdots \times K$$

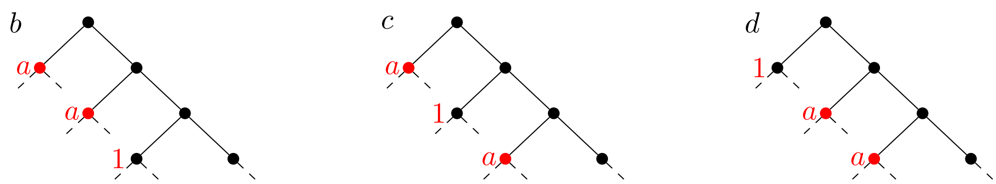
and the above containment is of finite index, then G is a *regular branch group over its subgroup* K.

This in particular implies that $K \leq \operatorname{rist}_G(v)$ for each $v \in V_1$ so that $\operatorname{rist}_G(1)$ has finite index in G. The argument can then be repeated to obtain that each $\operatorname{rist}_G(n)$ has finite index in G. Thus, if G acts transitively on each V_n, it is a branch group.

If there simply exists such a K (which may not have finite index in G), we say that G is *weakly regular branch over its subgroup* K. The same argument as above shows that, if G acts transitively on each V_n, then G is in particular a weakly branch group.

Despite not being a GGS group, the (first) Grigorchuk group[1] will be mentioned often enough in the remainder to justify its definition here.

Definition 2.1.4. The Grigorchuk group Γ is the group of automorphisms of the binary rooted tree generated by automorphisms a, b, c, d, where a is a rooted automorphism that swaps the vertices at the first level and $\psi(b) = (a, c), \psi(c) = (a, d), \psi(d) = (1, b)$. It is fractal and regular branch (see [17, Chapter VIII] for a nice exposition of these and other properties of this group). The figures below show (a part of) the action of the generators b, c, d.



[1] There is also a second Grigorchuk group, which is a GGS group, hence the first G in GGS.

Definition 2.1.5 (GGS groups)**.** Let $d > 2$ be an integer and let T be the d-regular rooted tree. Given a vector $\mathbf{e} = (e_1, \ldots, e_{d-1}) \in (\mathbb{Z}/d\mathbb{Z})^{d-1}$, the *GGS group with defining vector* $\mathbf{e}$ is the subgroup $G = \langle a, b \rangle$ of $\operatorname{Aut} T$ generated by two automorphisms a and b.

The automorphism a is rooted and acts on V_1 as the cyclic permutation $(0\ 1\ \ldots\ d{-}1)$. The element $b = (a^{e_1}, \ldots, a^{e_{d-1}}, b)$ is recursively defined by its action on subtrees rooted at V_1: it acts on T_{i-1} as a^{e_i} acts on T for $i = 1, \ldots, d-1$, and as b on T_{d-1}. An automorphism of this sort, which fixes a ray but acts non-trivially on T_v for infinitely many vertices v at distance 1 from the ray, is called *directed*.

We finish this section with some examples of GGS groups, the Gupta–Sidki p-groups and the Fabrykowski–Gupta group. As will be shown in the next section, they are fractal and indeed regular branch (Proposition 2.2.5).

Example 2.1.6 (Gupta–Sidki p-groups [37])**.** These are the motivating examples for the broader definition of GGS group. In this case, $d = p$ an odd prime and $\mathbf{e} = (1, -1, 0, \ldots, 0) \in \mathbb{F}_p^{p-1}$ so that $b = (a, a^{-1}, 1, \ldots, 1, b)$. Figure 2.1 shows the action of b on T for $p = 3$. We will see in the next section (Theorem 2.1) that they are p-groups, for their corresponding p.

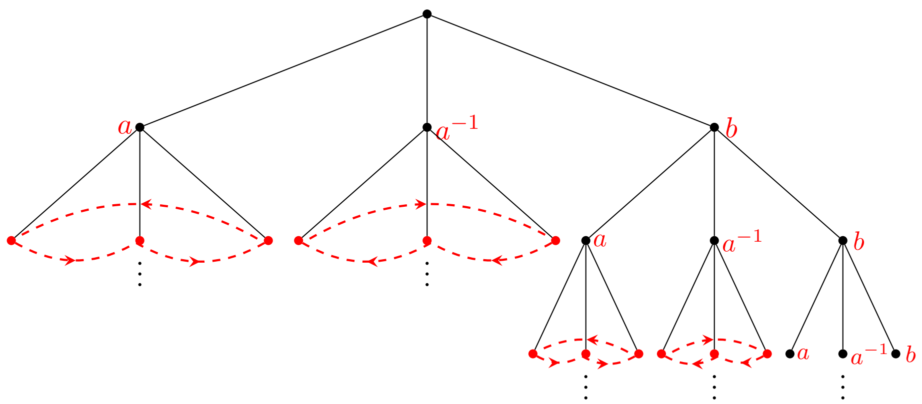


Figure 2.1: The directed generator b of the Gupta–Sidki 3-group.

Example 2.1.7 (Fabrykowski–Gupta group [19])**.** This example, with an even simpler definition than those by Gupta and Sidki, is notable for having intermediate word growth (proved in [19]). Here we have $d = 3$ and $\mathbf{e} = (1, 0)$ so $b = (a, 1, b)$, as shown in Figure 2.2.

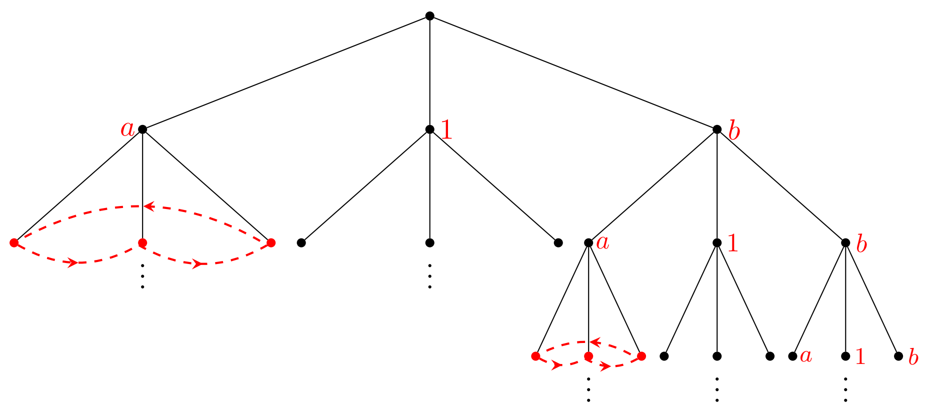


Figure 2.2: The directed generator b of the Fabrykowski–Gupta group.

2.2 Algebraic properties of GGS groups

The definition of a GGS group G ensures that it is self-similar, as every $g \in G$ is a group word in the generators a, b and, by definition of b, the vertex section g_v must be another word in the generators a, b.

Even better, GGS groups are in most cases of interest fractal and therefore infinite. Writing $b_i := b^{a^i}$ for $i = 0, \dots, d-1$, we have

$$\begin{aligned} \psi(b) &= (a^{e_1}, \dots, a^{e_{d-1}}, b), \\ \psi(b_1) &= (b, a^{e_1}, \dots, a^{e_{d-1}}), \\ &\vdots \\ \psi(b_{d-1}) &= (a^{e_2}, \dots, b, a^{e_1}). \end{aligned} \tag{2.1}$$

which makes it plain that $G_i = \varphi_i(\mathrm{St}_G(i)) = G$ for $i = 0, \dots, d-1$, provided that some e_i is a generator of $\mathbb{Z}/d\mathbb{Z}$ (this, in particular, excludes the case $e_1 = \cdots = e_{d-1} = 0$, which yields a finite GGS group). Note that this holds for the examples given above and indeed for the case $d = p$ a prime, which is of most interest here. A criterion for GGS groups acting on the p^n-regular tree to be infinite was found by Vovkivsky (see Theorem 2.2).

The above also yields that for a GGS group G acting on the d-regular tree, $\mathrm{St}_G(1) = B := \langle b \rangle^G = \langle b_0, \dots, b_{d-1} \rangle$ and $G = B \rtimes \langle a \rangle$.

In order to prove some more basic properties about GGS groups, we need to make an important aside on length in GGS groups.

2.2.1 Length in GGS groups

One of the main tools in the study of self-similar groups in general, and GGS groups in particular, is the use of length reduction of words, through the the projection maps shown above.

By way of illustration of this type of argument, which will be key in Chapter 3, we shall prove that the abelianization of a GGS group G acting on the d-regular tree is isomorphic to $\langle a \rangle \times \langle b \rangle \cong \mathbb{Z}/d\mathbb{Z} \times \mathbb{Z}/c\mathbb{Z}$ where c is the order of b in G. The argument presented here is a special case of [11, Section 4.4].

Consider the free product

$$H := \langle a, b \mid a^d = b^c = 1 \rangle \cong \mathbb{Z}/d\mathbb{Z} * \mathbb{Z}/c\mathbb{Z}$$

Clearly, G is a quotient of H. Call the quotient map ρ.

Every word $h \in H$ has a unique normal form:

$$h = a^{s_0} b^{r_1} a^{s_1} \dots b^{r_m} a^{s_m}$$

where $s_i \in \mathbb{Z}/d\mathbb{Z}, r_i \in \mathbb{Z}/c\mathbb{Z}$ and all s_i, r_i except possibly s_0, s_m are non-zero.

Define $|h|$ to be the syllable length of h in the free product, which ranges from $2m-1$ to $2m+1$ depending on whether or not $s_0, s_m = 0$. Each $h \in H$ can be rewritten as

$$(b^{r_1})^{a^{-s_0}} (b^{r_2})^{a^{-s_0-s_1}} \dots (b^{r_m})^{a^{-s_0-s_1-\dots-s_{m-1}}} a^{s_0+s_1+\dots+s_m} \tag{2.2}$$

Let H_1 denote the preimage of $\mathrm{St}_G(1)$ in H under ρ and let $h \in H_1$. Then h can be rewritten in the form (2.2), where $s_0 + s_1 + \dots + s_m \equiv 0 \mod d$, since $h \in H_1$. Now, keeping (2.1) in mind, define $\Psi : H_1 \to H \times \overset{d}{\dots} \times H$ by

$$\Psi(b) = (a^{e_1}, \dots, a^{e_{d-1}}, b),\ \Psi(b^a) = (b, a^{e_1}, \dots, a^{e_{d-1}}),\ \dots,\ \Psi(b^{a^{d-1}}) = (a^{e_2}, \dots, b, a^{e_1}).$$

Put $(h_0, \dots, h_{d-1}) := \Psi(h)$ (note that each h_i may not be a reduced word in H). Then each b^{a^i} in h contributes exactly one appearance of b in h_i and some power of a to the other h_j; thus $|h_i| \le m \le \lceil \frac{|h|}{2} \rceil$ for $i = 0, \dots, d-1$.

For $h \in H$, define h_b to be the word obtained from h by deleting all occurrences of a, thus obtaining a power of b. Define $K_b := \{h \in H \mid h_b = 1\}$, which is the kernel of $H \to \langle b \rangle$.

Lemma 2.2.1 (Cfr. [11, Lemma 4.7])**.** *If $h \in \ker\rho$ then $h \in K_b$. In other words, all relators of G come from K_b.*

Proof. The proof is by induction on $|h|$, the case of the empty word holding trivially. If $|h| = 1$ then h cannot represent the identity in G. Suppose that the result holds for all elements of $\ker\rho$ of syllable length less than n for $n \geq 2$. Let $h \in \ker\rho$ with $|h| = n$. Since h represents the identity in G, it must in particular be mapped to $\mathrm{St}_G(1)$, so we can apply Ψ to it to obtain $h_0, \dots, h_{d-1}$. By the above discussion, $|h_i| \leq \lceil \frac{|h|}{2} \rceil < n$, so $h_i \in K_b$ for $i = 0, \dots, d-1$. Since K_b is closed under concatenation, $h_0 \dots h_{d-1} \in K_b$. Now, each appearance of b in h corresponds to one appearance of b in exactly one h_i. Thus, $h \in K_b$, as required. The claim then follows by induction. □

We may therefore define the map $\pi_b : G \to \langle b \rangle$ by $\pi_b(g) = h_b$ for any $h \in H$ representing g. It is well defined by the above lemma.

Similarly, define h_a to be the word obtained from h by deleting all occurrences of b, and $K_a := \{h \in H \mid h_a = 1\}$, which is the kernel of $H \to \langle a \rangle$. Then $\ker\rho$ is contained in K_a. Thus the map $\pi_a : G \to \langle a \rangle$, taking $g \in G$ to h_a for any $h \in H$ representing g, is well defined.

Hence $\ker\rho \leq K_a \cap K_b$. We are now ready to prove:

Proposition 2.2.2. *Let G be a GGS group acting on the d-regular tree. Then*

$$G/G' \cong \langle a \rangle \times \langle b \rangle \cong \mathbb{Z}/d\mathbb{Z} \times \mathbb{Z}/c\mathbb{Z}$$

where c is the order of b in G. Moreover, $\gamma_i(G)$ has finite index in G for every i.

Proof. In the above discussion, we have just seen that $H/K_a \cong \langle a \rangle, H/K_b \cong \langle b \rangle$ and that $\ker\rho \leq K_a \cap K_b$. The latter implies that if $g \in G$ can be represented by an element of K_a then all other elements of H representing g are also in K_a. We may therefore identify K_a with its image in G under ρ, the kernel of the map $G \to \langle a \rangle$. Similarly, K_b may be identified with $\ker(G \to \langle b \rangle)$. Now, $G/(K_a \cap K_b) \cong \langle a \rangle \times \langle b \rangle$ is abelian, so $K_a \cap K_b \geq G'$. Conversely, any $h \in K_a \cap K_b$ represents the identity in G/G' so $K_a \cap K_b \leq G'$. Hence $G/G' \cong \langle a \rangle \times \langle b \rangle \cong \mathbb{Z}/d\mathbb{Z} \times \mathbb{Z}/c\mathbb{Z}$.

Once we know that the abelianization is finite, it follows from [54, Theorem 5.2.6] that all groups in the lower central series have finite index in G. □

From this we can show the following basic properties[2]:

Proposition 2.2.3. *Let G be a GGS group acting on the d-regular tree and let c be the order of b. Then*

(i) $G/B' \cong (\mathbb{Z}/c\mathbb{Z}) \wr (\mathbb{Z}/d\mathbb{Z})$;

(ii) $\mathrm{St}_G(2) \leq G' \leq B$.

Proof. The first property follows from the fact that B/B' is isomorphic to a direct product of d copies of $\mathbb{Z}/c\mathbb{Z}$ and $\langle a \rangle$ acts on it by cyclically permuting the factors.

For the second property, $G/B \cong \langle a \rangle$ implies that $G' \leq B$. For the other inclusion, first note that $\langle a \rangle \cap \mathrm{St}_G(2) = \langle b \rangle \cap \mathrm{St}_G(2) = 1$. Then, using Proposition 2.2.2, we have

$$cd = |G^{\mathrm{ab}}| \geq \left|\left(\frac{G}{\mathrm{St}_G(2)}\right)^{\mathrm{ab}}\right| = \left|\left(\frac{\langle a \rangle\, \mathrm{St}_G(2)}{\mathrm{St}_G(2)}\right)^{\mathrm{ab}}\right| \cdot \left|\left(\frac{\langle b \rangle\, \mathrm{St}_G(2)}{\mathrm{St}_G(2)}\right)^{\mathrm{ab}}\right| = |\langle a \rangle| \cdot |\langle b \rangle| = cd.$$

□

2.2.2 Torsion

The fractal property shown at the start of the section was first determined by Gupta and Sidki in [37], for the Gupta–Sidki p-groups, where they were also first defined. Gupta and Sidki also showed there that their groups are p-groups, using a similar length reduction argument to that shown in the previous subsection. The proof is reproduced here for two reasons. Firstly, because it is a very elegant, almost "bare-hands" proof of the fact that there exist finitely generated periodic and yet infinite groups, answering in the negative the General Burnside Problem (although it had been answered previously by other mathematicians, most notably, Golod and Shafarevich). Secondly, because it is another application of a length reduction argument, this time with a different length function, which shows how these arguments can be varied.

Theorem 2.1 ([37])**.** *For any odd prime p, the Gupta–Sidki p-group G is indeed a p-group.*

[2]The first item is a generalization of a result in [58] for the Gupta–Sidki 3-group. The second is shown in [22] for GGS groups on the p-regular tree with p an odd prime.

Proof. Recall from (2.2) that every element $g \in G$ may be written as

$$(b^{r_1})^{a^{-s_0}}(b^{r_2})^{a^{-s_0-s_1}}\dots(b^{r_m})^{a^{-s_0-s_1-\dots-s_{m-1}}}a^{s_0+s_1+\dots+s_m}$$

where $s_i, r_i \in \mathbb{F}_p$ and all s_i, r_i except possibly s_0, s_m are non-zero. In other words, every element $g \in G$ is represented by a (reduced) word wa^v in $(\langle b_0\rangle * \dots * \langle b_{p-1}\rangle) \rtimes \langle a\rangle$ with $w \in \langle b_0\rangle * \dots * \langle b_{p-1}\rangle$. We shall prove that g has p-power order by induction on the syllable length of a shortest such representative of g.

If the syllable length is 1 then g is either a power of a or a power of some b_i. The former is clearly of order p. To see that b has order p, note that

$$b^p = (1, \dots, 1, b^p) = (1, \dots, 1, (1, \dots, 1, b^p)) = \dots$$

fixes every vertex that is not of the form $(p-1)(p-1)\dots(p-1)$ and, since b fixes all vertices of this form, so does b^p. Hence b^p fixes all vertices of T and must therefore be trivial.

Assume then that every element with a representative of syllable length at most n, where $n \geq 1$, has been shown to have p-power order. Suppose that g is represented by wa^v, which has syllable length $n+1$. There are two cases to consider:

a) $v \neq 0$. This means that w has syllable length n and $n = l_0 + \dots + l_{p-1}$ where l_i is the contribution to the syllable length due to b_i. Then $g^p \in \mathrm{St}_G(1)$ is represented by $ww^{a^{-v}}w^{a^{-2v}}\dots w^{a^v}$, which has syllable length at most np, and the contribution from each b_i is at most $l_0 + \dots + l_{p-1} = n$. Thus, each vertex section $(g^p)_i$ has p-power order by inductive hypothesis and therefore g has p-power order too.

b) $v = 0$. Pick an arbitrary first level vertex section g_i. If $g_i \in \mathrm{St}_G(1)$ then g_i is represented by an element of syllable length $l_i \leq n$ and therefore has p-power order by inductive hypothesis. If $g_i \notin \mathrm{St}_G(1)$ then it is represented by an element of syllable length $l_i + 1$. The inductive hypothesis yields that g_i has p-power order if $l_i + 1 \leq n$. If $l_i + 1 = n + 1$ then the argument in the previous case shows that g_i must have p-power order. Thus each vertex section g_i has p-power order and therefore so does g. □

A more more sophisticated version of this argument enabled Vovkivsky to prove the following criterion for a GGS group defined on the p^n-regular tree to be an infinite p-group. In fact, Vovkivsky proves the necessity while the sufficiency is proved in a more general setting in [39].

Theorem 2.2 ([64])**.** *The GGS group $G = \langle a, b\rangle$ which acts on the p^n-regular tree and has defining vector $(e_1, e_2, \ldots, e_{p^n-1})$ is*

(i) *infinite if and only if either some e_i is not divisible by p or there exists k ($1 \le k < n$) and $i \in \{p^k, 2p^k, \ldots, p^n - p^k\}$ such that e_i is not divisible by p^{k+1};*

(ii) *a p-group if and only if for $k = 0, \ldots, n-1$,*

$$e_{p^k} + e_{2p^k} + \cdots + e_{p^n-p^k} \equiv 0 \mod p^{k+1}.$$

In particular, when $n = 1$, a GGS group is infinite if and only if the defining vector is non-zero and a p-group if and only if the sum of its entries is divisible by p.

Note, however, that when $n > 1$ and all entries in the defining vector are divisible by p, the GGS group, if it is infinite, will not be fractal. Rather, each vertex section will be a subgroup of finite index of G, generated by a^{p^k}, b for some k with $1 \le k < n$.

2.2.3 Branch structure and the just infinite property

Based on the methods of [38] (which is a continuation of [39]), Vovkivsky also proves the following criterion for an infinite torsion GGS group on the p^n-regular tree (that is, one which satisfies the conditions of Theorem 2.2) to be a branch group.

Theorem 2.3 ([64])**.** *The infinite torsion GGS group $G = \langle a, b\rangle$ which acts on the p^n-regular tree and has defining vector $(e_1, e_2, \ldots, e_{p^n-1})$ is a branch group if and only if e_i is not divisible by p for some $i \in \mathbb{Z}/p^n\mathbb{Z}$.*

Idea of proof. It is easy to see that G acts level-transitively (even if it is not torsion) if and only if some e_i is not divisible by p. This is because $\langle b\rangle$ acts transitively on the p^n vertices below i if and only if it acts like the full cyclic group $\mathbb{Z}/p^n\mathbb{Z}$, in other words, if e_i is a generator of this cyclic group. The action on the rest of the vertices of the second level is taken care of by the wreath product structure of $\mathrm{St}_G(1) \rtimes \langle a\rangle$. This can then be extended to the rest of the tree by the fact that G is fractal.

The crucial observation in the theorem (initially made by Gupta and Sidki in [38] for the group on the p-regular tree with defining vector $(1, -1, 1, -1, \ldots, 1, -1)$ and p odd) is that the group is actually *regular branch* over its second derived subgroup $G^{(2)}$ (in fact, this is the original motivation for the definition of regular branch). That is,

$\psi(G^{(2)}) > G^{(2)} \times \overset{p^n}{\cdots} \times G^{(2)}$ and therefore $\mathrm{St}_G(1) \geq \mathrm{rist}_G(1) \geq G^{(2)} \times \overset{p^n}{\cdots} \times G^{(2)}$. Now, since G is finitely generated and torsion, $|G : G^{(2)}|$ is finite and then

$$G \times \overset{p^n}{\cdots} \times G \geq \mathrm{St}_G(1) \geq \mathrm{rist}_G(1) \geq G^{(2)} \times \overset{p^n}{\cdots} \times G^{(2)}$$

implies that $\mathrm{rist}_G(1)$ has finite index in $\mathrm{St}_G(1)$ and hence in G. The argument is then repeated at vertices lower down the tree using the fractal nature of G. □

This almost immediately yields that any GGS branch group G as in the previous theorem is just infinite (which was also originally showed in [38] for the particular example studied there). The proof requires the following result by Grigorchuk to which we shall return in Chapter 4:

Theorem 2.4 ([28, Theorem 4])**.** *Every non-trivial normal subgroup of a branch group G contains $\mathrm{rist}_G(m)'$ for some $m \in \mathbb{N}$.*

Since every $\mathrm{rist}_G(m)$ has finite index in G, and G is finitely generated and torsion, $\mathrm{rist}_G(m)'$ also has finite index in G and therefore so does every non-trivial normal subgroup of G, by the above theorem.

In fact, we have just shown one implication of the following equivalence.

Theorem 2.5. *A branch group G is just infinite if and only if every $\mathrm{rist}_G(n)$ has finite abelianization.*

The other implication follows from the fact that the derived subgroup of any group is characteristic and therefore each $\mathrm{rist}_G(n)'$ is normal in G, hence of finite index.

In fact, in the case $n = 1$ we can do without the torsion condition and obtain that most defining vectors yield branch groups. Note that when $n = 1$ and $p = 2$ the only infinite GGS group $\langle a, b = (a, b)\rangle$ is the infinite dihedral group. Thus, from now on, p will be an odd prime. Following [52] and [22], we make the following definition.

Definition 2.2.4. Let $(e_1, \ldots, e_{p-1})$ be a vector in $\mathbb{F}_p^{p-1}$. The vector is *symmetric* if $e_i = e_{p-i}$ for $i = 1, \ldots, p-1$ and *non-symmetric* otherwise; it is *constant* if all e_i are equal and *non-constant* otherwise.

For the case of a torsion GGS group G, Lemma 1.2 of [52] (which is actually a result of Rozhkov, found in his habilitation thesis [55]), states that, if the defining vector is non-symmetric, then $\gamma_3(G) \geq G' \times \overset{p}{\cdots} \times G'$. The following results do not require the torsion condition.

Proposition 2.2.5 ([22]). *Let G be a GGS group with non-constant defining vector. Then*

$$\psi(\gamma_3(\mathrm{St}_G(1))) = \gamma_3(G) \times \overset{p}{\cdots} \times \gamma_3(G).$$

In particular, G is a regular branch group over $\gamma_3(G)$.

Moreover,

Proposition 2.2.6 ([22]). *Let G be a GGS group with non-symmetric defining vector. Then*

$$\psi(\mathrm{St}_G(1)') = G' \times \overset{p}{\cdots} \times G'.$$

In particular, G is a regular branch group over G'.

What about the constant vector case? It turns out that these groups are not regular branch but only weakly regular branch.

Proposition 2.2.7 ([22]). *Let G be a GGS group with constant defining vector and let $K := \langle ba^{-1}\rangle^G$. Then $\psi(K') \geq K' \times \overset{p}{\cdots} \times K'$ and G is weakly regular branch over K'.*

As an example, take the group defined by the vector $(1,1) \in \mathbb{F}_3^2$. It is shown in [9, Proposition 7.23] that K is torsion-free and therefore G cannot be just infinite. It is then reasonable to wonder what is the relation between being (weakly) regular branch and just infinite. Let us explore this now.

Another of the striking results of [38] is that the p-group with defining vector $(1,-1,1,-1\ldots,1,-1) \in \mathbb{F}_p^{p-1}$ contains a copy of every finite p-group. This relies on the fact that every finite p-group is embeddable in an iterated wreath product of cyclic groups of order p, so it suffices to show that the GGS group in question contains arbitrarily long iterated wreath products of such cyclic groups. This in turn is done by exploiting the fact that the group is torsion and regular branch over its commutator subgroup.

This can be generalized, following the approach of [29] (the exposition here is taken from [9, Section 6]). Note that here the tree T need not be regular, only level-homogeneous.

Lemma 2.2.8 ([9, Lemma 6.8])**.** *Let π be a set of primes. If $G \leq \operatorname{Aut} T$ is such that every $\operatorname{rist}_G(v)$ has non-trivial elements of order divisible by members of π, then G contains arbitrarily long wreath products of cyclic groups of the form*

$$C_{p_1} \wr C_{p_2} \wr \cdots \wr C_{p_n} \quad p_i \in \pi, n \in \mathbb{N}.$$

As a corollary of this lemma, we have the following connections between (weakly) branch groups and torsion.

Corollary 2.2.9. *Let $G \leq \operatorname{Aut} T$. If every $\operatorname{rist}_G(v)$ has an element of finite order then G has elements of unbounded finite order. Moreover,*

(i) *Every weakly branch group which is also torsion has unbounded exponent. Indeed, it generates the variety of all groups, i.e., it satisfies no non-trivial identical relation.*

(ii) *Every weakly branch group which is also a p-group for some prime p contains a copy of every finite p-group.*

(iii) *Every weakly regular branch group either is weakly branched over a torsion-free group or contains a copy of every finite p-group for some prime p.*

(iv) *Every regular branch group is either torsion-free or contains a copy of every finite p-group for some prime p.*

2.2.4 Pride minimality

In some sense, just infinite groups are quite "small" infinite groups, because they only have finite quotients. On the other hand, we think of free groups as quite "large" and so any group with free quotients should be considered "even larger". There is a notion of "largeness" of groups introduced by S. Pride given by a preorder $\preceq$ on groups. The definition commonly used now is a slightly modified one from [18]:

Definition 2.2.10. Let H, G be groups. Write $H \preceq G$ if there exist

1. a subgroup G_0 of finite index in G;

2. a subgroup H_0 of finite index in H and a finite normal subgroup Y of H_0;

3. an epimorphism from G_0 to H_0/Y.

Write $H \sim G$ if $H \preceq G$ and $G \preceq H$, and write $[G]$ for the equivalence class of G. The preorder $\preceq$ induces a preorder on the equivalence classes, the smallest being $[1]$, that of all finite groups. We say that G is *minimal* (or *atomic*) if $[G] \neq [1]$ and there are no classes strictly between $[G]$ and $[1]$. It is now common in the literature to call a group *large* if it is $\sim$-equivalent to the free group on two generators.

How does Pride's notion of largeness compare to our intuitive idea of just infinite groups being small? By the work of Wilson [65, 66], a given just infinite group is of one of four types:

1. Groups which are virtually a free abelian group of finite rank.
2. Groups which are virtually a direct product of finitely many copies of an infinite simple group.
3. Groups which are virtually a direct product of finitely many copies of a hereditarily just infinite group (that is, all its finite index subgroups are just infinite).
4. Just infinite branch groups.

Groups of the first three types are atomic if and only if, respectively, the rank is 1 and there is only one group in the direct product. In [49], Peter M. Neumann constructed (among other things) early examples of branch groups (although the definition had not appeared at the time), to answer some questions posed by Edjvet and Pride. As Neumann points out in Comment 5.4, his construction can be varied in many ways, one of which yields finitely generated just infinite branch groups which are *not* atomic. The question then becomes which just infinite branch groups are atomic? This was partially answered by Grigorchuk and Wilson in [35], where they proved the following.

Theorem 2.6 ([35])**.** *Let G be a regular branch group over the subgroup K so $K \leq \mathrm{rist}_G(v)$ for each $v \in T$. Suppose that $K' \geq \psi_n^{-1}(K \times \cdots \times K)$ for some n and that K is a subdirect product of finitely many just infinite groups each of which is commensurable with G. Then G is atomic.*

The theorem is then applied to show that all Gupta–Sidki p-groups are atomic. Very similar calculations show that all GGS groups with non-symmetric defining vector are also atomic.

2.2.5 Hausdorff dimension and the congruence subgroup property

The rest of this chapter is focused on GGS groups over the p-regular tree where p is an odd prime (because the case $p = 2$ yields either a finite group or the infinite dihedral group). One of the most obvious families of subgroups of GGS groups (or any group acting on a rooted tree) to study are the level stabilizers $\mathrm{St}_G(n)$. Since they have finite index, a natural question to ask is what are the sizes of the quotients $G_n := G/\,\mathrm{St}_G(n)$, which we call the *congruence quotients*. These were computed for special types of GGS groups by several authors:

For the Gupta–Sidki 3-group, Sidki himself showed in [58] that

$$\log_3 |G_n| = 2 \cdot 3^{n-2} + 1, \text{ for every } n \geq 2.$$

For the group with defining vector $(1, 1)$ studied by Bartholdi and Grigorchuk in [8], they show that

$$\log_3 |G_n| = \frac{3^n + 2n + 3}{4}, \text{ for every } n \geq 2.$$

More generally, Šunić showed (among other things)[3] in [61] that for the case with defining vector $(1, 0, \dots, 0)$, we have for all $n \geq 2$:

$$\log_p |G_n| = p^{n-1} + 1 \text{ if } p \geq 3.$$

The size of congruence quotients for all GGS groups over the p-regular tree was computed in [22]. In order to state the theorem we need some terminology which will be used in the next chapter as well.

Definition 2.2.11. Given a vector $\mathbf{a} = (a_1, \dots, a_n)$ over some field, the *circulant matrix* $\mathrm{C}(\mathbf{a})$ is the $n \times n$ matrix with $\mathbf{a}$ in its first row and whose other rows are obtained by cyclically permuting $\mathbf{a}$ as shown below:

$$\mathrm{C}(\mathbf{a}) = \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ a_n & a_1 & \dots & a_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & \dots & a_1 \end{pmatrix}.$$

[3] In fact, Šunić defined a large class of groups that are close relatives to the first Grigorchuk group, of which GGS groups with vector $(1, 0, \dots, 0)$ are only a special case. These come under the remit of *spinal groups* which were also investigated in [11] and [21].

Theorem 2.7 ([22]). *Let G be a GGS group with defining vector $\mathbf{e} = (e_1, \ldots, e_{p-1}) \in \mathbb{F}_p^{p-1}$. For every $n \geq 2$ we have*

$$\log_p |G_n| = tp^{n-2} + 1 - \delta \frac{p^{n-2} - 1}{p - 1} - \varepsilon \frac{p^{n-2} - (n-2)p + n - 3}{(p-1)^2},$$

where t is the rank of the circulant matrix $\mathrm{C}(e_1, \ldots, e_{p-1}, 0)$ and

$$\delta = \begin{cases} 1 & \text{if } \mathbf{e} \text{ is symmetric,} \\ 0 & \text{otherwise} \end{cases} \qquad \varepsilon = \begin{cases} 1 & \text{if } \mathbf{e} \text{ is constant,} \\ 0 & \text{otherwise.} \end{cases}$$

This immediately allows for the determination of the Hausdorff dimension of the closure of such GGS groups in $\operatorname{Aut} T$. Hausdorff dimension, a key concept in fractal geometry ([20]), is a measure of the prominence of the irregularities of a fractal when viewed at very small scales. It is usually defined over $\mathbb{R}$ (because we view fractals as subsets of real space), but can be defined in the same way for any metric space. In particular, following the seminal work of Abercrombie ([1]), and Barnea and Shalev ([4]), considerable attention has been given to metrics arising from profinite groups.

The Hausdorff dimension of a closed subgroup of a profinite group is a measure of the relative size of the subgroup in the profinite group. More precisely, given an infinite, countably based, profinite group Γ and a filtration $(\Gamma_n)_n$ (that is, a descending chain of open subgroups such that $\bigcap_n \Gamma_n = 1$), define an invariant metric on Γ with respect to this filtration by

$$d(x, y) = \inf\{\frac{1}{|\Gamma \colon \Gamma_n|} \mid xy^{-1} \in \Gamma_n\}.$$

With this metric, we can define Hausdorff dimension analogously to the case of subsets of $\mathbb{R}^n$ (see [20] or [4]), but it is not a nice definition to work with in the context of profinite groups. Luckily, by results of Abercrombie, and Barnea and Shalev, it is equivalent to the following, which we will take as the definition.

Definition 2.2.12. Given a profinite group Γ equipped with a filtration $(\Gamma_n)_n$, the *Hausdorff dimension* of a closed subgroup $H \leq \Gamma$ is

$$\dim(H) = \liminf_{n \to \infty} \frac{\log |H\Gamma_n \colon \Gamma_n|}{\log |\Gamma \colon \Gamma_n|}.$$

Intuitively, it is the prominence of the irregularities of H when viewed at very small scales within Γ, where the scale is determined by the filtration.

The choice of filtration is important. Indeed, different filtrations may yield different dimension functions for the same group, as demonstrated by an easy example given in [4, Example 2.5]. However, in many cases there are natural filtrations that ought to be used to extract meaningful information. This is the case of countably based pro-p groups (studied in [4]) where a natural filtration is $(\Gamma^{p^n})_n$ and is of course also the case for countably based (closed) subgroups of $\operatorname{Aut} T$ where the filtration is $(\mathrm{St}(n))_n$.

In fact, all the examples we are dealing with here are not just subgroups of $\operatorname{Aut} T$ but of its pro-p-Sylow subgroup $\mathcal{A}_p$ which is the inverse limit of finite wreath products W_n where $W_{n+1} = C_p \wr W_n$ and C_p is the cyclic group of order p. This makes the calculation of Hausdorff dimension easier.

Theorem 2.8 ([22])**.** *Let G be a GGS group with defining vector $\mathbf{e} \in \mathbb{F}_p^{p-1}$ and let $\bar{G}$ be its closure in $\mathcal{A}_p$. The Hausdorff dimension of $\bar{G}$ is $\mathcal{A}_p$ is*

$$\dim(\bar{G}) = \frac{(p-1)t}{p^2} - \frac{\delta}{p^2} - \frac{\varepsilon}{(p-1)p^2}$$

with t, δ, ε as in Theorem 2.7.

Since we know the size of congruence quotients it is natural to wonder how much information this gives us about all the other finite index subgroups of GGS groups. More concretely, we may ask whether every finite index subgroup contains some level stabilizer or, equivalently, whether every finite quotient of a GGS group factors through a congruence quotient. If this holds, we say that the group has the *congruence subgroup property* (more about this will be said in Chapter 6).

The next theorem generalizes [8, Proposition 8.4], whose proof contained an inaccuracy. Before stating it, we require a lemma.

Lemma 2.2.13. *Let G be a GGS group with non-constant defining vector $(e_1, \ldots, e_{p-1})$. Then $G' \leq_{\mathrm{s}} G \times \overset{p}{\ldots} \times G$.*

Proof. Since the defining vector is non-constant, at least one of the middle entries of $[a, b] = (b^{-1}a^{e_1}, a^{-e_1+e_2}, \ldots, a^{-e_{p-2}+e_{p-1}}, a^{-e_{p-1}}b)$ involving only powers of a is non-trivial. Choose j such that $[a, b]^{a^j}$ has this non-trivial power of a in its first coordinate and k such that $y := ([a, b]^{a^j})^k$ has a in its first coordinate. Now, $([a, b]y^{-e_1})^{-1}$ has b in its first coordinate. Thus, the first vertex section of G' is equal to G. Conjugating by powers of a, we obtain that $G' \leq_{\mathrm{s}} G \times \overset{p}{\ldots} \times G$. □

Theorem 2.9. *All GGS groups with non-symmetric defining vector over the p-regular tree have the congruence subgroup property.*

Proof. This is a consequence of results in [22]. By Proposition 2.2.6, if G is a GGS group with non-symmetric defining vector, then $\psi(\mathrm{St}_G(1)') = G' \times \cdots \times G'$ so G is regular branch over G'. This means that every $\mathrm{rist}_G(v)$ contains G' and, since every finite index subgroup contains some $\mathrm{rist}_G(n)'$ by Theorem 2.4, it suffices to show that $G'' \geq \mathrm{St}_G(m)$ for some m. Corollary 2.5 of [22] states that $\gamma_3(G) \geq \mathrm{St}_G(2)$, so it will be enough to prove that $G'' \geq \gamma_3(G) \times \cdots \times \gamma_3(G) \geq \mathrm{St}_G(3)$.

First note that, again by Proposition 2.2.6, there exists some $x \in \mathrm{St}_G(1)'$ such that $x = ([a,b], 1, \ldots, 1)$. By Lemma 2.2.13, there exist $y, z \in G'$ with a (respectively, b) in the first coordinate. Thus G'' contains

$$[x, y] = ([[a,b],a], 1, \ldots, 1) \text{ and } [x, z] = ([[a,b],b], 1, \ldots, 1).$$

Since $\gamma_3(G) = \langle [[a,b],a], [[a,b],b] \rangle^G$, we obtain that $G'' \geq \gamma_3(G) \times \overset{p}{\ldots} \times \gamma_3(G)$ by taking the normal closure of the above elements and conjugating by powers of a. □

Fernández-Alcober has informed me that he and his research student Uria have shown that GGS groups with symmetric non-constant defining vector also have the congruence subgroup property, while those with constant defining vector do not. This in particular implies that the closure in $\mathrm{Aut}\, T$ of a GGS group with non-constant defining vector is in fact its profinite completion.

I should also mention that Pervova proved in [52] that all *torsion* GGS groups on the p-regular tree have the congruence subgroup property. On the other hand, in that same paper, she proved that a family of groups closely related to GGS groups, which she calls EGS groups, do *not* have the congruence subgroup property. An EGS group is a group of automorphisms of the p-regular tree for p an odd prime, which is also defined by a vector $\mathbf{e} \in \mathbb{F}_p^{p-1}$. The group is generated by a, b, c where a, b are defined as for GGS groups and $c = (c, a^{e_1}, \ldots, a^{e_{p-1}})$. Theorem 3.1 of [52] states that no torsion EGS group with non-symmetric defining vector has the congruence subgroup property. However, the proof does not use the torsion assumption, so in fact the result holds for all EGS groups with non-symmetric defining vector.

2.2.6 Maximal subgroups

Another of the remarkable features of torsion GGS groups, witnessing their "finiteness", is the following result by Pervova, which will play a key role in the next chapter:

Theorem 2.10 ([51, Theorem 4.3])**.** *All maximal subgroups of torsion GGS groups are normal of prime index.*

The proof of this theorem is very technical, but it is in essence a clever use of length reduction arguments, which as with most length reduction arguments in the context of self-similar groups, are carried out for one element at a time. This is in contrast to the main result of the next chapter.

In a previous paper [50], Pervova also showed that all maximal subgroups of the Grigorchuk group are of index 2, which in turn played a crucial role in the analogous result of Chapter 3 by Grigorchuk and Wilson ([33]).

Furthermore, Alexoudas, Klopsch and Thillaisundaram in [3] were able to generalize Pervova's result to torsion multi-edge spinal groups. These are a generalization of GGS groups where there are several directed generators b, each with its own defining vector. They are torsion if and only if each defining vector $(e_1, \ldots, e_{p-1})$ satisfies $\sum e_i \equiv 0 \mod p$.

2.2.7 Width, lower central series and derived series

Another of the finiteness properties of some GGS groups, related to subgroup structure, is the fact that they have finite width.

Definition 2.2.14. Given any group G, its *width* is

$$\sup_n \operatorname{rank}(\gamma_n(G)/\gamma_{n+1}(G)).$$

Recall that the *rank* of an abelian group is the minimal number of generators. G has *finite width* if the above supremum is finite.

The study of groups of finite width is motivated by the classification of finite p-groups, as the definition of finite width is a generalization of finite coclass. Recall that a finite group of size p^n has coclass $n-c$ if its nilpotency class is c. A pro-p group has coclass $n-c$ if it is an inverse limit of finite groups of coclass $n-c$. The coclass of finite and pro-p groups has proved a fruitful invariant for their study. While all pro-p groups of finite coclass are p-adic analytic, this is not the case for pro-p groups of finite width as the following results show:

Theorem 2.11 ([56]). *Let* Γ *denote the first Grigorchuk group. Then*

$$\operatorname{rank}(\gamma_n(\Gamma)/\gamma_{n+1}(\Gamma)) = \begin{cases} 3 & \text{if } n = 1, \\ 2 & \text{if } n = 2^m + 1 + r,\ 0 \le r < 2^{m-1}, \\ 1 & \text{if } n = 2^m + 1 + r,\ 2^{m-1} \le r < 2^m. \end{cases}$$

and so Γ *(and its profinite completion) has finite width.*

Theorem 2.12 ([6]). *The Fabrykowski–Gupta group* $G = \langle a, b\rangle$ *with* $b = (a, 1, b)$ *has finite width (and therefore so does its profinite completion).*

This might lead one to think that all GGS groups have finite width. However, this is not the case even for the torsion ones:

Theorem 2.13 ([6]). *The Gupta–Sidki 3-group* G *is not of finite width. The asymptotic growth of* $\operatorname{rank}(\gamma_n(G)/\gamma_{n+1}(G))$ *is polynomial of degree* $(\log 3/\log(1+\sqrt{2}))-1$*. Moreover,* $|G\colon \gamma_{\beta_n+1}(G)| = 3^{(3^n+1)/2}$ *where*

$$\beta_n = \frac{1}{4}((1+\sqrt{2})^{n+1} + (1-\sqrt{2})^{n+1}) \text{ for } n \in \mathbb{N}.$$

Perhaps this difference in the behaviour of the lower central series is related to the fact that the Fabrykowski–Gupta group, like the Grigorchuk group, is of "spinal type" in the terminology introduced in [11]; in particular, its defining vector is of the form $(1, 0, 0, \ldots, 0)$. It would be interesting to apply the methods of [6] to GGS groups in general to see exactly which defining vectors yield GGS groups of finite width, and determine the growth of $\operatorname{rank}(\gamma_n(G)/\gamma_{n+1}(G))$ for those of infinite width.

The results of [6] yield that the derived series of the Fabrykowski–Gupta group satisfies

$$G^{(k)} = \gamma_5(G)^{\times 3^{k-2}} \text{ for } k \ge 2 \quad \text{and} \quad G^{(k)} \le \gamma_{2+3^{k-1}}(G) \text{ for } k \in \mathbb{N}.$$

For the Gupta–Sidki 3-group, Vieira showed in [63] that $G^{(k)} = \gamma_5(G)^{\times 3^{k-2}}$ for all $k \ge 2$. This, in conjunction with the result on width in [6] yields that $G^{(k)} \le \gamma_{\alpha_{k+1}}(G)$ for all $k \in \mathbb{N}$, where

$$\alpha_k = \frac{1}{2\sqrt{2}}((1+\sqrt{2})^k + (1-\sqrt{2})^k).$$

In fact, Vieira also showed in [63] that $G^{(k+1)} \leq \mathrm{St}_G(k+1) \leq G^{(k)} \leq \mathrm{St}_G(k)$ for all $k \geq 1$ and $|G\colon G^{(k)}| = 3^{5\cdot 3^{k-2}+1}$ for all $k \geq 2$.

However, as far as I am aware, the lower central and derived series of general GGS groups have not been investigated. This would of course be a natural step in deciding which GGS groups have finite width.

2.3 Algorithmic properties

Since M. Dehn introduced in 1911 the word, conjugacy and isomorphism problems for groups given by finite presentations, they have been intensely studied, even outside the realm of finitely presented groups. Recall that the *word problem* of a group is *solvable* if there is an algorithm that determines whether two given words in the generators of the group define the same group element (equivalently, whether a given word in the generators defines the identity). The *conjugacy problem* asks for an algorithm that determines whether or not two words define conjugate elements. The *membership problem*, introduced later, is a natural generalization of the word problem (which is why it is sometimes called the *generalized word problem*). It is solvable if there is an algorithm which determines, given finitely many elements $x, g_1, \ldots, g_n$, whether or not x is in the subgroup generated by $\{g_1, \ldots, g_n\}$. The *isomorphism problem* asks whether there is an algorithm that determines whether two group presentations yield isomorphic groups.

None of these problems are solvable in general so it is interesting to know for which groups (or classes of groups) they are solvable.

Separability properties play a key role in the solution of the word, conjugacy and membership problems. Intuitively, the notion of separability indicates whether subsets of the group can be "separated" in a finite quotient of the group, or equivalently, in the profinite topology. Recall that the profinite topology of a group has as a neighbourhood basis of the identity all normal finite index subgroups of the group. For instance, a group is residually finite if and only if every non-trivial element can be separated from the identity in a finite quotient; in other words, there is a finite quotient of the group where the non-trivial element is still non-trivial. Put differently, the identity is closed in the profinite topology of the group. We might call a residually finite group $\{1\}$-separable instead, but this latter term does not have the same ring to it. Similarly, a group is *conjugacy separable* if every conjugacy class is closed in the profinite topology (which implies that the group is residually finite). This means that if two elements of a conjugacy separable group are not conjugate, there is a finite quotient of the group in which their images are also not conjugate. A group is *subgroup separable* (or LERF, which stands for locally extended residually finite) if every finitely generated subgroup is closed in the profinite topology. 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 [40, 57, 46, 44]).

It is well known that if a group is finitely presented within a finitely based variety (one which is determined by a finite number of laws) and is residually finite (respectively, conjugacy separable, subgroup separable) then the group has solvable word (respectively, conjugacy, membership) problem. This is because, since the group is finitely presented there is an algorithm that decides whether or not a finite group is a homomorphic image of the group. However, it seems improbable that GGS groups are finitely presented within any variety – indeed, weakly branch groups do not satisfy any group law ([41, Theorem 14.1.5]) – so it is interesting to know for which of them various algorithmic problems are solvable. On the other hand, GGS groups have particularly nice presentations that allow for many of their properties to be investigated algorithmically.

2.3.1 Endomorphic presentations

It is known that the Gupta–Sidki 3-group is not finitely presentable [59], another feature in common with Grigorchuk's group. However, they both have a finite endomorphic presentation.

Definition 2.3.1. An *endomorphic presentation* is an expression of the form

$$L = \langle S | Q | \Phi | R \rangle$$

where S is a set of symbols, $Q, R \subset F_S$ are sets of reduced words with F_S denoting the free group on S and Φ is a set of group homomorphisms $\varphi : F_S \to F_S$. The presentation L is *finite* if Q, R, S, Φ are finite. Such a presentation gives rise to a group

$$G_L = F_S / \langle Q \cup \bigcup_{\varphi \in \Phi^*} \varphi(R) \rangle^{F_S}$$

where Φ^* denotes all finite compositions of homomorphisms in Φ.

The intuition behind this definition is that some of the defining relations can be mapped down the tree by the vertex sections (as is reflected in the use of the same notation).

An endomorphic presentation with exactly one homomorphism in Φ is called an L-*presentation*. The L stands for Lysenok who found in [45] such a presentation for the Grigorchuk group.

More generally, by a result of Bartholdi ([5]) every finitely generated, contracting, regular branch group has a finite endomorphic presentation but is not finitely presentable. A *contracting* self-similar group is one for which there are constants $C \geq 0$ and $\lambda < 1$ such that for every word $\psi(w) = (w_1, w_2, \ldots, w_d)$ in the generators representing an element in $\mathrm{St}_G(1)$,

$$|w_i| < \lambda |w|, i = 1, \ldots, d \text{ whenever } |w| > C,$$

where $|w|$ is the word length of w. This includes all GGS groups on the p-regular tree with non-constant vector.

Bartholdi gives in the appendix of [7] endomorphic presentations for GGS groups over the p-regular tree for p odd with defining vectors $(1, 0, 0, \ldots, 0)$ and $(1, 2, \ldots, p-1)$, which generalize the Fabrykowski–Gupta group and Gupta–Sidki 3-group, respectively. These presentations correct some errors in [5].

No finite presentation has been found for any branch group and it is in fact conjectured that no finitely generated branch group is finitely presentable. Nevertheless, finite endomorphic presentations are sufficiently close to finite presentations for groups given by finite endomorphic presentations to enjoy similar properties to finitely presented groups. In particular, the following theorem of Hartung[4] shows that the set of finite quotients of a group with a finite endomorphic presentation is recursive (there is an algorithm that decides whether or not a finite group is a quotient).

Theorem 2.14 ([42, Theorem 6])**.** *Let G be finitely endomorphically presented by $\langle S|Q|\Phi|R\rangle$ and denote the free group over S by F. There exists an algorithm which decides whether or not a homomorphism $\phi : F \to \mathrm{Sym}(n)$ induces a homomorphism $G \to \mathrm{Sym}(n)$.*

[4]Note that Hartung does not distinguish between "L-presentation" and "endomorphic presentation", calling them both "L-presentation".

Therefore, by the same arguments as for the case of finitely presented groups, any group with a finite endomorphic presentation which is residually finite (respectively, conjugacy separable, subgroup separable) has solvable word (respectively, conjugacy, membership) problem.

2.3.2 Word problem

The word problem is simple enough not to require Theorem 2.14, a straightforward direct argument sufficing. The word problem is solvable for all groups of spinal type (as defined in [9]) by a fast algorithm first worked out in [27] for Grigorchuk groups, and described in full generality in [9]. The particular case of GGS groups goes as follows. Let $w = a^{i_0}b^{j_1}a^{i_1}\dots a^{i_{k-1}}b^{j_k}a^{i_k}$ be a reduced word in the generators a, b (where i_0, i_k may be 0).

Step 1 Check that $a^{i_0}a^{i_1}\dots a^{i_{k-1}}a^{i_k} = 1$ in $\langle a\rangle$ (otherwise, the element represented by w is not in $\mathrm{St}_G(1)$ and cannot be the identity).

Step 2 Write w in the form $b_{g_1}^{j_1}\dots b_{g_k}^{j_k}$ where $g_n = -i_0 - i_1 \cdots - i_{n-1}$ and use the definition of b (that is, the defining vector) to find the (possibly unreduced) first level vertex projections $w_x := \varphi_x(w)$ of w. Then, each w_x has syllable length at most $\lceil \frac{|w|}{2} \rceil$ (as seen in Lemma 2.2.1).

Step 3 Repeat Steps 1 and 2 until either one of the vertex projections w_v is not in $\mathrm{St}_G(1)$ (and therefore w is non-trivial), or one of them has syllable length 1 (and is therefore a non-trivial power of b, hence w is non-trivial), or all the vertex sections w_v at some level m visibly represent the identity (because their length is short enough). The last case implies that w is the identity because the map $\psi_m : \mathrm{St}_G(m) \to G \times \cdots \times G$ is an embedding.

2.3.3 Conjugacy problem

Despite the very successful treatment of the word problem, the conjugacy problem has been harder to crack. First, it was shown in [67] that for p an odd prime, all torsion GGS groups which are regular branch over some subgroup K are conjugacy separable (the theorem is not stated like this but this is the right interpretation in our context).

The theorem essentially depends on two things. First, the observation that if for every $g \in G$ we have

$$g^{\bar{K}} \cap G = g^K \tag{2.3}$$

(where $\bar{K}$ is the closure of K in $\hat{G}$, the profinite completion of G) then $g^{\hat{G}} \cap G = g^G$ and G is conjugacy separable. This is because $\hat{G}/\bar{K} \cong G/K$ so $\hat{G}$ is the union of finitely many cosets $\bar{K}h$ with $h \in G$, and for each of them $g^{\bar{K}h} \cap G = (g^{\bar{K}} \cap G)^h \subseteq g^{Kh}$. Second, a proof by induction on length of elements of G that they all satisfy (2.3): there is some length function and a constant C such that elements of length at most C already satisfy the required condition and for any element $g = b_1 \dots b_k a^r$ of length larger than C either $r = 0$ and the first level vertex sections g_i have length shorter than g or $r \neq 0$ and then each $g_i g_{i+r} \cdots g_{i+(p-1)r}$ (which is the vertex section $(g^p)_i$) has shorter length than g. Thus, given an element g of length larger than C, we can obtain by repeating the above procedure "pseudo-vertex sections" which satisfy (2.3) and then some work needs to be done to "undo the vertex section maps" and obtain that g itself satisfies (2.3).

The length function mentioned above is in this case a modified syllable length which takes into account the (finite) order of the element. Thus if $g = b_{i_1}^{r_1} \cdots b_{i_m}^{r_m} a^v$ then the length of g in this context is $m +$ (order of g). A similar modified syllable length will be used in the next chapter.

The above was generalized to a much larger class of branch groups in [32], where a more sophisticated implementation of these length-reduction ideas, exploiting the regular branch property, was used to prove that these groups have solvable conjugacy problem. Moreover, it is shown that the conjugacy problem is also solvable for their subgroups of finite index (which is not true in general). Thus all GGS groups over the p-regular tree which are regular branch (that is, those with non-constant vector) have solvable conjugacy problem. Note, however, that the case of non-constant vector is still open as the results of [32] do not apply to it.

2.3.4 Membership problem

The only branch group, until now, which is known to be subgroup separable is the Grigorchuk group ([33]). This is the analogue of the main result in the next chapter. The proof follows similar length reduction strategies as the ones outlined above. The difference from the previous results on algorithmic properties is that the length reduction has to be carried out for finitely many elements (the generators of the finitely generated subgroup) simultaneously, which makes the reduction very delicate and hard to extrapolate to other contexts.

Isomorphism problem Very little attention has been given to the isomorphism problem in the context of GGS groups. If two defining vectors are non-zero scalar multiples of each other then they define the same GGS group. It would be interesting to know which defining vectors (apart from the ones already mentioned) yield isomorphic GGS groups.

2.3.5 GAP packages

Finally, the following GAP packages should be mentioned as very useful tools for investigating GGS or self-similar groups in general. The first one, in particular, was used to experimentally verify Lemmas 3.4.1, 3.4.2 and 3.4.3.

- Automata groups (AutomGrp) `http://www.gap-system.org/Packages/automgrp.html`
- Functionally recursive groups (FR) `http://www.gap-system.org/Packages/fr.html`

Chapter 3

Abstract commensurability and the Gupta–Sidki group

3.1 Introduction

In [33], the authors establish a 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 *(abstractly) 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 [33], we will see that an analogous result holds for the Gupta–Sidki 3-group:

Theorem 3.1. *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 Chapter 4. There, 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 3.1. It seems likely that this restriction is unnecessary. However, the proof relies heavily on a delicate length reduction argument that is only available for $p = 3$.

The first main theorem facilitates another result about the subgroup structure of G, also parallel to a result in [33].

Theorem 3.2. *The Gupta–Sidki 3-group is subgroup separable and hence has solvable membership problem.*

Recall that the "hence" above uses Theorem 2.14.

The proofs of the main theorems both rely on an auxiliary result, Theorem 3.4, on finitely generated subgroups of the Gupta–Sidki 3-group. It is worth mentioning that all the results hold for the general case where p is any odd prime, except for the length reduction argument in Theorem 3.4. For this reason, all preliminary results are stated for the general case in Section 3.2, leaving the restriction to $p = 3$ for the proof of Theorem 3.4 in Section 3.4. We depart from preliminary results in Section 3.3, where maximal subgroups are discussed. In [50, 51] it was shown that all maximal proper subgroups of the Grigorchuk and Gupta–Sidki groups have finite index. Theorem 3.3 shows that this maximal subgroups property passes to all finitely generated subgroups of the Gupta–Sidki 3-group, as a consequence of Theorem 3.1. The proofs of Theorem 3.1 and Theorem 3.2 are presented in the final Section 3.5. The arguments generalize easily to the case $p > 3$ assuming that the analogue of Theorem 3.4 holds. In this final section, the analysis carried out in Section 3.4 is used to establish the following characterization of finite subgroups of the Gupta–Sidki 3-group:

Theorem 3.5. *Let H be a finitely generated subgroup of the Gupta–Sidki 3-group G. Then either H is finite or one of its vertex sections is equal to G.*

Minor changes in the proof of this theorem and the detailed analysis carried out in [33] yield an identical characterization of the finite subgroups of the Grigorchuk group, whose proof can be found in Appendix A.

Theorem A.2. *Let H be a finitely generated subgroup of the Grigorchuk group Γ. Then either H is finite or one of its vertex sections is equal to Γ.*

This can be interpreted as a "strong fractal" property, in the sense that every infinite subgroup has some vertex section which is equal to the whole group. Thus the actions of G and Γ are reproduced somewhere down the rooted trees by each of their respective infinite subgroups.

3.2 Preliminary results

Notation. Throughout this section, G will denote a GGS group over the p-regular rooted tree where p is an odd prime. Its defining vector will be specified according to the needs of each result. Both $\mathrm{St}_G(1)$ and B may be used to denote the first level stabilizer of G. Also, $H \leq_{\mathrm{s}} G_1 \times \cdots \times G_n$ will mean that H is a subdirect product of the groups $G_1, \ldots, G_n$.

Recall that Proposition 2.2.5 stated that whenever G has non-constant defining vector, it is regular branch. This yields the following commensurability result.

Proposition 3.2.1. *Let G have non-constant defining vector. If $i \equiv j \mod p-1$ then $G^{\times i}$ is commensurable with $G^{\times j}$.*

Proof. For fixed $i \in \{0, \ldots, p-1\}$, we show by induction on n that $G^{\times i}$ is commensurable with $G^{\times n(p-1)+i}$. For the base case, we deduce from Proposition 2.2.5 that $G^{\times p-1+i} = G^{\times p} \times G^{\times i-1}$ is commensurable with $G \times G^{\times i-1} = G^{\times i}$. Now suppose the claim holds for n. Then $G^{\times (n+1)(p-1)+i} = G^{\times n(p-1)+i} \times G^{\times p-1}$ is commensurable with $G^{\times i} \times G^{\times p-1}$, which is in turn commensurable with $G^{\times i-1} \times G = G^{\times i}$. Hence $G^{\times (n+1)(p-1)+i}$ is commensurable with $G^{\times i}$ and the claim follows by induction. $\square$

The following will be very useful for the proofs of Theorems 3.1 and 3.2. It came out of conversations with B. Klopsch and A. Thillaisundaram. Furthermore, it also holds for torsion multi-edge spinal groups (those whose defining vectors have zero sum modulo p), by a similar proof.

Lemma 3.2.2. *Let G have non-constant defining vector $(e_1, \ldots, e_{p-1})$ with $\sum_{i=1}^{p-1} e_i \equiv 0 \mod p$. Then every finite index subgroup of G contains another finite index subgroup which is a subdirect product of finitely many copies of G.*

Proof. Let H be a finite index subgroup of G. Then the quotient of G by the normal core in G of H is a finite p-group (since G is a p-group by Theorem 2.2) and therefore nilpotent, in particular, soluble. Thus, H contains some derived subgroup $G^{(n)}$ of G.

Now, $G' \leq_{\mathrm{s}} G^{\times p}$ by Lemma 2.2.13 and it follows by a simple induction that $G^{(n)} \leq_{\mathrm{s}} G^{\times p^n}$ for all $n \in \mathbb{N}$, which finishes the proof. $\square$

Let us record the following observation for future use. It is a generalization of Proposition 2.2.2 of [58] and also holds for multi-edge spinal groups whose defining vectors all satisfy the hypothesis.

Proposition 3.2.3. *Let G have non-symmetric defining vector $(e_1, \dots, e_{p-1})$. An element $g \in G$ is in B if and only if there exist $n_0, \dots, n_{p-1} \in \mathbb{F}_p$, $c_0, \dots, c_{p-1} \in G'$ and $m_0, \dots, m_{p-1} \in \mathbb{F}_p$ such that*

$$g = (a^{m_0} b^{n_1} c_0, \dots, a^{m_{p-2}} b^{n_{p-1}} c_{p-2}, a^{m_{p-1}} b^{n_0} c_{p-1}).$$

Moreover, when this is the case, $m_0, \dots, m_{p-1}$ are determined by $n_0, \dots, n_{p-1}$ and the circulant matrix $\mathrm{C}(e_1, \dots, e_{p-1}, 0)$, and the representation is unique.

Proof. An element with such a representation is obviously in B. For the converse, let $g \in B$. The abelianization of B is just $\mathbb{F}_p^p$ with basis vectors $b_0, \dots, b_{p-1}$ modulo B', so there exist unique $n_0, \dots, n_{p-1} \in \mathbb{F}_p$ and $c = (c_0, \dots, c_{p-1}) \in B'$ such that $g = b_0^{n_0} \dots b_{p-1}^{n_{p-1}} c$. By Proposition 2.2.6, $B' = G'^{\times p}$ so $c_i \in G'$ for $i = 0, \dots, p-1$. Now, each b_i is the $(i+1)$th row of the following array:

$$\begin{array}{cccc} a^{e_1} & \dots & a^{e_{p-1}} & b \\ b & a^{e_1} & \dots & a_{p-1} \\ \vdots & \vdots & \vdots & \vdots \\ a^{e_2} & \dots & b & a^{e_1} \end{array}$$

which makes it easy to see that the powers of a in the coordinates of b_i are just the $(i+1)$th row of the circulant matrix $\mathrm{C} := \mathrm{C}(e_1, \dots, e_{p-1}, 0)$. Thus, the vector $(m_0, \dots, m_{p-1})$ in the statement is exactly $(n_0, \dots, n_{p-1})\mathrm{C}$. $\square$

Lemma 3.2.4. *Let G have non-symmetric defining vector and be torsion (so the sum of the entries in the defining vector is 0 modulo p). Suppose that $H \leq G$ is not contained in B. Then either all first level vertex sections of H are equal to G, or they are all contained in B so that $\mathrm{St}_H(1) = \mathrm{St}_H(2)$.*

Proof. Denote by $H_0, \dots, H_{p-1}$ the first level vertex sections of H. We claim that they are all conjugate in G. Since H is not contained in B there exists $as \in H$ with $s = (s_0, \dots, s_{p-1}) \in B$. Thus, for every $h = (h_0, \dots, h_{p-1}) \in \mathrm{St}_H(1)$ we have $h^{as} = (h_{p-1}^{s_0}, h_0^{s_1}, \dots, h_{p-2}^{s_{p-1}}) \in \mathrm{St}_H(1)$. From this we see that $H_i = H_{i-1}^{s_i}$ for $i = 0, \dots, p-1$.

Hence we may assume that no H_i is equal to G. We examine the image of H modulo B' (equivalently, the images of the H_i modulo G'). Since all these vertex sections are conjugate in G, they must have the same images modulo G' and the possibilities for these are $\langle ab^k \rangle$ for some $k \in \mathbb{F}_p$, or $\langle b \rangle$, or $\langle a \rangle \langle b \rangle$. We first discard the last case by showing that if $H_i G' = \langle a \rangle \langle b \rangle G'$ then $H_i = G$. Suppose not, then H_i is contained in some maximal subgroup $M < G$, which by Theorem 2.10 has index p in G. Thus $M \geq G'$ and $G = \langle a \rangle \langle b \rangle G' = H_i G' \leq MG' = M$, a contradiction.

So either $H_i G' = \langle b \rangle G'$ for all i or there exists $k \in \mathbb{F}_p$ such that $H_i G' = \langle ab^k \rangle G'$ for all i. Pick some $h = (h_0, \dots, h_{p-1}) \in \mathrm{St}_H(1)$ and consider its image modulo B'. If there exist i, j with $i \neq j$ such that $h_i G'$ and $h_j G'$ lie in different cyclic subgroups of G/G' then, conjugating by suitable elements of $H \setminus \mathrm{St}(1)$ we obtain $H_i G' = H_j G' = \langle a \rangle \langle b \rangle G'$, a contradiction to the above. Thus all $h_i G'$ lie in the same cyclic subgroup of G/G'. Our ultimate goal is to show that this cyclic subgroup is $\langle b \rangle G'$, so suppose for a contradiction that there is some $k \in \mathbb{F}_p \setminus \{0\}$ and some $(r_0, \dots, r_{p-1}) \in \mathbb{F}_p^p$ such that

$$(h_0 G', \dots, h_{p-1} G') = ((ab^k)^{r_0} G', \dots, (ab^k)^{r_{p-1}} G').$$

Now, by Proposition 3.2.3,

$$(h_0 G', \dots, h_{p-1} G') = (a^{m_0} b^{n_1} G', \dots, a^{m_{p-1}} b^{n_0} G')$$

where $(m_0, \dots, m_{p-1}) = (n_0, \dots, n_{p-1})\mathrm{C}$ and C is the circulant matrix $\mathrm{C}(e_1, \dots, e_{p-1}, 0)$. Comparing these representations of hB', we obtain the following equations describing the exponents of a and b:

$$(n_0, \dots, n_{p-1})\mathrm{C} = (r_0, \dots, r_{p-1}) \text{ and } (n_0, \dots, n_{p-1})\mathrm{P} = k(r_0, \dots, r_{p-1})$$

where P is the permutation matrix corresponding to $(1\,2\,\dots\,p)$. These are equivalent to $(n_0, \dots, n_{p-1})(k\mathrm{C}\mathrm{P}^{-1} - \mathrm{I}) = (0, \dots, 0)$. The matrix $k\mathrm{C}\mathrm{P}^{-1} - \mathrm{I}$ is still a circulant matrix over $\mathbb{F}_p$. Its determinant is equal to the sum of the entries in its first row, as we are working in $\mathbb{F}_p$. This sum is $-1 + k\sum_{i=1}^{p-1} e_i \equiv -1 \mod p$ because we assumed that the defining vector of G has zero sum. Hence $(n_0, \dots, n_{p-1}) = (0, \dots, 0)$ is the only solution and $h \in B'$, which finishes the proof of the lemma. □

The proof of Theorem 3.1 relies on a word length reduction argument. Recall the syllable length of elements of G introduced in Theorem 2.1: An element $g = b_{i_1}^{r_1} \cdots b_{i_m}^{r_m} a^v$ where $b_{i_j} \neq b_{i_{j+1}}$ for each j and $v \in \{0, \dots, p-1\}$ was defined to have syllable length m if $v = 0$ and $m+1$ otherwise. In the following a modified version of this will be used, which for simplicity will be called *length*, as no other length functions will appear.

Definition 3.2.5. Define the *length* $l(g)$ of $g = b_{i_1}^{r_1} \cdots b_{i_m}^{r_m} a^v$ to be m (independently of the value of v).

The next lemma records how the length of elements of G is reduced as we project down levels of the tree. Recall from the previous chapter that $\varphi_v : \mathrm{St}(v) \to \mathrm{Aut}\, T$ is the homomorphism which maps $g \in \mathrm{St}(v)$ to its vertex section g_v (its restriction to the subtree rooted at v).

Lemma 3.2.6. *Let $g \in \mathrm{St}_G(1)$ and suppose that $l(g) = m$. Then*

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

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

for $k, j \in \{0, \dots, p-1\}$.

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

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

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

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

(d) $b^{r_1} a^{r_2} \cdots b^{r_m}$, so $l(\varphi_k(g)) \leq \frac{m+1}{2}$.

This proves (i).

If $\varphi_k(g) \in \mathrm{St}(1)$ 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 $j \in \{0, \dots, p-1\}$. ∎

3.3 Maximal subgroups

This section establishes some results about groups whose maximal subgroups all have finite index. Once Theorem 3.1 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 [33].

The following is Lemma 1 of [33], but the proof given here is different and does not make use of the profinite topology.

Lemma 3.3.1. *Let Γ be an infinite finitely generated group and H a subgroup of finite index in Γ. If Γ 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 Γ, is of finite index in Γ and therefore $KM \leq \Gamma$ must be of finite index too. Now, M is a maximal subgroup of Γ contained in KM, so either $M = KM$ or $KM = \Gamma$. If the former holds, then M is of finite index in Γ, a contradiction. If the latter holds, then $H = K(M \cap H) \leq L$, another contradiction.

Since H is finitely generated, every proper subgroup is contained in a maximal subgroup (this can be shown without using Zorn's Lemma, see [48]) and, by the above, the maximal subgroup containing $H \cap M$ must be of infinite index in H. □

Recall that a *chief factor* of a group Γ is a minimal normal subgroup of a quotient group of Γ.

Lemma 3.3.2 ([33], 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 Δ are finite and all maximal subgroups of Δ have finite index.*

Theorem 3.3. *Let $\Gamma := G^{\times k}$ be the direct product of k copies of a torsion GGS group G. If H is a group commensurable with Γ then all maximal subgroups of H have finite index in H.*

Proof. Pick $K \leq_{\mathrm{f}} H$ and $J \leq_{\mathrm{f}} \Gamma$ with $K \cong J$. For $i = 1, \dots, k$, denoting by G_i the ith direct factor of Γ, we have that $J_i := J \cap G_i$ has finite index in G_i; so Lemma 3.2.2 yields some $n_i \in \mathbb{N}$ with $G_i^{(n_i)} \leq_{\mathrm{f}} J_i$. Thus $S := G_1^{(n_1)} \cdots G_k^{(n_k)} \leq_{\mathrm{f}} J$.

Now, G is residually finite and just infinite and so all of its chief factors are finite. Also, by Theorem 2.10, all of its maximal subgroups have finite index. Furthermore, we saw in Lemma 3.2.2 that $G^{(n)}$ is a subdirect product of p^n copies of G. Thus S satisfies the assumptions of Lemma 3.3.2 and all of its maximal subgroups have finite index. Lemma 3.3.1 then implies that all maximal subgroups of J must have finite index in J. The result now follows by applying Lemma 3.3.1 to $K \cong J$ and H. $\square$

3.4 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. Remember that H_v denotes the vertex section of a subgroup $H \leq \operatorname{Aut} T$ at vertex v; that is, $H_v = \varphi_v(\mathrm{St}_H(v)) = \varphi_{u_n} \circ \cdots \circ \varphi_{u_1}(\mathrm{St}_H(v))$ where $v = u_1 \dots u_n$ is considered as a string of symbols $u_i \in \{0, \dots, p-1\}$.

Theorem 3.4. *Let $\mathcal{X}$ be a family of subgroups of G satisfying*

(I) $1 \in \mathcal{X}$, $G \in \mathcal{X}$;

(II) *if $H \in \mathcal{X}$ then $L \in \mathcal{X}$ for all $L \leq G$ such that $H \leq_{\mathrm{f}} L$;*

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

Then all finitely generated subgroups of G are in $\mathcal{X}$.

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

Suppose for a contradiction that there are finitely generated subgroups of G which are not in $\mathcal{X}$. 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 $H \leq \mathrm{St}(1)$ then by (III) at least one of the first level vertex sections of H is not in $\mathcal{X}$ and the generating set of this vertex section has elements of length at most $\frac{1}{2}(D+1) < D$ by Lemma 3.2.6, contradicting the choice of H.

Therefore H is not contained in $\mathrm{St}(1)$. We will show that there exists some $v \in V_2$ such that the vertex section H_v is not in $\mathcal{X}$ and has a generating set consisting of elements of length less than D.

Since $\mathrm{St}_H(1)$ is a finitely generated subgroup of $\mathrm{St}(1)$, (III) implies that not all first level vertex sections of $\mathrm{St}_H(1)$ are in $\mathcal{X}$. However, if one of them is in $\mathcal{X}$ then, as all first level vertex sections of H are conjugate in G, they must all be in $\mathcal{X}$. Thus no first level vertex section of H is in $\mathcal{X}$; in particular, none of them is equal to G, and Lemma 3.2.4 asserts that they are all contained in $\mathrm{St}(1)$. For each $k \in \{0,1,2\}$, property (III) again implies that one of H_{k0}, H_{k1}, H_{k2} is not in $\mathcal{X}$. We claim that for some k every such vertex section is generated by elements of length less than D.

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

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

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 (ii) in Lemma 3.2.6. Our claim follows if $D > 3$. For $D = 3$ and $D = 2$ the claim follows from Lemma 3.4.3 and Lemma 3.4.2, respectively, while Lemma 3.4.1 shows that if $D = 1$ then $H \in \mathcal{X}$. $\square$

Lemma 3.4.1. *Let H be a subgroup of G with a finite generating set S consisting of elements of length at most* 1*. Then H is in the family $\mathcal{X}$ from Theorem 3.4.*

Proof. Any non-trivial $s \in S$ must be of the form a^k, b_i^r, $b_i^r a^k$ 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 we may suppose that S contains at least two elements and that they are not all 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_i^r. Then either $H = \langle b_0, b_1, b_2 \rangle = \mathrm{St}(1)$, so $H \in \mathcal{X}$ by (I), (III) and the fact that $\mathrm{St}(1) \leq_{\mathrm{s}} G^{\times 3}$; or $H = \langle b_i, b_j \rangle < \mathrm{St}(1)$ for $i \neq j \in \{0,1,2\}$, so two of the first level vertex sections of H are G and the other one is contained in $\langle a \rangle$; 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 $(b_i^r a)^{-1} = b_{i+1}^{-r} a^2$). If there is some $b_j^q \in S$ such that $j = i$ then $b_j^{p-r} b_i^r a = a \in H$ so $H = G \in \mathcal{X}$. If there is some $b_j^q \in S$ with $j \neq i$, then $b_j^q = (a^q, a^{-q}, b^q)^{a^j}$ has $a^{\pm q}$ in the ith coordinate; but $(b_i^r a)^3 = b_i^r b_{i-1}^r b_{i+1}^r = (b^r, b_1^r, b^r)^{a^i}$ has b^r in the ith coordinate, so $H_i = G \in \mathcal{X}$. Since $H \not\leq \mathrm{St}(1)$, 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 $b_i^r a$, $b_j^q a$, with $r \neq q$ 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

$$b_i^r a(b_j^q a)^{-1} = (a^r, a^{-1}, b^r)^{a^i}(a^{-q}, a^q, b^{-q})^{a^j}$$

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

Lemma 3.4.2. *Let $H \not\leq \mathrm{St}(1)$ be a subgroup of G with a finite generating set S consisting of elements of length at most 2 and such that $H_u \leq \mathrm{St}(1)$ for all $u \in V_1$. Then H_v is generated by elements of length at most 1 for all $v \in V_2$.*

Proof. First note that no generator of H is in $\mathrm{St}(1)$ since the only possibilities are elements of the form b_i^r and $b_{i_1}^{r_1} b_{i_2}^{r_2}$ whose first level vertex sections are not in $\mathrm{St}(1)$.

If some $t \in S \setminus \mathrm{St}(1)$ has length less than 2 then, as in the proof of Theorem 3.4, the set $T = \{1, t, t^{-1}\}$ is a transversal of $\mathrm{St}_H(1)$ in H and $\mathrm{St}_H(1)$ is generated by the set $X = \{t_1 s t_2^{-1} \mid t_i \in T, s \in S, t_1 s t_2^{-1} \in \mathrm{St}_H(1)\}$. The elements of X have length at most 4, so, since $H_u \leq \mathrm{St}(1)$ for all $u \in V_1$, we may apply (ii) from Lemma 3.2.6 to see that all second level vertex sections of H will be generated by elements of length at most 1.

Suppose then that all elements of $S \setminus \mathrm{St}(1)$ have length 2, that is, they are all of the form $b_{i_1}^{r_1} b_{i_2}^{r_2} a$. It suffices to consider only elements of this form as $(b_{i_1}^{r_1} b_{i_2}^{r_2} a)^{-1} = b_{i_2+1}^{-r_2} b_{i_1+1}^{-r_1} a^2$. Pick some $t = b_{i_1}^{r_1} b_{i_2}^{r_2} a \in S$ to form the transversal T so that $\mathrm{St}_H(1)$ is generated by X as above. Then every element of X is of the form $t^{-1}s$, st^{-1}, t^3 or tst where $s = b_{j_1}^{q_1} b_{j_2}^{q_2} a \in S$. 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 $t^{-1}s = b_{i_2+1}^{-r_2} b_{i_1+1}^{-r_1} b_{j_1+1}^{q_1} b_{j_2+1}^{q_2}$ and $st^{-1} = b_{j_1}^{q_1} b_{j_2}^{q_2} b_{i_2}^{-r_2} b_{i_1}^{-r_1}$ yield elements of length at most 4 which are in $\mathrm{St}(2)$ by assumption. Thus, by (ii) in Lemma 3.2.6, their second level vertex sections have length at most 1.

Since $i_1 \neq i_2$, an element of the form $t^3 = b_{i_1}^{r_1} b_{i_2}^{r_2} b_{i_1-1}^{r_1} b_{i_2-1}^{r_2} b_{i_1+1}^{r_1} b_{i_2+1}^{r_2}$ 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 $tst = b_{i_1}^{r_1} b_{i_2}^{r_2} b_{j_1-1}^{q_1} b_{j_2-1}^{q_2} b_{i_1+1}^{r_1} b_{i_2+1}^{r_2}$. 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

$$\begin{aligned} tst &= b_0^{r_1} b_2^{r_2} b_{j_1-1}^{q_1} b_{j_2-1}^{q_2} b_1^{r_1} b_0^{r_2} \\ &= \begin{cases} b_0^{r_1} b_2^{r_2} b_0^{q_1} b_{j_2-1}^{q_2} b_1^{r_1} b_0^{r_2} & \text{if } j_1 - 1 = 0 \\ b_0^{r_1} b_2^{r_2} b_{j_1-1}^{q_1} b_0^{q_2} b_1^{r_1} b_0^{r_2} & \text{if } j_2 - 1 = 0. \end{cases} \end{aligned}$$

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

$$\begin{cases} b^{r_1} a^{r_2} b^{q_1} a^{q_2} a^{-r_1} b^{r_2} & \text{if } j_1 - 1 = j_2 = 0; \\ b^{r_1} a^{r_2} b^{q_1} a^{-q_2} a^{-r_1} b^{r_2} & \text{if } j_1 - 1 = 0,\ j_2 = -1; \\ b^{r_1} a^{r_2} a^{q_1} b^{q_2} a^{-r_1} b^{r_2} & \text{if } j_2 - 1 = j_1 = 0; \\ b^{r_1} a^{r_2} a^{-q_1} b^{q_2} a^{-r_1} b^{r_2} & \text{if } j_2 - 1 = 0, j_1 = -1. \end{cases}$$

In the first and third cases we have $\varphi_2(t^{-1}s) = b^{-r_2} a^{r_1} a^{\pm q_1} a^{\mp q_2} = b^{-r_2}$ (as $H_2 \leq \mathrm{St}(1)$ by assumption); while, in the second and fourth, $\varphi_2(st^{-1}) = a^{\mp q_1} a^{\pm q_2} a^{-r_2} b^{-r_1} = b^{-r_1}$. 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. $\square$

Lemma 3.4.3. *Let $H \not\leq \mathrm{St}(1)$ be a subgroup of G with a finite generating set S consisting of elements of length at most 3 and such that $H_u \leq \mathrm{St}(1)$ for all $u \in V_1$. Then there exists $k \in V_1$ such that H_{kj} is generated by elements of length less than 3 for all $j \in \{0, 1, 2\}$.*

Proof. If there exists $t \in S \setminus \mathrm{St}(1)$ with $l(t) \leq 2$ then, by the same argument as in the proof of Lemma 3.4.2, the generating set X of $\mathrm{St}_H(1)$ 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 $S \setminus \mathrm{St}(1)$ have length 3, pick one of them, say t, to form the transversal T. Assume that t has the form $b_{i_1}^{r_1} b_{i_2}^{r_2} b_{i_3}^{r_3} a$ for $i_1, i_2, i_3 \in \{0, 1, 2\}$ with $i_2 \neq i_1$, $i_2 \neq i_3$ and $r_1, r_2, r_3 = \pm 1$. It suffices to consider elements of this form since $t^{-1} = b_{i_3}^{r_3} b_{i_2}^{-r_2} b_{i_1}^{-r_1} a^{-1}$. We may pick $k \in \{0, 1, 2\}$ such that neither $\varphi_k(b_{i_3+1})$ nor $\varphi_k(b_{i_1})$ 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, by (ii) in Lemma 3.2.6, the vertex sections H_{kj} for $j \in \{0, 1, 2\}$ are generated by elements of length at most $(8 + 3)/4 < 3$, as required. $\square$

3.5 The case $p = 3$: proof of main theorems

The length reduction arguments and results in the previous section yield a characterization of the finite subgroups of the Gupta–Sidki 3-group G.

Theorem 3.5. *Let H be a finitely generated subgroup of the Gupta–Sidki 3-group G. Then either H is finite or one of its vertex sections is equal to G.*

Proof. We start with 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 3.4.1, either H is finite or $H_u = G$ for every $u \in V_1$.

Assume that the assertion in the theorem holds whenever $D \leq n$ with $n \geq 1$. For $D = n + 1$ we consider the cases $H \leq \mathrm{St}(1)$ and $H \not\leq \mathrm{St}(1)$ separately. If $H \leq \mathrm{St}(1)$ 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 inductive hypothesis, each H_u is either finite or has a vertex section equal to G. If all H_u are finite, then so is H, as it is a subdirect product of its first level vertex sections.

If $H \not\leq \mathrm{St}(1)$ then $H_u \leq \mathrm{St}(1)$ for every $u \in V_1$ by Lemma 3.2.4. In the case $n = 1$, Lemma 3.4.2 shows that each second level vertex section H_v is generated by elements of length at most 1, and is therefore finite or equal to G. If $n = 2$, by Lemma 3.4.3, there exists $u \in V_1$ such that H_{ui} is generated by elements of length at most 2 for all $i \in \{0, 1, 2, \}$. Thus, for each $i \in \{0, 1, 2\}$, either H_{ui} is finite or it has a vertex section equal to G, by inductive hypothesis. If all H_{ui} are finite, so is H_u. Since $H \not\leq \mathrm{St}(1)$, all first level vertex sections of H are conjugate in G and are therefore finite, making H finite. Lastly, if $n > 2$, by the same argument as in the proof of Theorem 3.4, H_v is generated by elements of length at most $(3D + 3)/4 < D$ whenever $v \in V_2$. If some H_v is infinite, it has a vertex section equal to G by inductive hypothesis; otherwise, H is finite. The theorem now follows by induction. □

An identical characterization of finite subgroups of the Grigorchuk group $\Gamma = \langle a, b, c, d \rangle$ may be obtained using the same methods as in the above proof and the analysis carried out in [33, Theorem 3]. A proof of the non-trivial implication is given in Appendix A, where the necessary material from [33] is also included.

Let us now move on to the proof of the two main results. These can be easily generalized to the case $p > 3$, provided that an analogue of Theorem 3.4 holds.

Theorem 3.1. *Every infinite finitely generated subgroup of the Gupta–Sidki* 3*-group* G *is commensurable with* G *or* $G \times G$*.*

Notation. For simplicity, $\mathcal{G}$ will from now on mean 'G or $G \times G$'.

To prove this theorem it suffices to show that properties (I)—(III) in Theorem 3.4 hold for the family $\mathcal{C}$ of subgroups of G which are finite, or commensurable with $\mathcal{G}$. Clearly, (I) holds as the trivial subgroup and G are in this family. It is also easy to see that (II) is satisfied by $\mathcal{C}$: if $H \in \mathcal{C}$ is commensurable with $\mathcal{G}$ and $J \leq_f H$ is a subgroup isomorphic to a finite index subgroup of $\mathcal{G}$ 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 $\mathrm{St}(1)$ and all first level vertex sections of H are in $\mathcal{C}$ then $H \in \mathcal{C}$.

This will follow from the next lemma, which uses similar ideas to those in [33].

Lemma 3.5.1. *If* $H \leq_{\mathrm{s}} H_1 \times \cdots \times H_n$ *where each direct factor* H_i *is in* $\mathcal{C}$ *then* $H \in \mathcal{C}$*. In other words,* $\mathcal{C}$ *is closed for subdirect products.*

Proof. We proceed by induction on n, starting with $n = 2$ since there is nothing to prove for $n = 1$. Suppose then that $H \leq_{\mathrm{s}} H_1 \times H_2$ with $H_i \in \mathcal{C}$. If both factors H_i are finite then so is H and we are done. Assume that H_1 is commensurable with $\mathcal{G}$. Then there exists some subgroup of finite index in H_1 isomorphic to some L of finite index in $\mathcal{G}$. By Lemma 3.2.2, L contains some finite index subgroup J which is subdirect in a direct power of G. Denote by M the preimage of J in H under the projection map. Since M has finite index in H, so does its projection to H_2. Thus we may replace H by M, H_1 by J and H_2 by the projection of M in H_2 and obtain

$$H \leq_{\mathrm{s}} G_1 \times \cdots \times G_r \times H_2$$

for some finite r, which we take to be minimal, and where each G_i is a copy of G.

Write $A := G_1 \times \cdots \times G_r$. Then $AH = G_1 \times \cdots \times G_r \times H_2$ is commensurable with $\mathcal{G}$ by Proposition 3.2.1 and we claim that H has finite index in AH. Denote by K_i the kernel of the map from H to the product of all factors of A except G_i; that is, $K_i = H \cap (1 \times \cdots \times G_i \times \cdots \times 1) \triangleleft H$. Note that K_i is non-trivial since we took r to be minimal. Then $K_i \triangleleft G_i$ and, since G_i is just infinite, K_i has finite index in G_i. This way we obtain a finite index normal subgroup $K_1 \times \cdots \times K_r$ of A contained in H. Hence $|AH : H| = |A : A \cap H|$ is finite and our claim is proved, showing that H is commensurable with $\mathcal{G}$.

Assume that the assertion in the lemma holds for subdirect products of at most n factors in $\mathcal{C}$ where $n \geq 1$. If $H \leq_{\mathrm{s}} H_1 \times \cdots \times H_{n+1}$, then $K := \mathrm{im}(H \to H_1 \times \cdots \times H_n)$ is subdirect in $H_1 \times \cdots \times H_n$ and is therefore in $\mathcal{C}$ by inductive hypothesis. Since $H \leq_{\mathrm{s}} K \times H_{n+1}$, the argument for the case $n = 2$ yields that $H \in \mathcal{C}$. The theorem then follows by induction.

□

Theorem 3.2. *The Gupta–Sidki* 3*-group is subgroup separable and hence has solvable membership problem.*

To prove this theorem it suffices to show that the conditions of Theorem 3.4 hold for the family $\mathcal{S}$ of finitely generated subgroups of G all of whose subgroups of finite index are closed with respect to the profinite topology on G.

Clearly, $\mathcal{S}$ satisfies (I). To see that it also satisfies (II), let $H \leq_{\mathrm{f}} L$ for some H in $\mathcal{S}$; thus L is finitely generated. For any $K \leq_{\mathrm{f}} L$, we have $K \cap H \leq_{\mathrm{f}} H$ 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 $\mathcal{S}$ also satisfies (III) we need the following lemma.

Lemma 3.5.2.

(i) *Suppose that H_0 is a group all of whose quotients 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 $H \leq_{\mathrm{s}} H_0 \times G_1 \times \cdots \times G_n$. Then every quotient of H is residually finite.*

(ii) *If H is abstractly commensurable with $\mathcal{G}$ then every quotient of H is residually finite.*

Proof. This is essentially Lemma 12 in [33]. The first part is identical and the proof of the second only requires small modifications. Suppose that $K \triangleleft H$; we want to show that K is an intersection of subgroups of finite index in H. Now, H is commensurable with $\mathcal{G}$ and all finite index subgroups of the latter contain some derived subgroup of $\mathcal{G}$, which is a subdirect product of finitely many copies of G by Lemma 3.2.2. So H has a normal finite index subgroup N 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 is equivalent to $N/(K \cap N)$ being residually finite. □

We may now proceed to show that $\mathcal{S}$ also satisfies (III).

Lemma 3.5.3. *Let H be a finitely generated subgroup of* $\mathrm{St}(1)$ *such that its first level vertex sections H_0, H_1, H_2 are in $\mathcal{S}$. Then H is in $\mathcal{S}$.*

Proof. Let $K \leq_{\mathrm{f}} H$. The first level vertex sections of K have finite index in those of H and are therefore closed in G. So we only need to show that H is closed in G. We will show that $\psi(H)$ is closed in $G \times G \times G$, so that H is closed in $\mathrm{St}(1) \leq_{\mathrm{f}} G$ and hence in G. In fact, if suffices to show that $\psi(H)$ is closed in $H_0 \times H_1 \times H_2$ because in that case $\psi(H) = \bigcap K_i$ for $K_i \leq_{\mathrm{f}} H_0 \times H_1 \times H_2$ with each K_i closed in $G \times G \times G$ by our assumption on $\mathcal{S}$. The lemma will follow from the more general

Claim. For every $n \in \mathbb{N}$, if $H \leq_{\mathrm{s}} H_0 \times \cdots \times H_n$ with $H_i \in \mathcal{S}$ then H is closed in $H_0 \times \cdots \times H_n$.

Proof of claim. We proceed by induction on n. When $n = 1$, define $L_0 \times 1 := H \cap (H_0 \times 1)$. Then L_0 is normal in H_0, which is either finite or commensurable with $\mathcal{G}$ by Theorem 3.1, so H_0/L_0 is residually finite (either trivially, or by Lemma 3.5.2). Thus there is a collection $(N_j)_{j\in J}$ of normal subgroups of finite index in H_0 whose intersection is L_0. Each subgroup $(N_j \times 1)H$ is of finite index in $H_0 \times H_1$ so it is enough to show that $H = \bigcap_{j\in J}((N_j \times 1)H)$. Let u be an element of the intersection, say $u = (n_j, 1)(h_{j,0}, h_{j,1})$ for each $j \in J$. Then $h_{j,1}$ is constant so we can write $u = (n_j h_{j,0}, h_1)$ for each $j \in J$. Fix $i \in J$ and define $h := (h_{i,0}, h_1) \in H$. Then, for each $j \in J$ we have

$$uh^{-1} = (n_j, 1)(h_{j,0}h_{i,0}^{-1}, 1) \in (N_j \times 1)(H \cap H_0 \times 1) = N_j \times 1.$$

Thus $uh^{-1} \in \bigcap_{j\in J} N_j \times 1 = L_0 \times 1 \leq H$, so $u \in H$.

Now assume that H is closed in $H_0 \times \cdots \times H_{n-1}$ whenever $H \leq_{\mathrm{s}} H_0 \times \cdots \times H_{n-1}$ and each H_i is in $\mathcal{S}$. Suppose that $H \leq_{\mathrm{s}} H_0 \times \cdots \times H_n$ with $H_i \in \mathcal{S}$. Define $L_0 \times 1 \times \cdots \times 1 :=$ $H \cap (H_0 \times 1 \times \cdots \times 1)$. As in the base step, there exist normal subgroups $(N_j)_{j \in J}$ of finite index in H_0 whose intersection is L_0. Let $H^j := (N_j \times 1 \times \cdots \times 1)H$ for each $j \in J$. We claim that $H = \bigcap_{j \in J} H^j$ and that each H^j is closed in $H_0 \times \cdots \times H_n$. The first statement is proved similarly to the two-factor case: write an element u of the intersection as $u = (n_j h_{j,0}, h_1, \ldots, h_n)$ and put $h := (h_{i,0}, h_1, \ldots, h_n) \in H$ so that $uh^{-1} = (n_j, 1, \ldots, 1)(h_{j,0} h_{i,0}^{-1}, 1, \ldots, 1) \in N_j \times 1 \times \cdots \times 1$. For the second statement, fix $j \in J$ and let P be the pre-image of N_j in H^j under the canonical projection to H_0. Then P has finite index in H^j as N_j has finite index in H_0. Thus, for each $i \neq 0$, the projection P_i of P onto the ith factor has finite index in H_i, so $P_i \in \mathcal{S}$. Furthermore, $N_j \times 1 \times \cdots \times 1 \leq P \leq_{\mathrm{s}} N_j \times P_1 \times \cdots \times P_n$, so

$$P/(N_j \times 1 \times \cdots \times 1) \leq_{\mathrm{s}} P_1 \times \cdots \times P_n.$$

By the inductive hypothesis, $P/(N_j \times 1 \times \cdots \times 1)$ is closed in $P_1 \times \cdots \times P_n$. Thus P is closed in $N_j \times P_1 \times \cdots \times P_n \leq_{\mathrm{f}} H_0 \times H_1 \times \cdots \times H_n$, hence also in $H_0 \times H_1 \times \cdots \times H_n$. $\square$

Part II

The structure of (weakly) branch groups

Chapter 4

An invariant of subgroups of finite index

This chapter contains the main theorem on the structure of (weakly) branch groups, Theorem 4.2. Speaking loosely, it establishes a connection between subgroups with finitely many conjugates and orbits on layers of the tree. The main results in subsequent chapters make use of this theorem at some point.

Notation: Throughout the rest of the thesis, T will denote a level-homogeneous tree. It will be convenient to use a partial order on the vertices of a tree on which G acts: write $v \leq u$ and call v a *descendant* of u if u is on the unique path from v to the root. Two vertices will be said to be *incomparable* if there is no path (without backtracking) from the root going through both of them.

The starting point is the following result by Grigorchuk, which he stated for branch groups but whose proof is valid for weakly branch groups:

Theorem 4.1 ([28, Theorem 4])**.** *Suppose that G is a weakly branch group acting on a tree T and let $K \triangleleft G$ with $K \neq 1$. Then K contains the derived subgroup $\mathrm{rist}_G(n)'$ of $\mathrm{rist}_G(n)$ for some integer n.*

Because in a branch action each $\mathrm{rist}_G(n)$ has finite index, the above theorem yields that all branch groups are just non-(virtually abelian).

The main result of this chapter is the more general:

Theorem 4.2. *Suppose that G is a weakly branch group acting on a tree T. Let H be a subgroup of finite index and let $K \triangleleft H$. Then for all sufficiently large integers n there* *is a union X of H-orbits in V_n such that*

$$K \cap \mathrm{rist}_G(n)' = \mathrm{rist}_G(X)'. \tag{4.1}$$

More precisely,

$$\mathrm{rist}_G(X)' \leq K \quad \text{and} \quad K \cap \mathrm{rist}_G(V_n \setminus X) = [K, \mathrm{rist}_G(V_n \setminus X)] = 1. \tag{4.2}$$

The second part of (4.2) above is simply the statement that the subgroups K and $\mathrm{rist}_G(V_n \setminus X)$ generate their direct product in G.

4.1 Proof of Theorem 4.2

We begin by noting an immediate consequence of the definition of weakly branch groups.

Remark 2. Let G be a weakly branch group on a tree T with layers V_n. Since G acts transitively on each V_n, a subgroup $H \leq G$ of finite index will have at most $|G : H|$ orbits on V_n. If X is an H-orbit in V_n then the descendants of X in V_{n+1} comprise a union of H-orbits. Hence there is some n_0 such that for all $n \geq n_0$ the number of H-orbits on V_n is equal to the number of H-orbits on V_{n_0}.

Lemma 4.1.1. *Let G be a weakly branch group on a tree T. Then the following hold:*

(i) *For all $u \in T$ the subgroup $\mathrm{rist}_G(u)$ is not virtually abelian.*

(ii) *G has no non-trivial abelian subgroup K such that $K \lhd K^G \lhd G$.*

(iii) *$\mathrm{rist}_G(v)$ is not virtually soluble for each $v \in T$.*

Proof. (i) Suppose for a contradiction that there exist $u \in T$ and $A \trianglelefteq_{\mathrm{f}} \mathrm{rist}_G(u)$ with A abelian. As A is non-trivial, there exist $a \in A$ and $v \leq u$ such that $v^a \neq v$. Now, $\mathrm{rist}_G(v) \cap A \trianglelefteq_{\mathrm{f}} \mathrm{rist}_G(v)$ implies that $\mathrm{rist}_G(v) \cap A \neq 1$. Since A is abelian, we have $\mathrm{rist}_G(v) \cap A = (\mathrm{rist}_G(v) \cap A)^a = \mathrm{rist}_G(v^a) \cap A$ which implies that $\mathrm{rist}_G(v) \cap \mathrm{rist}_G(v^a) \neq 1$ and therefore $v = v^a$, a contradiction.

(ii) We first show that G has no non-trivial normal abelian subgroups. Assume for a contradiction that $A \lhd G$ is non-trivial and abelian. Then, by Theorem 4.1, there exists n such that $\mathrm{rist}_G(n)' \leq A$. Since A is non-trivial there exists a vertex v which we may take to be of level at least n and $a \in A$ such that $v^a \neq v$. Then $\mathrm{rist}_G(v)', \mathrm{rist}_G(v^a)' \leq \mathrm{rist}_G(n)' \leq A$ implies that $(\mathrm{rist}_G(v)')^a = \mathrm{rist}_G(v^a)' = \mathrm{rist}_G(v)'$ and so $v = v^a$ by (i), a contradiction.

The proof of the general case is similar to that of the corresponding fact when G is a branch group, shown in [34]. Suppose there exists a non-trivial, abelian $K \leq G$ such that $K \lhd K^G \lhd G$. Then there exists a non-trivial $k \in K$ and $u \in T$ such that $u^k \neq u$. Assume that u is of level n. We claim that $\mathrm{rist}_G(u) \cap \mathrm{N_G}(K)$ is an elementary abelian 2-group. Let $h \in \mathrm{rist}_G(u) \cap \mathrm{N_G}(K)$ so that k^{-1} and k^h commute. Then for any $v \in T_u$ we have

$$v^{k^{-1}k^h} = v^{k^{-1}h^{-1}kh} = v^{k^{-1}kh} = v^h$$

and

$$v^{k^hk^{-1}} = v^{h^{-1}khk^{-1}} = v^{h^{-1}kk^{-1}} = v^{h^{-1}},$$

so $h = h^{-1}$. Hence $\mathrm{rist}_G(u) \cap \mathrm{N_G}(K)$ is an elementary abelian 2-group and, since $K^G \leq \mathrm{N_G}(K)$, the subgroup $\mathrm{rist}_G(u) \cap K^G$ is abelian. Thus $\prod_{g \in G}(\mathrm{rist}_G(u^g) \cap K^G) = \mathrm{rist}_G(n) \cap K^G \lhd G$ is abelian and therefore trivial by the previous paragraph. But, since $K^G \lhd G$ there exists m such that $\mathrm{rist}_G(m)' \leq K^G$, which means that $\mathrm{rist}_G(n) \cap K^G \geq \mathrm{rist}_G(m)' \cap \mathrm{rist}_G(n)' \neq 1$ by (i). Thus, we have obtained a contradiction and K must be trivial.

(iii) If $\mathrm{rist}_G(v)$ is virtually soluble, then it must have a non-trivial abelian normal subgroup K. Since then $K \lhd K^G \lhd G$ this gives a contradiction to (ii).

□

From this we can prove something stronger.

Proposition 4.1.2. *Let G be a weakly branch group acting on a tree T. If $H \leq_{\mathrm{f}} G$ then H has no non-trivial virtually soluble normal subgroups.*

Proof. First we claim that if K is a finite normal subgroup of H then $K = 1$. Let $C = \mathrm{C}_H(K)$; then, since $H/C = \mathrm{N}_H(K)/C$ embeds in $\mathrm{Aut}(K)$, we have $C \leq_{\mathrm{f}} G$ and so $D \trianglelefteq_{\mathrm{f}} G$ where $D = \bigcap_{g \in G} C^g$. Moreover D centralizes K^G, thus $D \cap K^G$ is an abelian normal subgroup of G. From above $D \cap K^G = 1$ and so K^G is finite. If $K \neq 1$ then by Theorem 4.1 we have $\mathrm{rist}_G(n_1)' \leq K^G$ for some n_1. Because G is residually finite, there is some $L \trianglelefteq_{\mathrm{f}} G$ with $L \cap K^G = 1$, and we have $\mathrm{rist}_G(n_2)' \leq L$ for some n_2. But then $\mathrm{rist}_G(n)$ must be abelian where $n = \max(n_1, n_2)$, and this is a contradiction to Lemma 4.1.1 (ii).

Finally suppose that $1 \neq K \triangleleft H$ (in particular, K is infinite by the above) and that K is virtually soluble. The soluble normal subgroup K_0 of K of smallest index is normal in H and infinite, and the last non-trivial term A of the derived series of K_0 is an abelian normal subgroup of H. By the above, A is infinite and so $A \cap M$ is a non-trivial abelian normal subgroup of M where $M = \bigcap_{g \in G} H^g$. But this gives another contradiction to Lemma 4.1.1 (ii). $\square$

We are now ready to prove the main theorem.

Theorem 4.2. *Suppose that G is a weakly branch group acting on a tree T. Let H be a subgroup of finite index and let $K \triangleleft H$. Then for all sufficiently large integers n there is a union X of H-orbits in V_n such that*

$$K \cap \mathrm{rist}_G(n)' = \mathrm{rist}_G(X)'. \tag{4.1}$$

More precisely,

$$\mathrm{rist}_G(X)' \leq K \quad \text{and} \quad K \cap \mathrm{rist}_G(V_n \setminus X) = [K, \mathrm{rist}_G(V_n \setminus X)] = 1. \tag{4.2}$$

Proof. Firstly, we claim that for every n there is some m such that $\mathrm{rist}_G(m)' \leq \mathrm{rist}_H(n)'$: Writing C for the normal core of H in G, we have $\mathrm{rist}_C(n) \triangleleft G$, so $\mathrm{rist}_C(n)' \triangleleft G$. Thus, by Theorem 4.1, there is some m such that $\mathrm{rist}_G(m)' \leq \mathrm{rist}_C(n)' \leq \mathrm{rist}_H(n)'$; in particular, $\mathrm{rist}_G(m)' = \mathrm{rist}_H(m)'$.

Choose n_0 as in Remark 2. By the above, there is some $m_0 \geq n_0$ with $\mathrm{rist}_G(m_0)' \leq \mathrm{rist}_H(n_0)' \leq H$. Let $X_1, \ldots, X_r$ be the orbits of H on V_{m_0}.

For $i = 1, \ldots, r$ we have two possibilities: either (a) there is some descendant of a vertex in X_i which is moved by K or (b) X_i and all its descendants are pointwise fixed by K. If (a) holds, pick such a descendant u (at level n_i, say) of one of the vertices in X_i and some $k \in K$ such that $u^k \neq u$. Let $x, y \in \mathrm{rist}_H(u)$, so $y^k \in \mathrm{rist}_H(u^k)$ and $[x, y^k] = 1$. Hence $[x, y] = [x, (y^{-1})^k y] = [x, [k, y]] \in K$ and $\mathrm{rist}_H(u)' \leq K$. Because H normalizes K, it follows that K contains $\mathrm{rist}_H(Y_i)'$ where Y_i is the H-orbit containing u, consisting of descendants of X_i. If (b) holds, then $K \cap \mathrm{rist}_G(u) = 1$ for every vertex u in X_i.

Let n be the maximum of the n_i from case (a) and X the union of all descendants in V_n of the Y_i. Then

$$\begin{aligned} K \cap \mathrm{rist}_G(n)' = K \cap \mathrm{rist}_H(n)' &= K \cap (\mathrm{rist}_H(X)' \times \mathrm{rist}_H(V_n \setminus X)') \\ &= \mathrm{rist}_H(X)' \times (K \cap \mathrm{rist}_H(V_n \setminus X)') = \mathrm{rist}_H(X)'. \quad \square \end{aligned}$$

As a first application of the theorem, we have the following result about the subgroup structure of *branch* groups. Here we crucially use the assumption that each $\mathrm{rist}_G(n)$ has finite index in G; equivalently, that all proper quotients of a branch group are virtually abelian.

Corollary 4.1.3. *Let G be a branch group acting on a tree T, and $H \leq_{\mathrm{f}} G$. Then in any series of normal subgroups of H there are at most $|G : H|$ factors that are not virtually abelian.*

Proof. For an H-invariant subset X of a layer V_n of T we write $o(X)$ for the number of orbits of H on X.

Let $(K_i)_{i\in\mathbb{N}}$ be an ascending series of normal subgroups of H. Then, by Theorem 4.2, for each i there exist some n_i and some union X_i of H-orbits in V_{n_i} such that $\mathrm{rist}_G(n_i)' \leq H$ and $\mathrm{rist}_G(n_i)' \cap K_i = \mathrm{rist}_G(X_i)'$. Moreover n_i can be chosen so that $o(X_i)$ is as large as possible (by Remark 2). Clearly $o(X_{i-1}) \leq o(X_i)$ for each i. Thus there are at most $|G : H|$ indices i for which $o(X_{i-1}) < o(X_i)$. Suppose that $j \in \mathbb{N}$ is not one of these indices. We claim that the quotient K_j/K_{j-1} is virtually abelian. To see this, let $n \geq \max(n_{j-1}, n_j)$, replace X_{j-1} and X_j by their respective sets of descendants in V_n, put $R := \mathrm{rist}_G(n)'$ and notice that

$$\mathrm{rist}_G(X_{j-1})' = R \cap K_{j-1} \leq R \cap K_j = \mathrm{rist}_G(X_j)'.$$

Since $o(X_{j-1}) = o(X_j)$, equality holds above. Therefore the kernel of the obvious homomorphism from K_j to H/RK_{j-1} equals $K_j \cap (RK_{j-1}) = K_{j-1}$. By definition of a branch group, the quotient H/R is virtually abelian, and so therefore are H/RK_{j-1} and K_j/K_{j-1}. Our claim follows, and the result is proved. □

4.2 Abstract commensurability of weakly branch groups

The idea in this section is to use Theorem 4.2 to show that the maximum number of orbits of a finite index subgroup of a weakly branch group is an invariant of the subgroup. We thus reduce the problem of deciding whether groups are isomorphic to counting orbits on layers of trees.

Definition 4.2.1. For an infinite group H, denote by $b(H)$ be the largest number r such that H has r infinite normal subgroups that generate their direct product in H; if no such r exists write $b(H) = \infty$.

If H is a subgroup of finite index in a weakly branch group (or if H is any infinite group having no non-trivial finite normal subgroups) then $b(H)$ is the maximal number of direct factors in an irredundant subdirect product decomposition of H.

Corollary 4.2.2. *Let G be a weakly branch group on a tree T and let $H \leq_{\mathrm{f}} G$. Then $b(H)$ is finite and is the maximum number of H-orbits on any layer of T.*

Proof. By Theorem 4.1 and Remark 2, there exist n_0 and some $m \leq |G : H|$ such that $\mathrm{rist}_G(n_0)' \leq H$ and such that H has m orbits on V_n for all $n \geq n_0$. If $X_1, \ldots, X_m$ are the orbits of H on V_{n_0} then the m subgroups $\mathrm{rist}_G(X_i)'$ are infinite normal subgroups of H and generate their direct product.

Now let $K_1, \ldots, K_r$ be infinite normal subgroups of H that generate their direct product. By Theorem 4.2, for each i we can find an integer $n_i \geq n_0$ and a union X_i of H-orbits on V_{n_i} such that $\mathrm{rist}_G(X_i)' \leq K_i$. Let $n = \max\{n_1, \ldots, n_r\}$ and for each i choose an H-orbit Y_i of V_n consisting of descendants of elements of X_i. Obviously the sets Y_i are disjoint and so $r \leq m$. □

Lemma 4.2.3. *Let T be the p-regular tree for some prime p and let H be a subgroup of $\mathrm{Aut}(T)$ that acts on each layer as a p-group. Then the number of H-orbits on each layer is congruent to 1 modulo $p-1$. In particular, if H is also a subgroup of finite index in a weakly branch group on T, then $b(H) \equiv 1 \mod p-1$.*

Proof. This is elementary. Pick some layer V_n. Then the size of each H-orbit in V_n is a power of p. Writing s_i for the number of orbits of size p^i for $i = 0, \ldots, n$ we have $p^n = s_0 + ps_1 + \cdots + p^n s_n$, so we obtain $1 \equiv s_0 + s_1 + \cdots + s_n \mod p-1$. □

Using these results we show that direct products of certain weakly branch groups on p-regular trees can only be (abstractly) commensurable if the numbers of direct factors are congruent modulo $p-1$.

Lemma 4.2.4. *Let $\mathfrak{C}$ be the class of all groups G with no non-trivial abelian normal subgroups and with $b(G)$ finite. Let $H_1, \ldots, H_n$ be groups in $\mathfrak{C}$, and H a subgroup of finite index in the direct product $D = H_1 \times \cdots \times H_n$ such that each projection $\rho_j : H \to H_j$ is surjective. Then $b(H) = b(H_1) + \cdots + b(H_n)$.*

Proof. Let $K_1, \ldots, K_r$ be infinite normal subgroups of H that generate their direct product in H. For $i \in \{1, \ldots, r\}$ and $j \in \{1, \ldots, n\}$, write $K_{i,j} = \rho_j(K_i) \lhd H_j$. If $i \in \{1, \ldots, r\}$ and $K_{i,j}$ is finite for all j then K_i is finite, a contradiction. So for each i there is some j such that $K_{i,j}$ is infinite. Therefore $r \leq \sum r_j$, where for each j, we write r_j for the number of indices i with $K_{i,j}$ infinite. Now each subgroup K_i centralizes all other subgroups $K_{i'}$ with $i' \neq i$. Therefore for fixed j, each $K_{i,j}$ centralizes the product $P_{i,j}$ of all subgroups $K_{i',j}$ with $i' \neq i$; and since H_j has no non-trivial abelian normal subgroups, the normal subgroups $K_{i,j}$, $P_{i,j}$ intersect trivially. Thus the subgroups $K_{i,j}$ with $1 \leq i \leq r$ generate their direct product in H_j. It follows that $r_j \leq b(H_j)$, and hence $r \leq b(H_1) + \cdots + b(H_n)$. In particular, $b(H)$ is finite and bounded by $\sum b(H_j)$.

It remains to prove that $\sum b(H_j) \leq b(H)$. For each j, find a family of $b(H_j)$ infinite normal subgroups of H_j that generate their direct product in H_j. If $L \lhd H_j$ is one of these subgroups, then its image $\bar{L}$ under the natural injection $H_j \to D$ is an infinite normal subgroup of D; hence $\bar{L} \cap H$ is an infinite normal subgroup of H. In this way we find $\sum b(H_j)$ infinite normal subgroups of H that evidently generate their direct product. The result follows. $\square$

Corollary 4.2.5. *Let $\mathfrak{D}$ be the family of weakly branch groups which act as a p-group on each layer of the p-regular tree T (in particular, $\mathfrak{D}$ is contained in $\mathfrak{C}$). Let Γ_1 and Γ_2 be, respectively, direct products of n_1 and n_2 groups in $\mathfrak{D}$. If Γ_1 and Γ_2 are abstractly commensurable then $n_1 \equiv n_2 \mod p-1$.*

Proof. Suppose that $H_1 \leq_{\mathrm{f}} \Gamma_1$ and $H_2 \leq_{\mathrm{f}} \Gamma_2$ are isomorphic. Then, for $i = 1, 2$, each projection of H_i onto a direct factor of Γ_i is in $\mathfrak{C}$ by Corollary 4.2.2. Thus, Lemma 4.2.3 and Lemma 4.2.4 together imply that $b(H_i) \equiv n_i \mod p-1$, so we must have $n_1 \equiv n_2 \mod p-1$. $\square$

Applying the above corollary to the Gupta–Sidki 3-group we can now strengthen Theorem 3.1.

Theorem 4.3. *The Gupta–Sidki 3-group has precisely two abstract commensurability classes of finitely generated infinite subgroups.*

Chapter 5

The structure lattice of a branch group

In this chapter the structure lattice and structure graph of a branch group are defined. These very useful objects were first introduced in [65] and [66] and used to analyse just infinite groups. They were also used in [41] to characterize branch groups in purely group-theoretic terms. In those settings, they are defined as quotients of the lattice of subnormal subgroups with finitely many conjugates in a branch group. The definition given here is more direct, as it does not require the subnormality assumption. Theorem 4.2 will also be used here to show that the structure lattice corresponds exactly to unions of orbits of finite index subgroups of the branch group (Proposition 5.1.7) and that the structure graph corresponds to blocks of imprimitivity in the action of the branch group on each layer (Proposition 5.2.3). The structure graph will then be used in Chapter 6 to define an intrinsic congruence topology of a branch group (all relevant terminology will be explained there).

Throughout, G will denote a branch group acting on a tree T.

5.1 The structure lattice of a branch group

Let $L(G)$ be the collection of all subgroups of G which have finitely many conjugates (in other words, whose normalizer has finite index). If $H, K \in L(G)$, then clearly $H \cap K, \langle H, K\rangle \in L(G)$. Thus $L(G)$ forms a lattice with respect to subgroup inclusion, with $H \cap K$ and $\langle H, K\rangle$ respectively the meet and join of two elements H, K. By Proposition 4.1.2, $L(G)$ contains no non-trivial virtually soluble subgroups. We will use this without further comment in the remainder.

This allows us to prove the following, which is a generalization of [41, Theorem 8.3.1] and [66, Lemma 4.3].

Proposition 5.1.1. *Let $H, K \in L(G)$ with $K \trianglelefteq H$ and H/K virtually nilpotent. Then $\mathrm{C_G}(K) = \mathrm{C_G}(H)$.*

Proof. First we claim that if A is a subgroup with finitely many conjugates in a group Γ and A is virtually nilpotent then so is A^Γ. Let N be the normal core of $\mathrm{N}_\Gamma(A)$ in Γ, so that $A_0 := A \cap N$ is virtually nilpotent and normal in N. By Fitting's theorem ([54, 5.2.8]), A_0 has a unique maximal nilpotent normal subgroup, B say, which is normal in N. The finitely many Γ-conjugates of B are also nilpotent and normal in N; thus B^Γ is nilpotent, again by Fitting's theorem. It remains to show that B^Γ has finite index in A^Γ. Since A/B is finite so is the quotient $(AB^\Gamma)/B^\Gamma$ and this has finitely many conjugates in Γ/B^Γ because A has finitely many in Γ. Therefore the quotient $(A^\Gamma B^\Gamma)/B^\Gamma \cong A^\Gamma/B^\Gamma$ is finite by Dicman's lemma (see [54, 14.5.7]) and our claim is proved.

Suppose that $H, K \in L(G)$ with $K \trianglelefteq H$ and H/K virtually nilpotent, and put $C := \mathrm{C_G}(K) \in L(G)$. We will prove the proposition for $H_1 := H^C$, K and deduce the result for H, K from this. To see that H_1/K is virtually nilpotent, note that for each $c \in C$ the conjugate $(H/K)^c = H^c/K$ is isomorphic to H/K. There are finitely many of these C-conjugates, as $H \in L(G)$, so it follows from the above claim that H_1/K is virtually nilpotent. Now, $C \cap K \in L(G)$ is abelian, hence trivial, and we have

$$C \cap H_1 = (C \cap H_1)/(C \cap K) \cong K(C \cap H_1)/K \leq H_1/K.$$

Thus $C \cap H_1 \in L(G)$ is virtually nilpotent and therefore trivial. Note that C is normalized by H_1, since $K \trianglelefteq H_1$, whence $[C, H_1] \leq C$. As H_1 is also normalized by C we have $[C, H_1] \leq C \cap H_1 = 1$ and therefore $C \leq \mathrm{C_G}(H_1)$. The proof is complete as $K \leq H \leq H_1$ implies that $\mathrm{C_G}(H_1) \leq \mathrm{C_G}(H) \leq C$. $\square$

Notation. For $H, K \in L(G)$, we write $K \leq_{\mathrm{va}} H$ (respectively, $K \trianglelefteq_{\mathrm{va}} H$) if $K \leq H$ (resp. $K \trianglelefteq H$) and K contains the derived group of a finite index subgroup of H. Note that if $K \trianglelefteq_{\mathrm{va}} H$ then H/K is virtually abelian.

Proposition 5.1.1 has the following consequences.

Lemma 5.1.2. *Let $H_1, H_2 \in L(G)$. Then $H_1 \cap H_2 = 1$ if and only if $[H_1, H_2] = 1$.*

Proof. If $[H_1, H_2] = 1$ then $H_1 \cap H_2 \in L(G)$ is abelian and therefore trivial.

For the converse, let N be the normal core of the intersection $\mathrm{N_G}(H_1) \cap \mathrm{N_G}(H_2)$. Thus $N \trianglelefteq_{\mathrm{f}} G$ normalizes H_1 and H_2. For $i = 1, 2$, let $K_i = H_i \cap N$. Then $K_i \trianglelefteq_{\mathrm{f}} H_i$ and $K_i \in L(G)$. Now, since $K_i \trianglelefteq K_1K_2$, we have $[K_1, K_2] \leq K_1 \cap K_2 \leq H_1 \cap H_2 = 1$. Therefore, applying Proposition 5.1.1 to $K_i \leq_{\mathrm{f}} H_i$, we obtain $\mathrm{C_G}(H_1) = \mathrm{C_G}(K_1) \leq \mathrm{C_G}(K_2) = \mathrm{C_G}(H_2)$, and vice-versa, so $[H_1, H_2] = 1$. □

Lemma 5.1.3. *For every $H \in L(G)$ we have $\langle H, \mathrm{C_G}(H)\rangle = H \times \mathrm{C_G}(H) \leq_{\mathrm{va}} G$.*

Proof. Write $C = \mathrm{C_G}(H)$ and note that $\langle H, C\rangle = H \times C$ by Lemma 5.1.2. If the normal core N of H is non-trivial then $N \trianglelefteq_{\mathrm{va}} G$, so $H \leq_{\mathrm{va}} G$ and hence $H \times C \leq_{\mathrm{va}} G$. Suppose then that $N = 1$ and let $V \in L(G)$ be the intersection of a maximal number of conjugates of H such that $1 < V \leq H$. If W is a conjugate of V which is not contained in H then $1 \leq W \cap H \leq H$ is the intersection of one more conjugate of H than V. Therefore $W \cap H = 1$, by the choice of V and $W \leq C$ by Lemma 5.1.2. This implies that $H \times C$ contains all conjugates of V; in particular, it contains their product $V^G \trianglelefteq_{\mathrm{va}} G$. Thus $H \times C \leq_{\mathrm{va}} G$, as required. □

Lemma 5.1.4. *Let $H, K \in L(G)$. The following are equivalent:*

(i) $H \cap K \leq_{\mathrm{va}} H, K$;

(ii) $\mathrm{C_G}(H) = \mathrm{C_G}(K)$;

(iii) *there exists $D \in L(G)$ such that $H \times D \leq_{\mathrm{va}} G$ and $K \times D \leq_{\mathrm{va}} G$.*

Proof. If (i) holds, we have $A' \leq H \cap K$ for some $A \leq_{\mathrm{f}} H$. Let N and L be, respectively, the normal cores of A and $H \cap K$ in H. Then $N \trianglelefteq_{\mathrm{f}} H$ and $N' \trianglelefteq L \trianglelefteq H$; that is $L \trianglelefteq_{\mathrm{va}} H$. Thus $\mathrm{C_G}(L) = \mathrm{C_G}(H)$ by Proposition 5.1.1, but then $L \leq K \leq H$ implies that $\mathrm{C_G}(H \cap K) = \mathrm{C_G}(L) = \mathrm{C_G}(H)$. Repeating the procedure with H replaced by K yields $\mathrm{C_G}(H \cap K) = \mathrm{C_G}(K) = \mathrm{C_G}(H)$. This immediately implies (iii) by Lemma 5.1.3.

Suppose that (iii) is true; so $G_0' \leq H \times D$ for some $G_0 \leq_{\mathrm{f}} G$. Then $G_0 \cap K \leq_{\mathrm{f}} K$ and, since $D \cap K = 1$, we have $(G_0 \cap K)' \leq (H \times D) \cap K = H \cap K$; that is, $H \cap K \leq_{\mathrm{va}} K$. The same argument with H and K swapped gives $H \cap K \leq_{\mathrm{va}} H$. □

For $H, K \in L(G)$, write $K \sim H$ if any of the equivalent conditions of Lemma 5.1.4 holds. It is immediate that $\sim$ is an equivalence relation on $L(G)$. We show that $\sim$ is a congruence on $L(G)$.

Proposition 5.1.5. *Let $H_1, H_2, K_1, K_2 \in L(G)$ with $H_1 \sim H_2$ and $K_1 \sim K_2$. Then $H_1 \cap K_1 \sim H_2 \cap K_2$ and $\langle H_1, K_1 \rangle \sim \langle H_2, K_2 \rangle$.*

Proof. For $i = 1, 2$ we have $H_1 \cap H_2 \cap K_i \leq_{\mathrm{va}} H_i \cap K_i$ and $K_1 \cap K_2 \cap H_i \leq_{\mathrm{va}} H_i \cap K_i$ so that $(H_1 \cap H_2 \cap K_i) \cap (K_1 \cap K_2 \cap H_i) \leq_{\mathrm{va}} H_i \cap K_i$. Thus $H_1 \cap K_1 \sim H_2 \cap K_2$.

To show that the join operation is respected, write $M := \langle H_1 \cap H_2, K_1 \cap K_2 \rangle \in L(G)$. Then $\mathrm{C_G}(M)$ centralizes H_i and K_i for $i = 1, 2$ so that Lemma 5.1.2 yields $\mathrm{C_G}(M) \cap H_i = 1$ and $\mathrm{C_G}(M) \cap K_i = 1$. Therefore $\langle \mathrm{C_G}(M), H_i, K_i \rangle = \mathrm{C_G}(M) \times \langle H_i, K_i \rangle \geq \mathrm{C_G}(M) \times M$ and, since $\mathrm{C_G}(M) \times M \leq_{\mathrm{va}} G$ (by Lemma 5.1.3), we have $\mathrm{C_G}(M) \times \langle H_i, K_i \rangle \leq_{\mathrm{va}} G$ for $i = 1, 2$. □

Definition 5.1.6. By the above, the join and meet of two equivalence classes $[H], [K]$ of elements $H, K \in L(G)$ is well defined by

$$[H] \vee [K] = [\langle H, K \rangle] \quad \text{and} \quad [H] \wedge [K] = [H \cap K],$$

and the quotient $\mathcal{L} = L(G)/\sim$ is again a lattice with respect to these operations and the natural partial order inherited from $L(G)$:

$$[K] \leq [H] \quad \text{if and only if} \quad [K \cap H] = [K].$$

This quotient $\mathcal{L}$ is the *structure lattice* of G.

Note that $[G]$ and $[1] = \{1\}$ are, respectively, the greatest and least elements of $\mathcal{L}$. Observe also that $[H]^g = [H^g]$ for each $H \in L(G)$ and $g \in G$. Thus the action of G on its subgroups by conjugation induces a well defined action on $\mathcal{L}$. It is shown in [41] that $\mathcal{L}$ is a Boolean lattice (it is uniquely complemented and distributive), but we do not require this fact here.

Theorem 4.2 shows another way of describing this lattice:

Proposition 5.1.7. *There is a one-to-one correspondence between $\mathcal{L}$ and*

$$\{[\mathrm{rist}_G(X)] \mid X \text{ is a union of } H\text{-orbits in some } V_n, \text{ for } H \leq_{\mathrm{f}} G\}.$$

In other words, $\mathcal{L}$ corresponds precisely to the unions of orbits of finite index subgroups of G.

In particular, every element of $\mathcal{L}$ contains some $[\mathrm{rist}_G(v)]$.

Proof. Consider the family $M(G)$ of subgroups of the form $\mathrm{rist}_G(X)$ where X is an H-invariant subset of a layer of T for some subgroup $H \leq_{\mathrm{f}} G$. Write $\mathrm{rist}_G(X_1) \approx \mathrm{rist}_G(X_2)$ for such subsets X_1, X_2 if one consists of all descendants of the other in that layer. Then $\approx$ is an equivalence relation on $M(G)$. By Theorem 4.2, each $K \in L(G)$ is $\sim$ equivalent to some $\mathrm{rist}_G(X) \in M(G)$ (and each $\mathrm{rist}_G(X) \in M(G)$ is in $L(G)$ by definition of $M(G)$).

Furthermore, we claim that $\mathrm{rist}_G(X_1) \approx \mathrm{rist}_G(X_2)$ if and only if $\mathrm{rist}_G(X_1) \sim \mathrm{rist}_G(X_2)$. If $\mathrm{rist}_G(X_1) \approx \mathrm{rist}_G(X_2)$ with $X_1 \subseteq V_n$, say, then $\mathrm{rist}_G(V_n \setminus X_1) \in L(G)$ and $\mathrm{rist}_G(X_1) \times \mathrm{rist}_G(V_n \setminus X_1) \leq_{\mathrm{va}} G$, $\mathrm{rist}_G(X_2) \times \mathrm{rist}_G(V_n \setminus X_1) \leq_{\mathrm{va}} G$, so $\mathrm{rist}_G(X_1) \sim \mathrm{rist}_G(X_2)$ by Lemma 5.1.4. Conversely, if $\mathrm{rist}_G(X_1) \sim \mathrm{rist}_G(X_2)$ then assume without loss of generality that X_2 is further from the root than X_1. If there exists some $x \in X_2$ which is not a descendant of any element of X_1 then $\mathrm{rist}_G(x)$ centralizes $\mathrm{rist}_G(X_1)$ while it does not centralize $\mathrm{rist}_G(X_2)$; so $\mathrm{rist}_G(X_1)$ and $\mathrm{rist}_G(X_2)$ have different centralizers, a contradiction by Lemma 5.1.4.

Thus $M(G)/\approx$ is isomorphic to $\mathcal{L}$. $\square$

5.2 The structure graph

An element B of $L(G)$ is *basal* if $\langle B^G \rangle$ is the direct product of the finitely many conjugates of B; in particular, if $B^g \neq B$ then $B^g \cap B = 1$. Examples of basal subgroups include $\mathrm{rist}_G(v)$ for every vertex v of a tree on which G acts as a branch group.

Basal subgroups have the following useful properties.

Lemma 5.2.1 ([66]). *Let B, B_1, B_2 be basal subgroups of G. Then*

(i) *$B_1 \cap B_2$ is basal;*

(ii) *if $1 \neq [B_1] \leq [B_2]$ then $\mathrm{N_G}(B_1) \leq \mathrm{N_G}(B_2)$;*

(iii) *$\mathrm{N_G}(B)$ is the stabilizer of $[B]$ under the action of G by conjugation;*

(iv) *$\bigcap\{\mathrm{N_G}(B) \mid B$ is basal$\} = 1$.*

Proof. (i) First observe that $(B_1 \cap B_2) \in L(G)$. Suppose that $(B_1 \cap B_2)^g = B_1^g \cap B_2^g \neq B_1 \cap B_2$ for some $g \in G$. Then either $B_1^g \neq B_1$ or $B_2^g \neq B_2$. In each case we have $(B_1 \cap B_2)^g \cap (B_1 \cap B_2) = 1$.

(ii) Let $g \in \mathrm{N_G}(B_1)$. Then $1 \neq [B_1]^g = [B_1] \leq [B_2]^g \wedge [B_2] = [B_2^g \cap B_2]$ so $B_2^g = B_2$ and $g \in \mathrm{N_G}(B_2)$, as required.

(iii) This follows by a similar argument to that in the previous part.

(iv) Note that $\mathrm{N_G}([\mathrm{rist}_G(v)]) = \mathrm{St}_G(v)$ for every $v \in T$, where T is a tree on which G acts as a branch group. The claim follows from the observation that every $\mathrm{rist}_G(v)$ is basal and $\bigcap\{\mathrm{St}_G(v) \mid v \in T\} = 1$. □

Definition 5.2.2. The *structure graph* $\mathcal{B}$ of G has as vertices all non-trivial elements $[B] \in \mathcal{L}$ such that B is basal. Two distinct elements $[A] < [B]$ of $\mathcal{B}$ are joined by an edge if $[A]$ is maximal subject to this inequality. Again, the conjugation action of G on its basal subgroups induces an action on $\mathcal{B}$ by graph automorphisms.

In fact, every basal subgroup is $\sim$ equivalent to the direct product of rigid stabilizers of vertices in a block of imprimitivity in the action of G on some layer of T:

Proposition 5.2.3. *Suppose that G has a branch action on T. Then for every basal subgroup $B \leq G$ there exists $v \in T$ such that $[B] = [\prod_{h \in H} \mathrm{rist}_G(v)^h]$ where $H = \mathrm{N_G}(B) \geq \mathrm{St}_G(v)$.*

This is Lemma 15.1.1 of [41] but the proof can be simplified as follows.

Proof. By Theorem 4.2, there exists $v \in T$ such that $[\mathrm{rist}_G(v)] \leq [B]$ and therefore $[B] \geq [\mathrm{rist}_G(v)]^H$. Now, for any u in the same layer as v such that $[\mathrm{rist}_G(u)] \not\leq [\mathrm{rist}_G(v)]^H$ there exists $g \in G \setminus H$ such that $u = v^g$. Then $[\mathrm{rist}_G(u)] \leq [B]^g$ and $g \notin H$ implies that $\mathrm{rist}_G(u) \cap B = 1$, so that $\mathrm{rist}_G(u)$ and B commute (by Lemma 5.1.2). Thus $[B \times \mathrm{rist}_G(v)^{G \setminus H}] \geq [\mathrm{rist}_G(v)^H \times \mathrm{rist}_G(v)^{G \setminus H}] = [G]$, and $[B] = [\mathrm{rist}_G(v)^H]$ by Lemma 5.1.4. □

Thus we can think of each element of $\mathcal{B}$ as a block of imprimitivity in the action of G on a layer of T. This is the main idea driving Chapter 7, where the possible branch actions of G are studied using the structure graph.

Chapter 6

On the congruence subgroup property

The congruence subgroup problem for groups acting on rooted trees is defined by analogy with the classical problem for linear algebraic groups ([12]). Let a group G act faithfully on a locally finite rooted tree T. Consider, for each $n \geq 0$, the kernel $\mathrm{St}_G(n)$ of the action of G on $V_n \subset T$; can we "detect" every finite index subgroup of G (or, equivalently, every finite quotient) by looking at the finite quotients $G/\mathrm{St}_G(n)$? In other words, does every finite index subgroup of G contain some $\mathrm{St}_G(n)$? We can rephrase this question in terms of profinite completions. Taking the subgroups $\{\mathrm{St}_G(n) \mid n \geq 0\}$ as a neighbourhood basis for the identity gives a topology on G – the *congruence topology* – and the completion $\overline{G}$ of G with respect to this topology is a profinite group called the *congruence completion* of G. As G acts faithfully on T we have $\bigcap_n \mathrm{St}_G(n) = 1$, so there is an injection $G \hookrightarrow \overline{G}$. *A fortiori*, G is residually finite, so there is also an injection of G into its profinite completion $\widehat{G}$ and $\widehat{G}$ maps onto $\overline{G}$. Asking whether each finite index subgroup of G contains some stabilizer $\mathrm{St}_G(n)$ is tantamount to asking whether the map $\widehat{G} \to \overline{G}$ is injective. The *congruence subgroup problem* asks us to compute the *congruence kernel* C of this map. If C is trivial, we say that G has the *congruence subgroup property*.

This is, in fact, not just an analogy but a generalization of the classical congruence subgroup problem. Suppose for simplicity that we are dealing with a subgroup G of $\mathrm{GL}(d, \mathbb{Z})$. We can form the coset tree $T = \{a + n!\mathbb{Z}^d \mid a \in \mathbb{Z}^d,\ n \in \mathbb{N}\}$ where the edges are given by inclusion. Then G naturally acts faithfully on T by $g(a + n!\mathbb{Z}^d) = ga + n!\mathbb{Z}^d$ and the level stabilizers are precisely the usual principal congruence subgroups (kernels of homomorphisms to $\mathrm{GL}(d, \mathbb{Z}/n!\mathbb{Z})$). Note that we may take instead of $n!$ the product $q_1 \dots q_n$ where $(q_n)_n$ is any increasing sequence of natural numbers, to embed $\mathrm{GL}(d, \mathbb{Z})$ in the automorphism group of a different rooted tree (this is indeed done in [14] for regular rooted trees, but the action cannot be weakly branch as the group contains a non-trivial abelian normal subgroup, contradicting Lemma 4.1.1 (ii)).

In fact, the above may be generalized to automorphisms of residually finite groups. If H is a finitely generated residually finite group, then it has finitely many normal subgroups of any given finite index. Pick some descending chain $(H_n)_n$ of finite index characteristic subgroups of H with trivial intersection and form the coset tree T as above. Then $G = \mathrm{Aut}(H)$ acts on T by $g(rH_n) = g(r)H_n$ and we can formulate the congruence subgroup problem as above. Indeed, a question attributed to Ihara ([62]) asks whether G has the congruence subgroup property when H is a finitely generated non-abelian free group. The answer is 'yes' for the free group of rank 2 ([16]), but very much open for higher ranks.

The main general idea underlying the above variations of the congruence subgroup property can be summarised as 'take a natural family of finite index subgroups of a given group with trivial intersection and determine to what extent all finite index subgroups are determined by this family'. Since a branch group G has another obvious family of finite index subgroups, namely $\{\mathrm{rist}_G(n) \mid n \geq 0\}$, we may ask the same question for this family. Let $\widetilde{G}$ denote the *branch completion* of G; that is, the completion of G with respect to the *branch topology*, which is generated by taking $\{\mathrm{rist}_G(n) \mid n \geq 0\}$ as a neighbourhood basis of the identity. Then, as above, there is a surjective homomorphism $\widehat{G} \to \widetilde{G}$ and we are asked to determine the *branch kernel* B of this map. Note that the branch topology is stronger than the congruence topology so that we may also ask about the kernel R of the map $\widetilde{G} \to \overline{G}$, the *rigid kernel*.

The problem of determining the congruence, branch and rigid kernels for a branch group G was first posed in [10], where the authors also ask a 'preliminary question of great importance':

Question Do any of the kernels depend on the branch action of G?

In other words, is the nature of these kernels a property of the branch action or of the group? It should be noted that, in the classical context of arithmetic groups, it is well known that the congruence kernel does not depend on the the linear embedding of the arithmetic group (see [53]).

We provide a full answer to the above question by proving the following:

Theorem 6.1. *Let G have two branch actions on trees. Then the congruence kernels with respect to these actions coincide.*

Theorem 6.2. *Let G have two branch actions on trees. Then the branch kernels with respect to these actions coincide.*

The above immediately imply that the rigid kernels of a branch group with respect to any two branch actions are naturally isomorphic.

We will deduce these theorems from a more powerful observation, namely that the congruence and branch topologies of a branch group G can be defined in purely group-theoretic terms, with no reference to a branch action, using the *structure graph* $\mathcal{B}$ of G constructed in Chapter 5. This graph depends only on the subgroup structure of G and is related to every tree on which G acts as a branch group. As we shall see in Proposition 6.1.1, if G acts on T as a branch group then T is embeddable G-equivariantly in $\mathcal{B}$, where the action of G on $\mathcal{B}$ is that induced by conjugation on subgroups. Further, the image of T is coinitial in $\mathcal{B}$ (Proposition 5.1.7). Once we have analogues of $\mathrm{St}_G(n)$ and $\mathrm{rist}_G(n)$ for the action of G on $\mathcal{B}$, the above mentioned results are the key to showing that all congruence and branch topologies induced by a branch action coincide; they all agree with the topologies with respect to $\mathcal{B}$. We will also see how similar results hold for weakly branch groups.

6.1 The congruence and branch topologies

Notation. In this section we will deal with different branch actions of the same group G, so $\mathrm{St}_\rho(n)$ and $\mathrm{rist}_\rho(v)$ will denote the stabilizer of the nth layer and the rigid stabilizer of vertex v *with respect to a given branch action* $\rho : G \to \mathrm{Aut}(T)$. The subscript may be omitted when there is no risk of confusion.

Proposition 6.1.1 ([41]). *If G acts as a branch group on a tree T then there is an order-preserving G-equivariant embedding $\phi : T \to \mathcal{B}$ defined by $\phi : v \mapsto [\mathrm{rist}(v)]$.*

Proof. We have already seen that ϕ is G-equivariant as

$$\phi(v)^g = [\mathrm{rist}(v)]^g = [\mathrm{rist}(v)^g] = [\mathrm{rist}(v^g)] = \phi(v^g).$$

That ϕ is order-preserving is also clear since v is a descendant of w if and only if $\mathrm{rist}(v) \leq \mathrm{rist}(w)$.

To see that ϕ is injective, let $\mathrm{rist}(v) \sim \mathrm{rist}(w)$. If v and w were incomparable vertices, then $\mathrm{rist}(v) \cap \mathrm{rist}(w) = 1$ would imply that $\mathrm{rist}(v)$ and $\mathrm{rist}(w)$ are virtually abelian, a contradiction. Thus v and w are comparable; say $v \leq w$. Suppose for a contradiction that $v \neq w$; then there is a vertex $v_2 \neq v$ in the same layer as v such that $v_2 \leq w$ (a 'sibling' of v). Since $v, v_2 \leq w$, we have $\mathrm{rist}(v) \times \mathrm{rist}(v_2) \leq \mathrm{rist}(w)$ and $\mathrm{rist}(v) \leq_{\mathrm{va}} \mathrm{rist}(w)$, as $\mathrm{rist}(v) \sim \mathrm{rist}(w)$. It then follows that $\mathrm{rist}(v) \leq_{\mathrm{va}} (\mathrm{rist}(v) \times \mathrm{rist}(v_2))$ and that $\mathrm{rist}(v_2)$ is virtually abelian. This gives the desired contradiction, so $v = w$, as required. $\square$

Hence the structure graph "contains" all possible branch actions of G. It is then reasonable to define the congruence and branch topologies with respect to the action of G on $\mathcal{B}$.

The congruence topology

In order to define the congruence topology with respect to the action of G on $\mathcal{B}$ we must find analogues of the level stabilizers $\mathrm{St}_\rho(n)$ for a branch action ρ of G. As G acts level-transitively, the obvious analogue in $\mathcal{B}$ of a layer is an orbit b^G where $b = [B] \in \mathcal{B}$ and B is basal. Recall from Lemma 5.2.1 that $\mathrm{St}([B]) = \mathrm{N_G}(B)$, so the analogue of a level stabilizer $\mathrm{St}_\rho(n)$ is an *orbit stabilizer*

$$\mathrm{St}_{\mathcal{B}}(b^G) := \bigcap \{\mathrm{N_G}(B^g) \mid g \in G\}$$

for $b = [B] \in \mathcal{B}$. Since each basal subgroup has only finitely many conjugates, these orbit stabilizers have finite index in G. Furthermore, the intersection of all orbit stabilizers is trivial by Proposition 6.1.1. We take the family $\{\mathrm{St}_{\mathcal{B}}(b^G) \mid b \in \mathcal{B}\}$ as a neighbourhood basis of the identity to define the *congruence topology* of G with respect to the action of G on $\mathcal{B}$. The completion of G with respect to this topology is a profinite group $\overline{G}_{\mathcal{B}}$ onto which the profinite completion $\widehat{G}$ maps. We denote the kernel of this map by $C_{\mathcal{B}}$. In the following theorem we prove that for any branch action ρ of G, the topologies induced by $\{\mathrm{St}_\rho(n) \mid n \geq 0\}$ and by $\{\mathrm{St}_{\mathcal{B}}(b^G) \mid b \in \mathcal{B}\}$ are equal. Thus the congruence kernels with respect to each of them coincide and Theorem 6.1 as stated in the introduction follows.

Theorem 6.3. *For any branch action $\rho : G \to \mathrm{Aut}(T)$, denote by C_ρ the congruence kernel with respect to this action. Then $C_\rho = C_{\mathcal{B}}$.*

Proof. We must show that for every $n \geq 0$ there is some $b \in \mathcal{B}$ such that $\mathrm{St}_\rho(n) \geq \mathrm{St}_{\mathcal{B}}(b^G)$ and conversely, that for every $b \in \mathcal{B}$ there is some $n \geq 0$ such that $\mathrm{St}_{\mathcal{B}}(b^G) \geq \mathrm{St}_\rho(n)$. However, by Proposition 6.1.1, the nth layer V_n of the tree corresponds to an orbit of basal subgroups so that the former statement holds trivially. It therefore suffices to show the latter statement.

For a given $b = [B] \in \mathcal{B}$, Proposition 5.1.7 gives a vertex v of T such that $[\mathrm{rist}_\rho(v)] \leq b$. Suppose that v is in the nth layer of T and let $x \in \mathrm{St}_\rho(n)$. Note that $\mathrm{St}_\rho(n)$ is the pointwise stabilizer of the G-orbit of $[\mathrm{rist}_\rho(v)]$ by Proposition 6.1.1. Then

$$1 \neq [\mathrm{rist}_\rho(v^x)] = [\mathrm{rist}_\rho(v)] \leq b^x \wedge b.$$

Since B is basal this implies that $B^x = B$ (so $b^x = b$). Thus the normal subgroup $\mathrm{St}_\rho(n)$ fixes all elements of b^G and we conclude that $\mathrm{St}_\rho(n) \leq \mathrm{St}_{\mathcal{B}}(b^G)$, as required. $\square$

The branch topology

As with the congruence kernel, Theorem 6.2 will follow from the fact that the branch kernel with respect to a branch action is equal to the analogous object for the action on the structure graph. We must therefore define the branch topology with respect to $\mathcal{B}$. To do this we generalize the notion of a rigid stabilizer. For a non-trivial basal subgroup A of a branch group, define its *rigid normalizer* $\mathrm{R_G}(A)$ by

$$\mathrm{R_G}(A) := \bigcap \{\mathrm{N_G}(B) \mid B \text{ is basal and } A \cap B = 1\}.$$

We will use the following properties of rigid normalizers. These were proved in [41] but the proof given here is shorter.

Proposition 6.1.2. *Let A, A_1, A_2 be basal subgroups and v a vertex of a tree on which G acts as a branch group. Then*

(i) $A \leq \mathrm{R_G}(A) \leq \mathrm{N_G}(A)$;

(ii) *if* $[A_1] \leq [A_2]$ *then* $\mathrm{R_G}(A_1) \leq \mathrm{R_G}(A_2)$ *(in particular,* $\mathrm{R_G}(A_1) = \mathrm{R_G}(A_2)$ *if* $[A_1] = [A_2]$*);*

(iii) $\mathrm{R_G}(A)$ *is basal and the unique maximal element of* $[A]$*;*

(iv) $\mathrm{R_G}(\mathrm{rist}(v)) = \mathrm{rist}(v)$.

Proof. (i) Let B be a basal subgroup with $A \cap B = 1$. Then $[A, B] = 1$ by Lemma 5.1.2 and so $A \leq \mathrm{N_G}(B)$; thus $A \leq \mathrm{R_G}(A)$. If $g \in \mathrm{R_G}(A)$ then g normalizes each of the conjugates of A distinct from A as they are basal and have trivial intersection with A. Therefore g must normalize A itself.

(ii) Since $[A_1] \leq [A_2]$, there is some finite index subgroup K of A_1 such that $K' \leq A_1 \cap A_2$. Now, for a basal subgroup B with $B \cap A_2 = 1$, we have

$$(K \cap B)' \leq K' \cap B \leq A_1 \cap A_2 \cap B \leq A_2 \cap B = 1$$

and $B \cap K \leq_{\mathrm{f}} B \cap A_1$. That is to say, $B \cap A_1$ is virtually abelian and so $B \cap A_1 = 1$. Hence every $g \in \mathrm{R_G}(A_1)$ normalizes B and $\mathrm{R_G}(A_1) \leq \mathrm{R_G}(A_2)$, as required.

(iii) We first show that $\mathrm{R_G}(A) \sim A$ using Lemma 5.1.4(iii). Let D denote the product of the finitely many conjugates of A that are distinct from A; then $A^G = A \times D \trianglelefteq_{\mathrm{va}} G$. Since $A \leq \mathrm{R_G}(A)$, it suffices to show that $\mathrm{R_G}(A) \cap D = 1$. Let $x \in \mathrm{R_G}(A) \cap D$ and B be a basal subgroup. If $B \cap A = 1$ then $x \in \mathrm{R_G}(A)$ normalizes B. If $B \cap A \neq 1$ then $B \cap A \leq A$ is basal and is centralized by $x \in D \leq \mathrm{C_G}(A)$; hence $x \in \mathrm{N_G}(B \cap A) \leq \mathrm{N_G}(B)$ by Lemma 5.2.1. Thus $x \in \bigcap\{\mathrm{N_G}(B) \mid B \text{ is basal }\} = 1$.

To see that $\mathrm{R_G}(A)$ is basal, suppose that $\mathrm{R_G}(A)^g \neq \mathrm{R_G}(A)$ for some $g \in G$. Then, as $\mathrm{R_G}(A)^g = \mathrm{R_G}(A^g)$, we have $A \neq A^g$ by (ii) so that $A^g \cap A = 1$ and $[A^g, A] = 1$ by Lemma 5.1.2. Now, $\mathrm{R_G}(A) \sim A$ and $\mathrm{R_G}(A)^g \sim A^g$, so that $\mathrm{R_G}(A) \cap \mathrm{R_G}(A)^g \sim A \cap A^g = 1$ by Proposition 5.1.5.

For any $H \in [A]$ and any basal subgroup B such that $B \cap A = 1$ we have $[B, A] = 1$, so $B \leq \mathrm{C_G}(A) = \mathrm{C_G}(H)$. This means that $[B, H] = 1$, in particular, $H \leq \mathrm{N_G}(B)$. Hence $H \leq \mathrm{R_G}(A)$.

(iv) By part (i), it suffices to show that $\mathrm{R_G}(\mathrm{rist}(v)) \leq \mathrm{rist}(v)$. Let $g \in \mathrm{R_G}(\mathrm{rist}(v))$ and $w \in T \backslash T_v$. If w is incomparable with v then $\mathrm{rist}(w) \cap \mathrm{rist}(v) = 1$ and so g fixes w. If $v \leq w$ then $\mathrm{rist}(v) \leq \mathrm{rist}(w)$ and $\mathrm{R_G}(\mathrm{rist}(v)) \leq \mathrm{N_G}(\mathrm{rist}(v)) \leq \mathrm{N_G}(\mathrm{rist}(w))$ by Lemma 5.2.1. Thus $w = w^g$ and the claim follows. □

By analogy with the rigid stabilizers of layers, the *rigid normalizer of an orbit* $a^G = [A]^G$ in $\mathcal{B}$ is defined to be $\mathrm{R_G}(a^G) := \mathrm{R_G}(A)^G$. That these rigid normalizers of orbits have finite index in G follows from Proposition 5.1.7. It also follows that the intersection of all of them is trivial. We may therefore take $\{\mathrm{R_G}(a^G) \mid a \in \mathcal{B}\}$ as a neighbourhood basis of the identity to generate the *branch topology* of G with respect to the action of G on $\mathcal{B}$. The completion $\widetilde{G}_{\mathcal{B}}$ with respect to this topology is a profinite group in which G can be embedded. We denote by $B_{\mathcal{B}}$ the kernel of the map $\widehat{G} \to \widetilde{G}_{\mathcal{B}}$. In our last theorem we prove that for any branch action ρ of G, the topologies induced by $\{\mathrm{rist}_\rho(n) \mid n \geq 0\}$ and by $\{\mathrm{R_G}(a^G) \mid a \in \mathcal{B}\}$ are equal; consequently, the branch kernels with respect to each of them coincide and Theorem 6.2 as stated in the introduction follows.

Theorem 6.4. *For a branch action $\rho : G \to \mathrm{Aut}(T)$, denote by B_ρ the branch kernel with respect to this action. Then $B_\rho = B_{\mathcal{B}}$.*

Proof. As with the proof for the congruence kernels, we must show that for every $n \geq 0$ there is some $b \in \mathcal{B}$ such that $\mathrm{rist}_\rho(n) \geq \mathrm{R_G}(b^G)$ and that for every $b \in \mathcal{B}$ there is some $n \geq 0$ such that $\mathrm{R_G}(b^G) \geq \mathrm{rist}_\rho(n)$. The former claim follows from Proposition 6.1.2(iv) since all rigid stabilizers of vertices are basal. It therefore suffices to prove the latter claim. Given $b \in \mathcal{B}$, Proposition 5.1.7 yields that $[\mathrm{rist}_\rho(v)] \leq b$ for some $v \in T$. Suppose that $v \in V_n \subset T$. We then have $\mathrm{rist}_\rho(v) = \mathrm{R_G}(\mathrm{rist}_\rho(v)) \leq \mathrm{R_G}(b)$ by Proposition 6.1.2, and the transitivity of G on each of the layers of T implies that $\mathrm{rist}_\rho(n) \leq \mathrm{R_G}(b^G)$, as required. □

6.2 The congruence subgroup problem for weakly branch groups

One may ask whether Theorem 6.3 can be applied in the weakly branch setting. In order to follow the same strategy, we would have to define the congruence topology in terms of some analogue of the structure graph of a weakly branch group. However, the definition of the structure graph crucially depends on the fact that all proper quotients of a branch group are virtually abelian. To mimic the construction in the weakly branch setting we would need to know that every proper quotient of a weakly branch group is in some reasonable class $\mathfrak{X}$. Then two subgroups with finitely many conjugates in a weakly branch group would be equivalent if they "differ by a group in $\mathfrak{X}$". It seems that any such $\mathfrak{X}$ would have to be quite large as there are examples ([60]) of weakly branch groups which are just non-(abelian-by-P-type). A group K is of *P-type* for some group P if K is isomorphic to a subgroup of $wr^k(P)$, a regular wreath product of P iterated a finite number k of times.

Fortunately, the arguments used in the proof of Theorem 6.3 also apply in the case of weakly branch groups, using the results of Chapter 4. The congruence kernel C_ρ with respect to a weakly branch action $\rho : G \to \mathrm{Aut}(T)$ is defined exactly as in the branch action setting.

Theorem 6.5. *Let G be a weakly branch action. The congruence kernels of any two weakly branch actions of G coincide.*

Proof. Let $\rho : G \to \mathrm{Aut}(T_\rho)$ and $\sigma : G \to \mathrm{Aut}(T_\sigma)$ be two weakly branch actions of G on rooted trees T_ρ and T_σ. We will show that every level stabilizer with respect to ρ contains a level stabilizer with respect to σ. The argument is symmetrical, so we conclude that the sets $\{\mathrm{St}_\rho(n) \mid n \geq 0\}$ and $\{\mathrm{St}_\sigma(n) \mid n \geq 0\}$ are cofinal. Thus, on taking each of $\{\mathrm{St}_\rho(n) \mid n \geq 0\}$ and $\{\mathrm{St}_\sigma(n) \mid n \geq 0\}$ as a neighbourhood basis of the identity, we obtain the same topology on G and hence the same congruence kernel.

For a given $n \geq 0$ choose a vertex v in T_ρ of level n. Then $\mathrm{rist}_\rho(v)$ is a non-trivial normal subgroup of a finite index subgroup of G, so Theorem 4.2 gives a vertex u of T_σ such that $1 \neq \mathrm{rist}_\sigma(u)' \leq \mathrm{rist}_\rho(v)$. Suppose that u is in the mth layer of T_σ and let $x \in \mathrm{St}_\sigma(m)$. Then

$$1 \neq (\mathrm{rist}_\sigma(u)')^x = \mathrm{rist}_\sigma(u^x)' = \mathrm{rist}_\sigma(u)' \leq \mathrm{rist}_\rho(v^x) \cap \mathrm{rist}_\rho(v).$$

This implies that $\mathrm{rist}_\rho(v^x) = \mathrm{rist}_\rho(v)$ and x stabilizes v. Thus the normal subgroup $\mathrm{St}_\sigma(m)$ fixes all elements of the G-orbit of v; that is, it fixes pointwise the nth layer of T_ρ (as G acts transitively on every layer). We conclude that $\mathrm{St}_\sigma(m) \leq \mathrm{St}_\rho(n)$, as required. $\square$

Examples of weakly branch (but not branch) groups include:

- the just non-solvable torsion-free Brunner–Sidki–Vieira group acting on the binary tree, generated by two copies of the adding machine ([15]),
- the Basilica group studied in [36],
- the just non-(abelian-by-P-type) groups constructed and studied in [60].

Question Which of the above have the congruence subgroup property?

Chapter 7

Uniqueness of branch actions

The fact that the congruence subgroup problem is independent of the (weakly) branch action of a group prompts one to ask whether all (weakly) branch actions of a group are in some sense "equivalent". The focus here will be on branch actions since there is a structure theory for them. From now on, G denotes a branch group. Recall that we have put a partial order on the tree T on which G acts so that $v \leq u$ if u is on the unique path from v to the root. Thus, the *supremum* $\sup(u, v)$ of vertices u, v is the vertex furthest from the root in the common path from u, v to the root. Continuing the "descendant" terminology from previous chapters, we will say that v is a *child* of u if $v < u$ and they are adjacent.

The study of the possible branch actions of G reduces to studying its structure graph $\mathcal{B}$, since all trees on which G has a branch action can be embedded G-equivariantly into this graph (Proposition 6.1.1). The first thing to note is that any branch action of G completely determines $\mathcal{B}$ as, by Proposition 5.2.3, every element of $\mathcal{B}$ corresponds to a block of imprimitivity in the action of G on some layer of T.

7.1 Unique branch actions

If $\mathcal{B}$ is (isomorphic to) a tree, then the choice of a tree on which G can act as a branch group is very limited. Indeed, G acts on $\mathcal{B}$ as a branch group: each layer of $\mathcal{B}$ consists of blocks of imprimitivity in the action of G on some layer of T, so G must act transitively on these blocks and the rigid stabilizer of each layer of $\mathcal{B}$ contains the rigid stabilizer of some layer of T, so they must all have finite index in G. Moreover, every other tree on which G has a branch action consists of infinitely many layers of $\mathcal{B}$, with the same order as in $\mathcal{B}$ (this is proved in [41, Corollary 15.2.3] but it follows almost immediately from Proposition 6.1.1, Proposition 5.1.7 and the assumption that $\mathcal{B}$ is a tree). We say that every other tree is obtained from $\mathcal{B}$ by *deleting layers* of $\mathcal{B}$, or that every other tree is a *deletion of layers* of $\mathcal{B}$. Indeed, this is a sufficient condition for $\mathcal{B}$ to be a tree.

Proposition 7.1.1. *If every tree on which G has a branch action is obtained from some T by deleting layers, then T and $\mathcal{B}$ are isomorphic as G-spaces.*

Proof. Since T can be embedded G-equivariantly in $\mathcal{B}$, it suffices to show that this embedding is also a surjection. Let $b \in \mathcal{B}$. Then, by Proposition 5.2.3, $b = [\mathrm{rist}(u)^H]$ for some $u \in T$ and $H \geq \mathrm{St}(u)$. We may assume without loss of generality that $\mathrm{St}(u)$ is a maximal subgroup of H so that $[\mathrm{rist}(u)]$ is a child of b in $\mathcal{B}$, as follows: $\mathrm{St}(u)$ is a maximal subgroup of some $K \leq H$. If $[\mathrm{rist}(u)^K]$ is again not equal to any $[\mathrm{rist}(v)]$ with $v \in T$, then replace H by K; if $[\mathrm{rist}(u)^K] = [\mathrm{rist}(v)]$ for some $v \in T$ then replace $[\mathrm{rist}(u)]$ by $[\mathrm{rist}(v)]$ and repeat the procedure. After finitely many steps, we will find some $w \in T$ and $K \leq H$ with $\mathrm{St}(w)$ maximal in K such that $[\mathrm{rist}(w)^K]$ is not equal to any $[\mathrm{rist}(v)]$ with $v \in T$, so we take $w = u$ and $K = H$.

Now, since $H \leq_{\mathrm{f}} G$ we can find a finite chain $H = H_n \leq H_{n-1} \leq \cdots \leq H_0 = G$ where H_i is maximal in H_{i-1}. We shall use these subgroups to construct the top of a tree on which G acts as a branch group. Each of the subgroups in the above chain is the stabilizer of some block of imprimitivity in the action of G on $\mathcal{B}$ so $[\mathrm{rist}(u)], b = [\mathrm{rist}(u)^H], [\mathrm{rist}(u)^{H_{n-1}}], \ldots, [\mathrm{rist}(u)^{H_1}], [\mathrm{rist}(u)^{H_0}] = [G]$ is a path in $\mathcal{B}$ from $[\mathrm{rist}(u)]$ to $[G]$ passing through b. Now, since G acts transitively on the cosets of each H_i, each set $\{[\mathrm{rist}(u)^{H_i g}] \mid g \in G\}$ is the orbit of $[\mathrm{rist}(u)^{H_i}]$ in $\mathcal{B}$. Thus, taking these orbits, we obtain a finite rooted tree on which G acts level-transitively with finite index rigid level stabilizers. To form an infinite tree, attach the rest of T below the layer containing u. By construction, G acts as a branch group on this new tree, which must by assumption be a deletion of layers of T. Hence $b = [\mathrm{rist}(v)]$ for some $v \in T$ and $\mathcal{B} \cong T$. □

Consider the collection of all rooted trees on which G acts as a branch group. The relation 'T_1 is obtained from T_2 by deleting layers' (in other words, all layers of T_1 are layers of T_2) is a partial order on the collection of all rooted trees on which G acts as a branch group. Given any chain of trees, their union is an upper bound (it must be a tree since, if there were any cycles, they would already be contained in some tree in the chain). Zorn's lemma then yields a maximal tree from which all other trees in the chain are obtained by deleting layers. This proves

Proposition 7.1.2. *The structure graph $\mathcal{B}$ of G is isomorphic to a tree if and only if it is the unique maximal tree on which G acts as a branch group.*

Intuitively, this means that there is essentially only one branch action of G. This raises two questions:

1. When is $\mathcal{B}$ a tree?

2. Are there any examples?

Obviously, it is not practical to check every tree on which G acts to see whether it is a deletion of layers of some other tree on which G acts; we may not even be able to conceive of all the possible trees. So we need something more practicable. The answer (explicitly proved in [41]) follows from Proposition 5.2.3, which shows that each vertex in $\mathcal{B}$ corresponds to a block of imprimitivity in the action of G on a layer of the tree.

Proposition 7.1.3. *The structure graph $\mathcal{B}$ of G is a tree if and only if there exists a tree T on which G has a branch action where any of the following equivalent conditions holds:*

(i) *for every $v \in T$, the set of blocks of imprimitivity containing v is linearly ordered by inclusion;*

(ii) *for each $v \in T$ there is a unique chain $\mathrm{St}(v) = H_0 \leq H_1 \leq \cdots \leq H_n = G$ such that H_i is a maximal subgroup of H_{i+1} for each i (indeed, each H_i must be the stabilizer of a vertex in the unique path from v to the root);*

(iii) *the smallest block containing two incomparable vertices $u, v \in V_n$ consists precisely of all the vertices $w \in V_n$ such that $w \leq \sup(u, v)$.* □

This also provides a trivial example: the whole automorphism group of T satisfies this condition as it is the inverse limit $\varprojlim_{n\in\mathbb{N}}(\mathrm{Sym}(m_n) \wr \cdots \wr \mathrm{Sym}(m_1))$. Of course, we are interested in less trivial examples but it is not easy to check that a given branch group satisfies the above condition.

It was proved in [34] that if a branch group acts on a tree T satisfying $(*)$ and $(**)$ below then $\mathcal{B}$ is isomorphic to T:

($*$) for each $v \in T$, the stabilizer $\mathrm{St}(v)$ acts primitively and regularly on the children of v (and therefore $\mathrm{St}(v)$ acts like a cyclic group of prime order on these children);

($**$) for every pair $u, u' \in T$ of incomparable vertices and every $v < u$ there exists $g \in G$ which fixes u' and moves v.

In fact, we can do better than this.

Proposition 7.1.4. *The structure graph $\mathcal{B}$ of a branch group G is isomorphic to T if and only if the branch action of G on T satisfies*

(A) *for each $v \in T$, the stabilizer $\mathrm{St}(v)$ acts primitively on the children of v,*

(B) *for every pair $u, u' \in T$ of incomparable vertices and every $v < u$ there exists $g \in G$ which fixes u, u' and moves v.*

Proof. Suppose first that the action of G on T satisfies (A) and (B). As pointed out in the above discussion, it will suffice to show that for any pair u, v of vertices in V_n, the smallest block containing both of them consists of all $w \in V_n$ with $w \leq \sup(u, v)$. Since there is some $g \in G$ taking u to v, this is equivalent to showing[1] that $\langle \mathrm{St}(u), g\rangle = \mathrm{St}(\sup(u, u^g))$. We proceed by induction on the distance n between u and $s := \sup(u, u^g)$. If $n = 0$ then $g \in \mathrm{St}(u)$ and there is nothing to prove. If $n = 1$ then u, u^g are children of s and (A) implies that $\mathrm{St}(u)$ is a maximal subgroup of $\mathrm{St}(s)$, so $\langle \mathrm{St}(u), g\rangle = \mathrm{St}(s)$. Suppose then that $n > 1$ and that the claim is true for all pairs of vertices whose supremum is at distance less than n from them. Let $t > u$ be a child of s. Then u^g and t are incomparable, so (B) implies that there exists $h \in \mathrm{St}(u^g) \cap \mathrm{St}(t)$ moving u. This means that $\sup(u, u^h) = t$ is at distance $n-1$ from u, so the induction hypothesis ensures that $\langle \mathrm{St}(u), h\rangle = \mathrm{St}(t)$. But now, $h \in \mathrm{St}(u^g)$ is equal to k^g for some $k \in \mathrm{St}(u)$, so we have

$$\mathrm{St}(t) = \langle \mathrm{St}(u), k^g\rangle \leq \langle \mathrm{St}(u), g\rangle \leq \mathrm{St}(s)$$

and, since $\mathrm{St}(t)$ is maximal in $\mathrm{St}(s)$ by (A), we conclude that $\langle \mathrm{St}(u), g\rangle = \mathrm{St}(s)$, which finishes the induction.

Conversely, suppose that for any pair of vertices u, v in a layer, the smallest block containing both of them consists of all vertices in that layer below $\sup(u, v)$. If u, v are

[1] This is the approach used in [41] to show that ($*$) and ($**$) imply that $\mathcal{B} \cong T$ and it was this proof that inspired the definition of (A) and (B).

children of $\sup(u, v)$ then the above supposition implies that $\mathrm{St}(\sup(u, v))$ acts primitively on the children of $\sup(u, v)$, so (A) follows. Consider now a pair u, u' of incomparable vertices and another vertex $v < u$. Recall that our hypothesis is equivalent to there being for each $w \in T$ a unique chain $\mathrm{St}(w) \leq \cdots \leq G$ of subgroups, each maximal in the next one, and that these subgroups must in fact be stabilizers of vertices in the unique path from w to the root. Thus, since u' and v are incomparable, $\mathrm{St}(u') \not\leq \mathrm{St}(v)$ so $\langle \mathrm{St}(u'), \mathrm{St}(v) \rangle > \mathrm{St}(v)$. By hypothesis, $\langle \mathrm{St}(u'), \mathrm{St}(v) \rangle = \mathrm{St}(w)$ for some $w > v$. In particular, $w \geq u$ and $\langle \mathrm{St}(u'), \mathrm{St}(v) \rangle \geq \mathrm{St}(u)$. Hence there exists $g \in \mathrm{St}(u') \cap \mathrm{St}(u)$ moving v, so (B) follows. $\square$

On to checking which branch groups have essentially unique branch actions. In the case of GGS groups, it is somewhat easier to work with the slightly stronger conditions $(*), (**)$. The authors of [34] already show there that the (first) Grigorchuk group and the Gupta–Sidki 3-group satisfy $(*)$ and $(**)$. A slight modification of the proof for the Gupta–Sidki 3-group yields the following more general result.

Proposition 7.1.5. *For any GGS-group acting on the p-regular tree with non-zero defining vector, the given action satisfies $(*), (**)$.*

Proof. We follow the same argument as in [34]. Since every GGS-group is self-similar (that is, every vertex section is isomorphic to G as a permutation group) and G cyclically permutes the p vertices below the root, condition $(*)$ is satisfied.

To prove that $(**)$ also holds, let $u, u' \in T$ be incomparable vertices and suppose that s is their supremum. For given $v < u$, if there exists $h \in G_s$, the vertex section of G at s, which fixes u' while moving v then any preimage $g \in G$ of h also fixes u' while moving v. But G_s is isomorphic to G as a permutation group, so we may assume that s is the root.

Now, G acts level-transitively on T, so we can suppose that $u' = p^n$ for some n (having labelled the vertices of T as words in $\{1, \ldots, p\}$). Since u and u' are incomparable, v is in a subtree rooted at $i \neq p$. If $e_i \neq 0$ then b fixes u' while moving all vertices below i, so $(**)$ holds. If $e_i = 0$, consider $b^{a^{p-i}}$ which has a^{e_i} in the pth position (therefore fixes u') and $a^{e_{2i}}$ in the ith position. If $e_{2i} \neq 0$ then $b^{a^{p-i}}$ finishes the proof. If $e_{2i} = 0$ then we repeat the above procedure, reaching $b^{a^{p-(p-2)i}} = b^{a^{2i}}$ if necessary. Then $b^{a^{2i}}$ has $a^{e_{(p-2)i}}$ in pth position (fixing u') and $a^{e_{(p-1)i}}$ in ith position. But now $e_{(p-1)i}$ must be non-zero as we have cycled through all other entries in $\mathbf{e}$, discovering that each is zero, and have assumed that $\mathbf{e}$ is non-zero. Thus $b^{a^{2i}}$ is an element which fixes u' and moves v. This shows that $(**)$ is satisfied. $\square$

Note that these examples also have the congruence subgroup property, so one might wonder whether there is any connection between these properties. It turns out that there is no such connection, as there are examples of branch groups with non-trivial congruence kernel and satisfying $(*), (**)$.

Example 7.1.6. Recall from Chapter 2 the definition of an EGS group Γ over the p-regular rooted tree. Given $(e_1, \ldots, e_{p-1}) \in \mathbb{F}_p^{p-1}$, a defining vector, Γ is generated by a, b, c where a, b are defined as for GGS groups and $c = (c, a^{e_1}, \ldots, a^{e_{p-1}})$. Pervova showed in [52] that all EGS groups with non-symmetric defining vector (which are just infinite branch groups) do not have the congruence subgroup property. The congruence kernel of such a group is a profinite abelian group of exponent p. Since they contain the GGS group generated by a, b, the proof of Proposition 7.1.5 shows that all EGS groups with nonzero defining vector satisfy $(*), (**)$. In particular, those with non-symmetric defining vector provide our desired examples.

Here is another example which is not a GGS group. The *Hanoi tower group* H was introduced in [30] and its properties are further studied in [31]. It models the Hanoi towers game on three pegs and is defined as the group of automorphisms of the ternary tree generated by

$$a = (1, 1, a)(12), \quad b = (1, b, 1)(13), \quad c = (c, 1, 1)(23).$$

It was shown in [10] that the Hanoi tower group does not have the congruence subgroup property. Its congruence kernel is a metabelian torsion-free profinite group.

Proposition 7.1.7. *The Hanoi towers group satisfies* $(*), (**)$ *and therefore its structure graph is isomorphic to the ternary tree.*

Proof. As shown in in [31], H acts transitively on all levels of the ternary tree and every vertex section of H is equal to H. Thus, since H acts primitively and regularly on the first level vertices, it satisfies $(*)$.

To show that it satisfies $(**)$, let $u, u' \in T$ be incomparable and $v < u$. As in the proof of Proposition 7.1.5, we can suppose that the supremum of u, u' is the root vertex.

Since H acts transitively on all levels, we may assume that $u' = 3^n$ for some n; that is, u' is in the right-most ray of the tree, so it is fixed by $a = (1, 1, a)(12)$. As u, u' are incomparable, v must be either in the subtree rooted at 1 or that rooted at 2; hence v is moved by a. $\square$

Remark 3. Having an essentially unique branch action is not merely a curious or nice property of a branch group, it also has important rigidity consequences. It was shown in [34] that if G is a branch group whose structure graph is (isomorphic to) a tree T then every automorphism of G is induced by a conjugation of G in $\operatorname{Aut} T$. This had been proved by Sidki for the Gupta–Sidki 3-group in [58], where he moreover computed the automorphism group. It was also shown in [43] for weakly branch groups satsifying an extra condition on subgroups, examples of which include the Grigorchuk and Gupta–Sidki groups, and the weakly branch Brunner–Sidki–Vieira group from [15].

This brings two questions to mind. Firstly, might it be that if a branch group is rigid (meaning its automorphism group is equal to its normalizer in $\operatorname{Aut} T$ for some T on which it acts) then its structure graph is isomorphic to T? Secondly, what is the connection between the ideas in [34] and [43]? Could it happen that branch groups with a tree as structure graph are actually a special case of the weakly branch groups considered in [43]?

7.2 Non-unique branch actions

We have just seen several examples of branch groups with an essentially unique branch action as their structure graph is a tree. Are there examples of branch groups whose structure graph is not a tree? Is there an easy way of checking?

We saw in the last section that $\mathcal{B}$ is a tree if and only if the action of G on it satisfies conditions (A) and (B) from Proposition 7.1.4 and that any other tree T' on which G has a branch action is obtained from $\mathcal{B}$ by deleting layers. Upon closer inspection of (A) and (B), it becomes clear that the action on T' need not satisfy (A) as the children of a vertex in T' may be its "grandchildren" in $\mathcal{B}$ and therefore $\operatorname{St}(v)$ cannot act primitively on them. However, the missing layers of T' do not affect condition (B), which will be satisfied. Indeed, if the action of G on T' does not satisfy (B) then neither does the action on any tree from which T' was obtained by deleting layers, so $\mathcal{B}$ cannot be a tree.

Example 7.2.1. The second Grigorchuk group Γ is a GGS group acting on the 4-regular tree (whose vertices are labelled by finite words in the alphabet $\{1,2,3,4\}$). The defining vector of Γ is $(1,0,1)$ so that $b=(a,1,a,b)$. It follows from Theorem 2.2 and Theorem 2.3 that Γ is an infinite 2-group which is fractal and a branch group. We shall now see that its structure graph is not a tree, as the action on the 4-regular tree does not satisfy (B). To avoid confusion, $\mathrm{St}(V_1)$ will denote the first level stabilizer. It suffices to find two incomparable vertices u,u' and $v<u$ such that $\mathrm{St}(u')\cap\mathrm{St}(u)\leq\mathrm{St}(v)$. Take, for example, $u=1, u'=31, v=11$. Then $\mathrm{St}(u)=\mathrm{St}(V_1)=\langle b,b^a,b^{a^2},b^{a^3}\rangle$ and each of these generators has the same action on the vertices $\{11,12,13,14\}$ as on $\{31,32,33,34\}$. Thus, any element of $\mathrm{St}(u)$ fixing $u'=31$ must also fix $v=11$, and our claim follows.

Furthermore, by the same argument, any GGS group on the p^n-regular tree whose defining vector satisfies the following condition will also fail to satisfy (B): $e_i=0$ for all $i\in(p^{n-1}\mathbb{Z}/p^n\mathbb{Z})\setminus\{0\}$ and for each nontrivial coset H of $p^{n-1}\mathbb{Z}/p^n\mathbb{Z}$ either $e_h=0$ for all $h\in H$ or $e_h\neq 0$ for all $h\in H$. An example would be, for $p=3, n=2$ the vector $(1,2,0,1,4,0,1,2)$, which defines $b=(a,a^2,1,a,a^4,1,a,a^2,b)$. Note that this can only hold when $n>1$.

What can we say about the possible branch actions of G when its structure graph is not a tree? How different can these actions be? It follows from Proposition 5.1.7 that any two branch actions of G must have quite a lot in common as every $\mathrm{rist}(n)$ with respect to one action contains some $\mathrm{rist}(m)$ with respect to the other. Thus two branch actions can only really differ in their action on "finite tops" of the trees on which they act. Pushing this simple idea a bit further yields restrictions on the types of trees on which G can have a branch action.

Proposition 7.2.2. *Let $(m_n)_n$, $(k_n)_n$ be two sequences of integers bigger than 2 such that either k_0 is coprime to all m_n or m_0 is coprime to all k_n. Denote by T_k, respectively T_m, the rooted tree of type (k_n), respectively (m_n). Then G cannot have a branch action on both T_k and T_m.*

Proof. Assume without loss of generality that k_0 is coprime to all m_n and suppose for a contradiction that G acts as a branch group on both T_k and T_m. Let u be a vertex in the first layer of T_k. Then, since $\mathrm{rist}(u)$ is a basal subgroup of G, it is equivalent to $\langle\mathrm{rist}(v)\rangle^L$ for some $v\in T_m$ and $\mathrm{St}(v)\leq L\leq G$. Suppose that v is in the nth layer of T_m. Now, $\mathrm{rist}(u)$ has k_0 conjugates (including itself), so the L-orbit of v must be one of k_0 blocks of imprimitivity in the action of G on the nth layer of T_m. Since there are $N := m_0 m_1 \dots m_n$ vertices in the mth layer of T_m and k_0 is coprime to N, the layer cannot possibly be partitioned into k_0 blocks. This contradiction finishes the proof. $\square$

So, in particular, a group cannot have two branch actions on regular trees of coprime degree.

Question: In the case when the structure graph is not isomorphic to a tree, what does it look like?

Appendix A

The finite subgroups of the Grigorchuk group

Here is the proof of the characterization of finite subgroups of the Grigorchuk group promised in Chapter 3. Since it makes heavy use of results from [33], these are reproduced here with minimal changes (mainly so that notation agrees with that introduced previously; so, for instance, "projection" has been changed to "vertex section"). It is also worth noting that the length function used here is that introduced in Subsection 2.2.1, which agrees with the usual word length of finitely generated groups in the case of the Grigorchuk group. All other notation agrees with the rest of the thesis; in particular, Γ denotes the Grigorchuk group generated by a, b, c, d (see Definition 2.1.4).

Proposition A.0.3. *The Grigorchuk group Γ is an infinite 2-group and $\langle b, ac\rangle$, $\langle c, ad\rangle$, $\langle d, ab\rangle$ are subgroups of index 2 (see [17, p. 221, p. 226]). Moreover, if S is a subset of $\{a, b, c, d\}$ then $\langle S\rangle$ is either equal to Γ or finite.* $\square$

Lemma A.0.4 ([33, Lemma 6])**.** *Let H be a subgroup of Γ. Suppose that $H \not\leq \mathrm{St}(1)$ and that not all of the eight third level vertex sections of $\mathrm{St}_H(3)$ are equal to Γ. Let P be either first level vertex section of $\mathrm{St}_H(1)$.*

(i) *If $H\Gamma = \langle a\rangle\Gamma$ then one of the following holds.*
  - (a) *$P\Gamma = \langle ac\rangle\Gamma$.*
  - (b) *$P\Gamma = \langle ad\rangle\Gamma$.*
  - (c) *$P \leq \mathrm{St}(1)$.*

(ii) *If $H\Gamma = \langle ab\rangle\Gamma$ then $P\Gamma = \langle ac\rangle\Gamma$.*

(iii) *If $H\Gamma = \langle ac\rangle\Gamma$ then $P\Gamma = \langle ad\rangle\Gamma$.*

(iv) *If* $H\Gamma = \langle ad\rangle\Gamma$ *then* $P \leq \mathrm{St}(1)$.

(v) *If* $H\Gamma = \langle a, b\rangle\Gamma$ *then* $P\Gamma = \langle a, c\rangle\Gamma$.

(vi) *If* $H\Gamma = \langle a, c\rangle\Gamma$ *then* $P\Gamma = \langle a, d\rangle\Gamma$.

(vii) *If* $H\Gamma = \langle a, d\rangle\Gamma$ *then* $P \leq \mathrm{St}(1)$.

Lemma A.0.5 ([33, Lemma 7]). *Let* $g = (g_0, g_1) \in \mathrm{St}_\Gamma(1)$ *and let* w *be a word of length* $l(g)$ *in* a, b, c, d *which represents* g. *Let* l_a *be the number of factors in* w *equal to* a *and write* $\varepsilon = l(g) - 2l_a$; *thus* $\varepsilon \in \{-1, 0, 1\}$.

(a) For $i = 0, 1$, we have $l(g_i) \leq \frac{1}{2}(l(g) + \varepsilon)$. Moreover, $l(g_0) \leq \frac{1}{2}l(g)$ unless the first and last factors in w are from $\{b, c\}$, and $l(g_0) \leq \frac{1}{2}(l(g) - 1)$ if the first or last factor is d or if both the first and last factors are a.

(b) Let $g \in \mathrm{St}_\Gamma(2)$. Then $l(\varphi_i(g_0)) \leq \frac{1}{4}(l(g) + 1)$ for $i = 0, 1$.

Lemma A.0.6 ([33, Lemma 8]). *Let* $g \in \Gamma$.

(a) *If* g *stabilizes each of the vertices* 0000 *and* 0001 *and if* $l(g) \leq 13D$, *where* $D > 2$, *then* $l(\varphi_i\varphi_0^3(g)) < D$ *for* $i = 0, 1$.

(b) *If* g *stabilizes each of the vertices* 000 *and* 001 *and if* $l(g) \leq 5D$, *where* $D > 2$, *then* $l(\varphi_i\varphi_0^2(g)) < D$ *for* $i = 0, 1$.

Lemma A.0.7 ([33, Lemma 9]). *Let* H *act transitively on a set* Ω *with* n *elements where* $n > 2$. *Let* $\alpha \in \Omega$ *and let* $L = \mathrm{stab}_H(\alpha)$. *Suppose that* $H = \langle S\rangle$, *where* $S = S^{-1}$, *and that* $H = LS^{n-2}$. *Then all elements of* $S \setminus L$ *map* α *to the same element and all products of two elements of* $S \setminus L$ *lie in* L.

Theorem A.1 ([33, Theorem 3]). *Let* $\mathfrak{X}$ *be a family of subgroups of* Γ *with the following properties.*

(i) $1 \in \mathfrak{X}$ *and* $\Gamma \in \mathfrak{X}$.

(ii) *If* $H \in \mathfrak{X}$, *then* $L \in \mathfrak{X}$ *for each subgroup* L *such that* $H \leq L$ *and* $|H : L|$ *is finite.*

(iii) *If* H *is a finitely generated subgroup of* $\mathrm{St}(1)$ *and both first level vertex sections of* H *are in* $\mathfrak{X}$ *then* $H \in \mathfrak{X}$.

Then all finitely generated subgroups of Γ *are in* $\mathfrak{X}$.

Proof. We note first that if $\mathfrak{X}$ satisfies conditions (i)–(iii) above then so does the family $\{H \mid H^g \in \mathfrak{X} \text{ for all } g \in \Gamma\}$. We replace $\mathfrak{X}$ by this subfamily; thus we may suppose that if H is in $\mathfrak{X}$ then so is each conjugate of H in Γ. Suppose that there are finitely generated subgroups which are not in $\mathfrak{X}$, and among such subgroups choose H to be one having a finite generating set S such that $D = \max(l(s) \mid s \in S)$ is as small as possible. Then $D > 1$, by Proposition A.0.3. If $H \not\leq \mathrm{St}(1)$ then from (ii) at least one of the two first level vertex sections of H is not in $\mathfrak{X}$, and this vertex section has a generating set with elements of length at most $\frac{1}{2}(D+1)$ by Lemma A.0.5. This is a contradiction to the choice of H. Therefore one of the cases of Lemma A.0.4 applies.

(1) If case (i)(a), case (ii) or case (v) applies, then let $v = 000$.

(2) If case (i)(b), case (iii) or case (vi) applies, then let $v = 00$.

(3) If case (i)(c), case (iv) or case (vii) applies, then let $v = 0$.

Let L be the stabilizer of v in H. Then the corresponding vertex section L_v of L lies in $\mathrm{St}(1)$, by Lemma A.0.4. We shall show that the two vertex sections of L_v at the vertices below v have generating sets of elements of length less than D and that at least one of these vertex sections is not in $\mathfrak{X}$. This contradiction will finish the proof of the theorem. We consider separately the three choices for the definition of v.

Case (1): The elements of $H \setminus \mathrm{St}_H(1)$ interchange the two sets of four vertices of the third layer below the vertices $0, 1$, and the two vertex sections P_0 , P_1 of $\mathrm{St}_H(1)$ in Γ are conjugate in Γ; thus P_0 , P_1 are not in $\mathfrak{X}$ from (i) and (ii). By Lemma A.0.4, each of P_0, P_1 contains an element of the form acy with $y \in \Gamma$. The groups P_0, P_1 act on the subtrees with roots 0, 1; their elements of form acy interchange the two sets of two vertices of the second layer below the vertices of the first layer in these subtrees, and the two vertex sections of each $\mathrm{St}_{P_i}(1)$, which are the four vertex sections at level 2 of $\mathrm{St}_H(2)$, are conjugate in Γ, not in $\mathfrak{X}$, and by Lemma A.0.4 they contain elements of the form adw with $w \in \Gamma$. Finally, all eight vertex sections of $\mathrm{St}_H(3)$ are not in $\mathfrak{X}$ and we see that H acts transitively on the vertices of the third layer, and from Lemma A.0.4 that $L_v \leq \mathrm{St}(1)$. Since L_v is not in $\mathfrak{X}$, one of its first level vertex sections is not in $\mathfrak{X}$.

We may add to S the element 1 and the inverses of the elements of S. Suppose that $H \neq LS^6$. If the hypothesis of Lemma A.0.4 case (v) applies, then $H\Gamma = \langle a, b\rangle\Gamma$ and S must have elements of at least two of the types ax, aby, bz with $x, y, z \in \Gamma$; but the square of aby does not fix v and the images of v under elements ax, bz are clearly distinct from v and each other, and we have a contradiction by Lemma A.0.7. If instead the hypothesis of case (i)(a) or case (ii) holds, then the discussion of these cases in Lemma A.0.4 shows that S contains an element whose square $q = (q_0, q_1)$ has q_0 congruent to ac modulo Γ, so that q cannot fix v, and again we have a contradiction by Lemma A.0.7.

Hence $H = LS^6$, and so there is a transversal T to L in H consisting of products of at most six elements of S. Therefore the elements of the generating set $S_L = L \cap \{t_1^{-1} s t_2 \mid t_1, t_2 \in T, s \in S\}$ for L are products of at most 13 elements of S and have length at most $13D$. From Lemma A.0.6, the images in the two first level vertex sections of L_v of the elements of S_L have length less than D, as required.

Case (2): The argument in this case is similar to the one for case (1), but easier. First we see that H acts transitively on the vertices of the second layer, that $L_v \leq \mathrm{St}(1)$, and that at least one of the first level vertex sections of L_v is not in $\mathfrak{X}$.

We may assume that $1 \in S = S^{-1}$. If $H \neq LS^2$, then the argument of the second paragraph of case (1), with b, c replaced by c, d, gives a contradiction.

Hence $H = LS^2$, and so there is a transversal T to L in H consisting of products of at most two elements of S. Therefore the elements of the generating set $S_L = L \cap \{t_1^{-1} 1 s t_2 \mid t_1, t_2 \in T, s \in S\}$ have length at most $5D$. By Lemma A.0.6, the images under the two projection maps from L_v to Γ of the elements of S_L have length less than D, as required.

Case (3): Arguing as for cases (1) and (2), we see that L_v is in $\mathrm{St}(1)$, and that its first level vertex sections are not in $\mathfrak{X}$.

By hypothesis we have $\mathrm{St}_H(1) = \mathrm{St}_H(2)$, so that $L = \mathrm{St}_H(1)$. Choose $s_0 \in S \setminus L$. Then L is generated by the elements of $S \cap L$, their conjugates under s_0 , and the products of pairs of elements of $S \setminus L$. In particular, L has a generating set S_L of elements of length at most $3D$. By Lemma A.0.6, the images of L in Γ corresponding to the vertex sections at vertices 00 and 01 have a generating set of elements of length at most $\frac{1}{4}(3D + 1)$, as required. $\square$

Theorem A.2. *Let H be a finitely generated subgroup of the Grigorchuk group Γ. Then either H is finite or one of its vertex sections is equal to Γ.*

Proof. We start with 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 Proposition A.0.3, either H is finite or $H = \Gamma$.

Assume that the assertion in the theorem holds whenever $D \leq n$ with $n \geq 1$. For $D = n + 1$ we consider the cases $H \leq \mathrm{St}(1)$ and $H \nleq \mathrm{St}(1)$ separately. If $H \leq \mathrm{St}_\Gamma(1)$ 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 Lemma A.0.5. By inductive hypothesis, each H_u is either finite or has a vertex section equal to Γ, in which case H has a vertex section equal to Γ. If both H_u are finite, then so is H, as it is a subdirect product of its first level vertex sections.

If $H \nleq \mathrm{St}_\Gamma(1)$ there are three cases to consider depending on the vertex section H_0 at vertex 0. If $H_0 \leq \mathrm{St}_\Gamma(1)$ then, by Case 3 in the proof of Theorem A.1, H_{00} and H_{01} are generated by elements of length less than D. Thus each of H_{00} and H_{01} is either finite or has a vertex section equal to Γ, by inductive hypothesis. If the latter, then H has a vertex section equal to Γ. If both H_{00} and H_{01} are finite then so is H_0, as it is a subdirect product of H_{00} and H_{01}. Now, H_0 and H_1 are conjugate in Γ as $H \nleq \mathrm{St}_\Gamma(1)$; hence H is finite.

If $H_0\Gamma' = \langle ad\rangle\Gamma'$ or $\langle a, d\rangle\Gamma'$, then $H_{00} \leq \mathrm{St}_\Gamma(1)$ by Lemma A.0.4. Case 2 of the proof of Theorem A.1 shows that H_{000} and H_{001} are generated by elements of length less than D so that either one of them has a vertex section equal to Γ, or they are both finite. If the latter holds, then 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 $g = ((g_{00}, g_{01}), g_1) \in \mathrm{St}_H(00)$ we have $g^h = ((g_{01}^{s_1}, g_{000}^{s_0}), g_1^{h_1})$, whence $H_{00}^{s_1} = H_{01}$. 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 $H_{000} \leq \mathrm{St}_\Gamma(1)$ by Lemma A.0.4. Case 1 of the proof of Theorem A.1 shows that H_{0000} and H_{0001} are generated by elements of length less than D so that either one of them has a vertex section equal to Γ or both of them and therefore H_{000} are finite. Since H acts transitively on the third layer of the tree, there is an element $h \in H$ swapping the vertices 000 and 001. Then H_{000} and H_{001} are conjugate in Γ, 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. $\square$

Bibliography

[1] A. G. Abercrombie. Subgroups and subrings of profinite rings. *Math. Proc. Cambridge Philos. Soc.*, 116(2):209–222, 1994.

[2] S. V. Aleshin. Finite automata and burnside's problem for periodic groups. *Math. Notes*, 11(3):199–203, 1972.

[3] Theofanis Alexoudas, Benjamin Klopsch, and Anitha Thillaisundaram. Maximal subgroups of multi-edge spinal groups. *Groups Geom. Dyn.* To appear. Pre-print http://arxiv.org/abs/1312.5615.

[4] Yiftach Barnea and Aner Shalev. Hausdorff dimension, pro-p groups, and Kac–Moody algebras. *Trans. Amer. Math. Soc.*, 349(12):5073–5091, 1997.

[5] Laurent Bartholdi. Endomorphic presentations of branch groups. *J. Algebra*, 268(2):419 – 443, 2003.

[6] Laurent Bartholdi. Lie algebras and growth in branch groups. *Pacific J. Math.*, 218(2):241–282, 2005.

[7] Laurent Bartholdi, Bettina Eick, and René Hartung. A nilpotent quotient algorithm for certain infinitely presented groups and its applications. *Internat. J. Algebra Comput.*, 18(08):1321–1344, 2008.

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

[9] Laurent Bartholdi, Rostislav I. Grigorchuk, and Zoran Šunić. Branch groups. In *Handbook of algebra, Vol. 3*, pages 989–1112. North-Holland, Amsterdam, 2003.

[10] Laurent Bartholdi, Olivier Siegenthaler, and Pavel A. Zalesskii. The congruence subgroup problem for branch groups. *Israel J. Math.*, 187(1):419–450, 2012.

[11] Laurent Bartholdi and Zoran Šunić. On the word and period growth of some groups of tree automorphisms. *Comm. Algebra*, 29(11):4923–4964, 2001.

[12] Hyman Bass, John Milnor, and J P Serre. Solution of the congruence subgroup problem for SL_n ($n \geq 3$) and Sp_{2n} ($n \geq 2$). *Publ. Math. Inst. Hautes Études Sci.*, 33(1):59–137, 1967.

[13] Gilbert Baumslag. *Topics in combinatorial group theory*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1993.

[14] A. M. Brunner and Said Sidki. The generation of $\mathrm{GL}(n, \mathbf{Z})$ by finite state automata. *Internat. J. Algebra Comput.*, 8(1):127–139, 1998.

[15] A.M. Brunner, Said Sidki, and Ana Cristina Vieira. A just-nonsolvable torsion-free group defined on the binary tree. *J. Algebra*, 211(1):99 – 114, 1999.

[16] Kai-Uwe Bux, Mikhail V. Ershov, and Andrei S. Rapinchuk. The congruence subgroup property for $\mathrm{Aut}\, F_2$: a group-theoretic proof of Asada's theorem. *Groups Geom. Dyn.*, 5(2):327–353, 2011.

[17] Pierre de la Harpe. *Topics in geometric group theory*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 2000.

[18] M. Edjvet and Stephen J. Pride. The concept of "largeness" in group theory. II. In *Groups – Korea 1983 (Kyoungju, 1983)*, volume 1098 of *Lecture Notes in Math.*, pages 29–54. Springer, Berlin, 1984.

[19] Jacek Fabrykowski and Narain Gupta. On groups with sub-exponential growth functions. II. *J. Indian Math. Soc. (N.S.)*, 56(1-4):217–228, 1991.

[20] Kenneth Falconer. *Fractal geometry*. John Wiley & Sons, Ltd., Chichester, 1990. Mathematical foundations and applications.

[21] Gustavo A. Fernández-Alcober and Amaia Zugadi-Reizabal. Spinal groups: semidirect product decompositions and Hausdorff dimension. *J. Group Theory*, 14(4):491–519, 2011.

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

[23] Alejandra Garrido. Abstract commensurability and the Gupta–Sidki group. *Groups Geom. Dyn.* Accepted. Pre-print http://arxiv.org/abs/1310.0493.

[24] Alejandra Garrido. On the congruence subgroup problem for branch groups. *Israel J. Math.* Accepted. Pre-print http://arxiv.org/abs/1405.3237.

[25] Alejandra Garrido and John S. Wilson. On subgroups of finite index in branch groups. *J. Algebra*, 397(0):32 – 38, 2014.

[26] Rostislav I. Grigorchuk. On Burnside's problem on periodic groups. *Funktsional. Anal. i Prilozhen.*, 14(1):53–54, 1980.

[27] Rostislav I. Grigorchuk. Degrees of growth of finitely generated groups and the theory of invariant means. *Izv. Akad. Nauk SSSR Ser. Mat.*, 48(5):939–985, 1984.

[28] Rostislav I. Grigorchuk. Just infinite branch groups. In *New horizons in pro-p groups*, volume 184 of *Progr. Math.*, pages 121–179. Birkhäuser Boston, Boston, MA, 2000.

[29] Rostislav I. Grigorchuk, W. N. Herfort, and Pavel A. Zalesskii. The profinite completion of certain torsion p-groups. In *Algebra (Moscow, 1998)*, pages 113–123. de Gruyter, Berlin, 2000.

[30] Rostislav I. Grigorchuk and Zoran Šunić. Asymptotic aspects of Schreier graphs and Hanoi Towers groups. *C. R. Math. Acad. Sci. Paris*, 342(8):545–550, 2006.

[31] Rostislav I. Grigorchuk and Zoran Šunić. Self-similarity and branching in group theory. In *Groups St. Andrews 2005. Vol. 1*, volume 339 of *London Math. Soc. Lecture Note Ser.*, pages 36–95. Cambridge Univ. Press, Cambridge, 2007.

[32] Rostislav I. Grigorchuk and John S. Wilson. The conjugacy problem for certain branch groups. *Tr. Mat. Inst. Steklova*, 231(Din. Sist., Avtom. i Beskon. Gruppy):215–230, 2000.

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

[34] Rostislav I. Grigorchuk and John S. Wilson. The uniqueness of the actions of certain branch groups on rooted trees. *Geom. Dedicata*, 100:103–116, 2003.

[35] Rostislav I. Grigorchuk and John 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.

[36] Rostislav I. Grigorchuk and Andrzej Żuk. On a torsion-free weakly branch group defined by a three state automaton. *Internat. J. Algebra Comput.*, 12(1-2):223–246, 2002. International Conference on Geometric and Combinatorial Methods in Group Theory and Semigroup Theory (Lincoln, NE, 2000).

[37] Narain Gupta and Said Sidki. On the Burnside problem for periodic groups. *Math. Z.*, 182(3):385–388, 1983.

[38] Narain Gupta and Said Sidki. Some infinite p-groups. *Algebra i Logika*, 22(5):584–589, 1983.

[39] Narain Gupta and Said Sidki. Extension of groups by tree automorphisms. In *Contributions to group theory*, volume 33 of *Contemp. Math.*, pages 232–246. Amer. Math. Soc., Providence, RI, 1984.

[40] Marshall Hall, Jr. Coset representations in free groups. *Trans. Amer. Math. Soc.*, 67:421–432, 1949.

[41] P. Hardy. *Aspects of abstract and profinite group theory.* PhD thesis, University of Birmingham, 2002.

[42] René Hartung. Coset enumeration for certain infinitely presented groups. *Internat. J. Algebra Comput.*, 21(8):1369–1380, 2011.

[43] Yaroslav Lavreniuk and Volodymyr Nekrashevych. Rigidity of branch groups acting on rooted trees. *Geom. Dedicata*, 89:159–179, 2002.

[44] Darren Long and Alan W. Reid. *Surface subgroups and subgroup separability in 3-manifold topology.* Publicações Matemáticas do IMPA. IMPA, Rio de Janeiro, 2005.

[45] I.G. Lysenok. A system of defining relations for a Grigorchuk group. *Math. Notes*, 38:784–792, 1985.

[46] Anatoly 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.

[47] Volodymyr Nekrashevych. *Self-similar groups*, volume 117 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2005.

[48] B. H. Neumann. Some remarks on infinite groups. *J. Lond. Math. Soc.*, 12:120–127, 1937.

[49] Peter M. Neumann. Some questions of Edjvet and Pride about infinite groups. *Illinois J. Math.*, 30(2):301–316, 1986.

[50] 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.

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

[52] E. L. Pervova. Profinite completions of some groups acting on trees. *J. Algebra*, 310(2):858–879, 2007.

[53] Gopal Prasad and Andrei S. Rapinchuk. Developments on the congruence subgroup problem after the work of Bass, Milnor and Serre. In *Collected papers of John Milnor. V. Algebra*, pages 307–325. American Mathematical Society, Providence, RI, 2010. Edited by Hyman Bass and T. Y. Lam.

[54] Derek J. S. Robinson. *A course in the theory of groups*, volume 80 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1993.

[55] A. V. Rozhkov. *Finiteness conditions in groups of tree automorphisms*. Habilitation thesis, Chelyabinsk State University, 1996.

[56] A. V. Rozhkov. Lower central series of a group of tree automorphisms. *Math. Notes*, 60(2):165–174, 1996.

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

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

[59] Said Sidki. On a 2-generated infinite 3-group: the presentation problem. *J. Algebra*, 110(1):13–23, 1987.

[60] Said Sidki. Just non-(abelian by P-type) groups. In *Infinite groups: geometric, combinatorial and dynamical aspects*, volume 248 of *Progr. Math.*, pages 389–402. Birkhäuser, Basel, 2005.

[61] Zoran Šunić. Hausdorff dimension in a family of self-similar groups. *Geom. Dedicata*, 124:213–236, 2007.

[62] B. Sury. Bounded generation does not imply finite presentation. *Comm. Algebra*, 25(5):1673–1683, 1997.

[63] Ana Cristina Vieira. On the lower central series and the derived series of the Gupta-Sidki 3-group. *Comm. Algebra*, 26(4):1319–1333, 1998.

[64] Taras Vovkivsky. Infinite torsion groups arising as generalizations of the second Grigorchuk group. In *Algebra (Moscow, 1998)*, pages 357–377. de Gruyter, Berlin, 2000.

[65] John S. Wilson. Groups with every proper quotient finite. *Proc. Cambridge Philos. Soc.*, 69:373–391, 1971.

[66] John S. Wilson. On just infinite abstract and profinite groups. In *New horizons in pro-p groups*, volume 184 of *Progr. Math.*, pages 181–203. Birkhäuser Boston, Boston, MA, 2000.

[67] John S. Wilson and Pavel A. Zalesskii. Conjugacy separability of certain torsion groups. *Arch. Math. (Basel)*, 68(6):441–449, 1997.