Tree languages and branched groups

Laurent Bartholdi[1] · Marialaura Noce[2]

Abstract

We study the portraits of isometries of rooted trees—the labelling of the tree, at each vertex, by the permutation of its descendants—in terms of languages. We characterize regularly branched self-similar groups in terms of ω-regular languages. We deduce the algorithmic decidability of some problems, such as the comparison of regularly branched contracting groups, and their orbit structure on the boundary of the rooted tree.

1 Introduction

There is a rich interplay between group theory and theoretical computer science, and more precisely the theory of formal languages. We develop it in the context of *self-similar groups* acting on rooted trees.

1.1 Regular languages

Let A be a finite set called *alphabet*. A *language* is a subset of A^*, where A^* denotes the set of all finite words in A. Some languages can be described by a *finite state automaton*: a device $\mathcal{M}$ with a finite amount of memory that reads letter-by-letter words over the alphabet A as input and decides at the end whether to accept the word. An automaton $\mathcal{M}$ *accepts* the word w if there exists a run $q_0, \ldots, q_n$ on w such that $q_n \in Q_f$, where Q_f denotes the set of final states. The language $\mathcal{L}(\mathcal{M})$ is the set of words accepted by the automaton $\mathcal{M}$. A

✉ Marialaura Noce
mnoce@unisa.it

Laurent Bartholdi
laurentbartholdi@gmail.com

1 Fakultät für Mathematik und Informatik, Universität des Saarlandes, 66123 Saarbrücken, Germany

2 Mathematisches Institut, Georg-August Universität zu Göttingen, 37073 Göttingen, Germany

language K is *regular* if there exists a finite state automaton $\mathcal{M}$ such that $K = \mathcal{L}(\mathcal{M})$. For more information we refer to [15].

Consider a group G generated (as a monoid) by a set A. The set of words in A^* that represent the identity in G is a language $W(G)$, called the *word problem* of G, and determines G. It is a still active area of research to relate the complexity of $W(G)$, as a language, with structural properties of G. The first result in this direction is due to Anisimov, who in 1970 proved that $W(G)$ is a regular language if and only if the group G is finite [2].

The group $G = \langle A \rangle$ may also be understood via a *normal form* for its elements: a language $N(G) \subseteq A^*$ that is in bijection with G under the natural evaluation map; for example, the free Abelian group $\mathbb{Z}^2$ can be represented by the regular language $(\{x\}^* \cup \{x^{-1}\}^*)(\{y\}^* \cup \{y^{-1}\}^*)$.

Automata may also be used to represent relations: adding an extra padding symbol $\diamond$ to A, one may represent a relation $R \subseteq A^* \times A^*$ by its language

$$\{(u \diamond^{\max(0, |v|-|u|)}, v \diamond^{\max(0, |u|-|v|)}) : (u, v) \in R\},$$

and the relation R is called *regular* if its language is regular as a subset of $((A \sqcup \{\diamond\}) \times (A \sqcup \{\diamond\}))^*$. The rich class of *automatic groups* are those groups $G = \langle A \rangle$ admitting a regular normal form $N \subseteq A^*$ and for which the multiplication relations $\{(u, v) \in N^2 : ua =_G v\}$ are regular for all $a \in A$; see [12].

Kharlampovich, Khoussainov, and Miasnikov introduced Cayley-automatic groups in [16] as groups admitting a regular normal form (not necessarily coming from words over a generating set) for their elements and the group operation. They prove that this class of groups strictly extends that of automatic groups, while retaining some of the most important algorithmic properties (such as fast solvability of the word problem). In this paper, the groups we are interested in shall not admit regular normal forms in the above sense; but rather as "tree languages", see below.

The theory of regular languages has been extended to infinite words by Büchi, see [9]. In contrast to a classical finite automaton, a Büchi machine is a finite state automaton that takes infinite words as input and and accepts a run if the set of recurring states intersects the set of accepting states. An ω-*regular* language is a language $\subset A^\omega$ recognized by a Büchi machine. The class of ω-regular languages is closed under all Boolean operations; this was used to prove decidability of Pressburger arithmetic [8].

The theory was further extended to *tree languages*. Finite words are seen as A-labels on the vertices of a finite path, and ω-words are A-labelled infinite paths. In this vein, *tree words* are A-labelled rooted trees (infinite in our case), and a *tree automaton* is a finite machine accepting certain tree words. It reads trees starting from the root, one node at a time, and clones itself to process independently and in parallel all descendant subtrees. These tools are at the heart of Rabin's proof [21] that the monadic second-order logic of two successor functions (S2S) is decidable.

1.2 Self-similar groups

Finite state automata also appear in group theory as representations of the actual *elements* of a group. For an alphabet X, the set of words X^* naturally forms the vertex set of a regular rooted tree $\mathcal{T}$, rooted at the empty word and with an edge between u and ux for all $u \in X^*$, $x \in X$. Consider a group G acting by tree isometries on $\mathcal{T}$, and $g \in G$. Then the action of g on $\mathcal{T}$ is entirely characterized by its action σ_g on X, the neighbours of the root, and for each $x \in X$ the induced bijection between the subtrees xX^* and $\sigma_g(x)X^*$. Identifying these subtrees with $\mathcal{T}$, we obtain a permutation σ_g and a collection $(g@x)_{x \in X}$ of tree isometries called the

states of g. If $g@x \in G$ for all $x \in X$, we call the group G *self-similar*. In this case, the action of G on $\mathcal{T}$ is conveniently recorded by a homomorphism $\psi : G \to G \wr X$, defined by $g \mapsto ((g@x), \sigma_g)$, where $G \wr X$ stands for the permutational wreath product $G^X \rtimes \mathrm{Sym}(X)$.

Furthermore, the operations $g \mapsto g@x$ may be composed; for $v = ux \in X^*$ we write $g@v = (g@u)@x$. If g is such that $\{g@v : v \in X^*\}$ is finite, then g may be represented by a finite automaton with alphabet $X \times X$: its stateset is $Q = \{g@v : v \in X^*\}$ and it has for all $h \in Q, x \in X$ a transition from h to $h@x$ with label $(x, \sigma_h(x))$. As an example, consider $X = \{0, 1\}$, the stateset $Q = \{a, b, c, d, e\}$, and the transitions

$$
\begin{aligned}
(a, (0, 1)) &\mapsto e, & (a, (1, 0)) &\mapsto e, \\
(b, (0, 0)) &\mapsto a, & (b, (1, 1)) &\mapsto c, \\
(c, (0, 0)) &\mapsto a, & (c, (1, 1)) &\mapsto d, \\
(d, (0, 0)) &\mapsto e, & (d, (1, 1)) &\mapsto b, \\
(e, (0, 0)) &\mapsto e, & (e, (1, 1)) &\mapsto e,
\end{aligned}
$$

written in graph form (with the state written inside nodes) as follows:
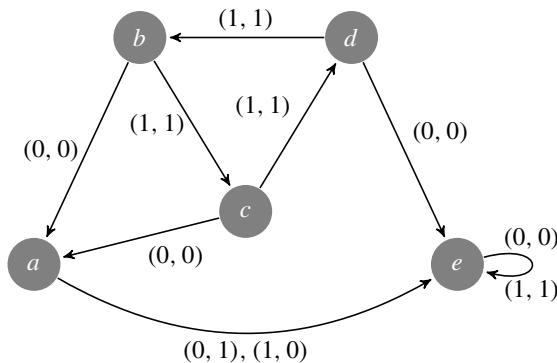
Every choice of initial state $q \in Q$ yields a relation $R_q \subset X^* \times X^*$ which is the graph of a permutation of X^*; the group generated by these permutations is known as the (first) Grigorchuk group [13] and possesses a wealth of striking properties: it is a finitely generated torsion group which is infinite but all its proper quotients are finite; it is a group of intermediate word growth; it is amenable but not elementarily amenable, …. Another close example is the Gupta-Sidki 3-group, defined as follows. Consider $X = \{0, 1, 2\}$, the stateset $Q = \{a, a^{-1}, b, b^{-1}, e\}$, and the transitions

$$
\begin{aligned}
(a, (0, 1)) &\mapsto e, & (a, (1, 2)) &\mapsto e, & (a, (2, 0)) &\mapsto e, \\
(a^{-1}, (0, 2)) &\mapsto e, & (a^{-1}, (1, 0)) &\mapsto e, & (a^{-1}, (2, 1)) &\mapsto e, \\
(b, (0, 0)) &\mapsto a, & (b, (1, 1)) &\mapsto a^{-1}, & (b, (2, 2)) &\mapsto b, \\
(b^{-1}, (0, 0)) &\mapsto a^{-1}, & (b^{-1}, (1, 1)) &\mapsto a, & (b^{-1}, (2, 2)) &\mapsto b^{-1}, \\
(e, (0, 0)) &\mapsto e, & (e, (1, 1)) &\mapsto e, & (e, (2, 2)) &\mapsto e,
\end{aligned}
$$

and again in graph form as follows:
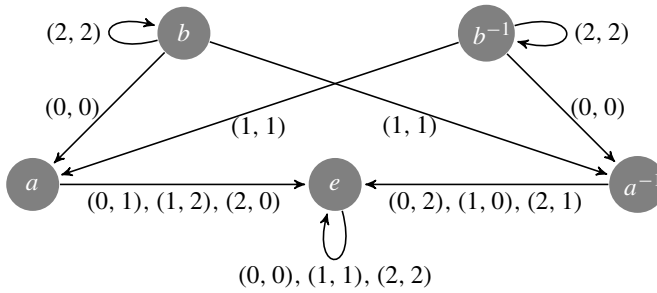
This example easily generalizes to $X = \{0, 1, \ldots, p - 1\}$ for any prime $p > 2$, yielding the family of Gupta-Sidki p-groups. The Grigorchuk group and the Gupta-Sidki p-groups are instances of *GGS-groups*, and provide a large family of counterexamples to the General Burnside Problem.

Important subclasses of groups acting on trees have emerged: *recurrent groups* are those for which the map '$g \mapsto g@v$' is surjective from the stabilizer of $v \in \mathcal{T}$ onto G. *Level-transitive groups* are those that act transitively on X^n for all $n \in \mathbb{N}$. Loosely speaking, a branched group is a group whose lattice of subnormal subgroups is similar to the corresponding structure in the full isometry group Isom $\mathcal{T}$ of the tree $\mathcal{T}$. For a self-similar group G, there is a stronger notion: G is *regularly branched* if there is a finite-index subgroup K such that G contains, for all $v \in \mathcal{T}$, a copy of K acting only on the subtree vX^*. The aforementioned examples of the Grigorchuk and Gupta-Sidki torsion groups are recurrent, level-transitive and regularly branched.

Every $f \in$ Isom $\mathcal{T}$ can be described by providing, at every vertex v of the tree, a permutation $\sigma_{f@v} \in \text{Sym}(X)$ that describes how f acts on the children of v. The *portrait* of f is this Sym(X)-labelling $v \mapsto \sigma_{f@v}$ of $\mathcal{T}$, and there is a one-to-one correspondence between isometries of $\mathcal{T}$ and portraits.

The relations between a self-similar group G and the collection of portraits of its elements have already been considered in the literature, in particular by Siegenthaler in [22] and his 2008 doctoral thesis; see also [18, 19, 24, 26]. In [20], Penland and Šunić prove that the closures of certain self-similar groups of rooted trees that satisfy an algebraic law do not have a regular language of portraits.

1.3 Main results and their consequences

This paper's main results relate tree languages with group-theoretical properties of branched groups.

Theorem A (= Theorem 1) *Let G be a level-transitive, recurrent, closed self-similar subgroup of* Isom $\mathcal{T}$. *Then G is regularly branched if and only if its set of portraits is a tree regular language.*

We point out that in the "if" direction of the theorem above, we do not need that G is level-transitive.

Theorem B (= Theorem 2) *Let G be a contracting regularly branched group. Then its set of portraits is a tree regular language.*

Theorems A and B solve some algorithmic decidability problems in regularly branched groups. In the following statements, a regularly branched group is given by a finite collection of automata generating it, with the promise that the group is regularly branched.

Corollary C *Let* $G, H \leq \mathrm{Isom}\, \mathcal{T}$ *be two contracting finitely generated regularly branched groups; then there exists an algorithm that determines whether* $G \leq H$. *As a consequence, it is decidable if* $G = H$.

For a group G acting on a tree X^* and therefore on its boundary X^ω, the *orbit relation* is the equivalence relation

$$\mathcal{O} = \{(\xi, \eta) \in X^\omega \times X^\omega : \eta \in G\xi\} \subseteq (X \times X)^\omega.$$

Corollary D *Let* $G \leq \mathrm{Isom}\, \mathcal{T}$ *be a finitely generated, contracting, regularly branched group; then its orbit relation is an* ω-*regular language, and is computable.*

The *Hausdorff dimension* of profinite groups was first considered by Abercrombie in [1]. In our situation of a closed group G acting on a rooted tree X^*, its Hausdorff dimension is computed as follows: let G_n be the natural quotient of G acting on X^n; then

$$\mathrm{Hdim}(G) = \liminf_{n \to \infty} \frac{\log \# G_n}{\log \# (\mathrm{Isom}\, X^*)_n}.$$

Corollary E *Let* $G \leq \mathrm{Isom}\, \mathcal{T}$ *be a finitely generated, closed self-similar regularly branched group; then the Hausdorff dimension of* G *is computable.*

The paper is organized as follows: in Sect. 2 we define groups acting on trees, with a particular emphasis on branched groups. In Sect. 3, we give basic notions about tree languages, in Sect. 4 we prove Theorems A and B, and in the last Sect. 5 we present some applications and we prove some corollaries of Theorems A and B.

1.4 Finitely constrained groups

We conclude this introduction by pointing out some interesting applications of the corollaries above that give a better understanding of the closure and algebraic properties of branched groups.

The article [20] by Penland and Šunić considers closely related questions; in particular, they point to self-similar branched groups whose language of portraits enjoy various properties. The definition of regular tree language that we gave is the most general; another, more restrictive accepting condition is known as the *Büchi condition*, and an example of a group whose language is regular but not Büchi is $\{g \in \mathrm{Isom}\, \mathcal{T} : g(v)$ and v differ in finitely many positions for all $v \in X^*\}$. A yet more restrictive condition is being a *sofic tree shift*. In our terminology, this amounts to all subsets of states being residual, and precisely captures topologically closed tree languages. In [14], Grigorchuk considered *finitely constrained* groups, a notion we now recall.

Definition 1 A *pattern* is an isometry f of a finite truncation of the regular tree $\mathcal{T}$, and an element $g \in \mathrm{Isom}\, \mathcal{T}$ is said to *avoid the pattern* f if its action at every subtree is different from f.

A closed subgroup $G \leq \mathrm{Isom}\, \mathcal{T}$ is *finitely constrained* if there exists a finite set F of patterns such that $g \in \mathrm{Isom}\, \mathcal{T}$ belongs to G if and only if it avoids all patterns in F.

Not every regular language is defined by a finite collection of forbidden patterns, and this also applies to groups. In fact, it is not difficult to see that, while admitting a regular

language of patterns corresponds to being branched, admitting a finitely constrained language of patterns corresponds to being branched over a level stabilizer ($G \cap$ the kernel of a map π_m as above).

For closed groups, these notions correspond. Let indeed G be a closed subgroup of Isom $\mathcal{T}$, regularly branched over the subgroup K. Then G is also regularly branched over $\overline{K}$, the topological closure of K; now $\overline{K}$ is closed and finite-index, and therefore open in G. It follows that $\overline{K}$ contains a level stabilizer $G \cap \ker \pi_m =: L$. We have $L^X \le \psi(\overline{K})$ because $L \le \overline{K}$ which is branched, and $L^X \le \psi(\ker \pi_{m+1}) \le \psi(\ker \pi_m)$, so $L^X \le \psi(L)$ and G is regularly branched over L.

More generally, Penland and Šunić prove for a closed group G that, under a technical condition (G must be normalized by a level-transitive, self-similar, recurrent group), the language of portraits of G is regular if and only if it is finitely constrained.

In [26], Šunić showed that a group defined by forbidden patterns of depth $s + 1$ is the topological closure of a self-similar, countable, regularly branched group over its stabilizer of level s. Furthermore, it is proven in [25] and [3] that all possible values of the Hausdorff dimension of finitely constrained groups acting on the binary tree can be determined by understanding its forbidden patterns. Now it is clear that the tree language defined by a finite collection of forbidden patterns is regular, but there are regular languages that require an infinite number of forbidden patterns; for example, the GGS-groups defined by a constant vector, and the "Basilica group" (This is akin to the distinction, in symbolic dynamics, between subshifts of finite type and sofic shifts).

The first author and Siegenthaler constructed in [4] an example of group, called the "twisted twin" of the Grigorchuk group, which is not isomorphic to the Grigorchuk group but has the same closure. Following the notation of the Introduction, its automaton is given by the alphabet $X = \{0, 1\}$, the stateset $Q = \{a, b, c, d, e\}$, and the transitions

$$
\begin{aligned}
(a, (0, 1)) &\mapsto e, & (a, (1, 0)) &\mapsto e, \\
(b, (0, 0)) &\mapsto c, & (b, (1, 1)) &\mapsto a, \\
(c, (0, 0)) &\mapsto a, & (c, (1, 1)) &\mapsto d, \\
(d, (0, 0)) &\mapsto e, & (d, (1, 1)) &\mapsto b, \\
(e, (0, 0)) &\mapsto e, & (e, (1, 1)) &\mapsto e
\end{aligned}
$$

(note the "twist" on the entries for b), in graph form
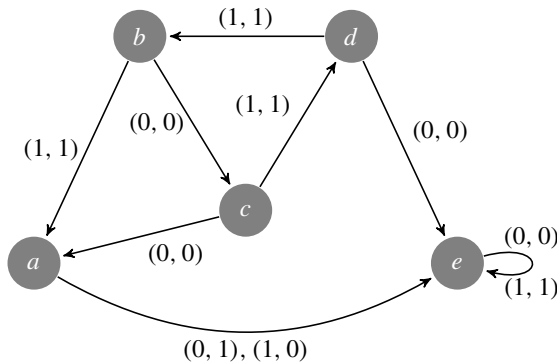
It can be checked algorithmically that the Grigorchuk group and its twisted twin have the same closure, using Theorem B: since their pattern languages are computable and regular, so are their closures, and equality of regular tree languages is decidable. On the other hand, it can also be checked algorithmically that the Grigorchuk group and its twisted twin are not

isomorphic, by computing their abelianizations. There is code in the first author's computer algebra package "FR" that performs these tasks automatically.

Finally, we remark that the growth of the tree language of in the profinite case gives the Hausdorff dimension, and in the finitely generated case gives growth of portraits. This growth has been considered for some examples, and in both cases it is doubly exponential with computable constants, see [27].

However, and this is the main point we wish to stress in this article, these notions do not correspond for general (non-closed) self-similar groups. This is best seen in the context of the *congruence property*. There are groups, such as the "twisted twin " mentioned above, that satisfy all the hypotheses of our main theorem—they are self-similar, level-transitive, contracting, recurrent—and are branched, but not on a level stabilizer. Even more strikingly, the derived subgroup of the "twisted twin" group $\widetilde{G}$ is not closed in the tree topology, equivalently does not contain $\widetilde{G} \cap \ker \pi_m$ for any m. Thus its set of portraits is regular, but the group is not finitely constrained. This answers an open question in [20].

2 Branched groups

In this Section we briefly summarize and recall the most important notions in the theory of groups of isometries of rooted trees and, more precisely, of branched groups. For more information on the topic see, for example, [5–7, 11].

2.1 Groups of isometries of rooted trees

Let $\mathcal{T}$ be a *d-regular rooted* tree: an oriented tree with a designated vertex called the root, and such that every vertex has d immediate descendants and one immediate ancestor (except the root which has no ancestor).

Let X be a finite alphabet with d elements, X^n the set of all words of length n and $X^* = \bigcup_{n \in \mathbb{N}} X^n$; then the vertex set of $\mathcal{T}$ may be identified with X^*. We denote by ϵ the empty word, namely the root of the tree. The descendants of $u \in X^n$ are all words $ux \in X^{n+1}$, for $x \in X$. For every vertex $u \in X^*$, the subtree of $\mathcal{T}$ hanging from u has vertex set uX^* and is naturally identified with $\mathcal{T}$ via $uw \leftrightarrow w$. Let $\mathrm{Isom}\,\mathcal{T}$ denote the group of bijective maps of the set X^* which preserve incidence, so are isometries of the graph $\mathcal{T}$. Let $\mathrm{Sym}(X)$ be the set of all permutations over the alphabet X. Considering separately the action of $\mathrm{Isom}\,\mathcal{T}$ on X and on each subtree xX^* for $x \in X$, we obtain an isomorphism

$$\psi \colon \mathrm{Isom}\,\mathcal{T} \to (\mathrm{Isom}\,\mathcal{T}) \wr X := (\mathrm{Isom}\,\mathcal{T})^X \rtimes \mathrm{Sym}(X)$$

decomposing every isometry into a permutation of X and a tuple of elements of $\mathrm{Isom}\,\mathcal{T}$. For $g \in \mathrm{Isom}\,\mathcal{T}$, we write $\psi(g)_1$, respectively $\psi(g)_2$ the elements of $\mathrm{Isom}\,\mathcal{T}^X$, respectively $\mathrm{Sym}(X)$; then $\psi(g)_1(x)$ is the map $v \mapsto xv \overset{g}{\mapsto} g(xv) = yw \mapsto w$, and $\psi(g)_2$ is the map $x \overset{g}{\mapsto} g(x)$, which we abbreviate σ_g.

Definition 2 (Portraits) Anticipating Sect. 3, a *tree word* is a map $\mathcal{T} \to A$ for some alphabet A. The set of tree words is naturally a topological space for the product topology on $A^{\mathcal{T}}$; it is a Cantor set as soon as A has more than one element and X is non-empty.

The *portrait* of $g \in \mathrm{Isom}\,\mathcal{T}$ is the tree word $p_g \colon \mathcal{T} \to \mathrm{Sym}(X)$ defined inductively by $p_g(\epsilon) = \sigma_g$ and $p_g(xv) = p_{\psi(g)_1(x)}(v)$. The portrait map induces a homeomorphism between $\mathrm{Isom}\,\mathcal{T}$ and the set of all tree words $\mathcal{T} \to \mathrm{Sym}(X)$.

For $g \in \operatorname{Isom} \mathcal{T}$ and $v \in X^*$, the *state* of g at v, written $g@v$, is defined inductively by $g@\epsilon = g$ and $g@(xv) = \psi(g)_1(x)@v$.

When considering subgroups $G \leq \operatorname{Isom} \mathcal{T}$, the following definitions clarify additional properties of the restriction to G of the map $\psi \colon \operatorname{Isom} \mathcal{T} \to \operatorname{Isom} \mathcal{T} \wr X$.

Definition 3 (Self-similar and recurrent groups) Consider $G \leq \operatorname{Isom} \mathcal{T}$. It is *self-similar* if ψ restricts to a map $G \to G \wr X$, and *recurrent* if furthermore for every $g \in G$, $v \in X^*$ there exists $h \in G$ with $h(v) = v$ and $h@v = g$.

In other words, G is self-similar precisely when $g@v \in G$ for all $g \in G$, $v \in X^*$, and is recurrent if $(G_v)@v = G$ for all $v \in X^*$.

The portrait of $g \in \operatorname{Isom} \mathcal{T}$ may then be defined directly by $p_g(v) := \sigma_{g@v}$. Note that if G is self-similar then its set of portraits is a "differentially closed" language, namely a language $\mathcal{L}$ such that for every $w \in \mathcal{L}$ and every $x \in X$ one has $\partial w / \partial x \in \mathcal{L}$, where $(\partial w / \partial x)(v) = w(xv)$.

Definition 4 (Level-transitive groups) A group G acting on $\mathcal{T}$ is *level-transitive* if it acts transitively on X^n for all $n \in \mathbb{N}$; so the quotient of $\mathcal{T}$ by the action is a single ray.

Definition 5 (Contracting groups) A self-similar group $G \leq \operatorname{Isom} \mathcal{T}$ is *contracting* if there exists a finite subset $F \subseteq G$ such that for every $g \in G$ the state $g@v$ belongs to F for all $v \in X^*$ except finitely many. There is a minimal such set F called the *nucleus* of G, and denoted by $\mathcal{N}$.

2.2 Branched groups and branch structure

In this section we describe one of the most important classes of groups of isometries of rooted trees: branched groups.

Definition 6 (Regularly branched groups) A group G acting on $\mathcal{T}$ is *regularly branched* (over K) if there exists a finite index subgroup $K \leq G$ with $K^X \leq \psi(K)$.

This definition can be conveniently expressed as the existence of a *branch structure*: a finite quotient Q of G, a subgroup $Q_1 \leq Q \wr X$, and epimorphisms $\pi \colon G \twoheadrightarrow Q$ and $\phi \colon Q_1 \twoheadrightarrow Q$. Indeed, with K as above and assumed without loss of generality normal, set $Q := G/K$ and $Q_1 := \psi(G)/K^X$, with $\pi \colon G \to Q$ the natural map and $\phi \colon \psi(G)/K^X \to \psi(G)/\psi(K) \cong G/K$.

3 Automata on infinite trees

In this section we define automata on infinite trees, and defer to [10] for a thorough introduction on automata on trees and tree languages.

Recall that an *alphabet* is a finite set, and a *regular tree* is a set of the form $\mathcal{T} = X^*$ for an alphabet X. Its vertices are thus finite sequences (X^*), and its boundary consists of infinite sequences (X^ω) called *rays*. For example, if $X = \{t\}$, the corresponding regular tree is geometrically an infinite half-line, with vertex set $\{t^n : n \in \mathbb{N}\}$, and it has a single ray t^ω.

A *tree word* is a map $w \colon \mathcal{T} \to A$, for a regular tree $\mathcal{T}$ and an alphabet A. For example, if $\mathcal{T} = \{t\}^*$, a tree word is a usual ω-word. As we saw in Definition 2, the portrait of a tree isometry is a tree word, with $A = \operatorname{Sym}(X)$.

A *tree language* is a set of tree words, all with the same tree and alphabet; so again if a group acts on a tree $\mathcal{T} = X^*$ then the set of portraits of its elements is a tree language with alphabet $\mathrm{Sym}(X)$.

Definition 7 (Müller automata) A *Müller tree automaton* on the tree $\mathcal{T} = X^*$ is a quintuple $\mathcal{M} = (Q, A, \delta, Q_i, Q_f)$, where

- Q is a finite (non-empty) set of *states*;
- A is an alphabet;
- $\delta \subseteq Q \times A \times Q^X$ is the *transition rule*;
- $Q_i \subseteq Q$ is the set of *initial states*;
- $Q_f \subseteq 2^Q$ is the set of *residual statesets*.

Definition 8 (Run of automata, see [17]) A *run* of a Müller tree automaton $\mathcal{M}$ is a tree word w with alphabet δ such that

- $w(\epsilon)_1 \in Q_i$;
- for every $v \in X^*$ one has $w(v)_3(x) = w(vx)_1$;
- for every ray $\xi \in X^\omega$ one has $\{w(v)_1 : v$ is a prefix of $\xi text of length that least n\} \in Q_f$ for all n large enough.

In words, the run must start at an initial state, the states match along edges according to some rules in δ, and for every ray in $\mathcal{T}$ the set of states appearing infinitely often along it is residual.

The *value* of a run w is the tree word $\overline{w}$ with alphabet A defined by $\overline{w}(v) = w(v)_2$. A tree word $\mathcal{T} \to A$ is *accepted* by $\mathcal{M}$ if it is the value of a run of $\mathcal{M}$.

Definition 9 (Regular tree languages) A tree language is *regular* if it is the set of accepted tree words of some Müller tree automaton.

We note the fundamental (and non-trivial) fact that the class of languages definable by non-deterministic Müller tree automata is closed under all Boolean operations.

4 Main results

Here we prove our main results Theorems A and B.

Theorem 1 *Let G be a level-transitive, recurrent, closed self-similar subgroup of* $\mathrm{Isom}(X^*)$. *Then G is regularly branched if and only if its set of portraits is a regular language.*

Proof In the first direction, suppose G is regularly branched, and let $(\pi : G \to Q, \phi : Q \wr X \geq Q_1 \to Q)$ be a branch structure. Consider the following Müller automaton $\mathcal{M}$:

- Q (defined above) is the set of states;
- $A = \mathrm{Sym}(X)$;
- $\delta \subseteq Q \times A \times Q^X$ with $(q, a, (q_x)) \in \delta$ whenever $((q_x), a)$ defines an element $q_1 \in Q_1$ and $\phi(q_1) = q$;
- $Q_i = Q$;
- $Q_f = 2^Q$.

Given $g \in G$, construct a run w of $\mathcal{M}$ by $w(v) = (\pi(g@v), \sigma_{g@v}, (\pi(g@vx))_{x \in X})$, and note that its value is the portrait of g. Conversely, given a run w of $\mathcal{M}$, note that its value defines a portrait and therefore an isometry t of X^*. Furthermore, the run defines an

element $w(\epsilon)_1 \in Q$ so there exists $g \in G$ such that g and t have same image in Q. Now the construction of $\mathcal{M}$ implies that for every $v \in X^*$ there exists $q_{1,v} = (w(v)_3, w(v)_2) \in Q_1$ with $\phi(q_{1,v}) = w(v)_1$; thus there exists a lift $g_1 \in G$ of $q_{1,\epsilon}$ such that g_1 and t have same image in Q_1. Continuing, there exists a subgroup $Q_2 \le Q_1 \wr X$ and an element $q_{2,v} \in Q_2$ obtained by assembling $q_{1,v}$, $(q_{1,vx})$, and a lift $g_2 \in G$ of $q_{2,\epsilon}$ such that g_2 and t have same image in Q_2. We obtain thus for all n an element $g_n \in G$ such that g_n and t have same image in Q_n. Since the Q_n converge to G and G is closed, we get $t = \lim g_n \in G$.

Note that we did not use the hypotheses that G be level-transitive nor recurrent.

In the other direction: assume that the language of portraits of the self-similar group G is a regular language, and let $\mathcal{M}$ be a Müller machine recognizing it, with stateset Q. Without loss of generality, assume $\mathcal{M}$ is *trim*: every state can be reached from an initial state, and leads to a residual set. Fix a letter $x \in X$. For every $g \in G$ and $n \in \mathbb{N}$, define

$$Q_n(g) = \{\rho(x^n)_1 : \rho \text{ is a run accepting an element } f \text{ with } f(x^n) = x^n \text{ and } f @ x^n = g\}.$$

Since G is recurrent, $Q_n(g)$ is non-empty for all $g \in G, n \in \mathbb{N}$; and $Q_{n+1}(g) \subseteq Q_n(g)$, so their intersection $Q_\infty(g)$ is also non-empty for all $g \in G$. Let K_0 denote the subgroup of G generated by $\{gh^{-1} : Q_\infty(g) \cap Q_\infty(h) \ne \emptyset\}$.

We first note that K_0 is a finite index subgroup of G; more precisely, the index of K_0 is at most $\#Q$. Indeed, let $g_0, g_1, \ldots, g_{\#Q}$ be arbitrary elements of G. Then the sets $Q_\infty(g_0), \ldots, Q_\infty(g_{\#Q})$ cannot all be disjoint, since they are non-empty and more numerous than Q. Thus there exists i and j such that $Q_\infty(g_i)$ and $Q_\infty(g_j)$ intersect, so $g_i g_j^{-1} \in K_0$, and so $g_0, \ldots, g_{\#Q}$ cannot be left coset representatives of K_0.

Consider next an element $k = gh^{-1} \in K_0$, and choose $q \in Q_\infty(g) \cap Q_\infty(h)$. In particular, there exists a run ρ_g accepting an element $f_g \in G$ and such that $\rho_g(x)_1 = q$ and $f_g(x) = x$ and $f_g @ x = g$; and similarly a run ρ_h accepting an element $f'_h \in G$ with $\rho_h(x)_1 = q$ and $f'_h(x) = x$ and $f'_h @ x = h$. Define f_h now to coincide with f_g on $X^* \setminus xX^*$, and acting as h on the subtree xX^*. The tree isometry f_h thus defined belongs to G; indeed it is accepted by the run equal to ρ_g on $X^* \setminus xX^*$ and equal to ρ_h on xX^*. Furthermore, f_g, f_h agree everywhere except on the xX^*, so $f_g f_h^{-1}$ acts as k on that subtree and trivially elsewhere. The elements f_g, f_h are accepted by runs with the same initial state q', so $Q_\infty(f_g) \cap Q_\infty(f_h) \ni q'$ is non-empty. Thus, for every generator k of K_0, we have constructed an element $\ell = f_g f_h^{-1} \in K_0$ with $\ell @ x = k$ and $\ell(x) = x$ and $\ell(v) = v$ for all $v \notin xX^*$.

We now let K denote the normal closure of K in G, all the more of finite index, and continue our reasoning with $\ell @ v = k$. Since G is recurrent and acts transitively on X, we obtain for arbitrary $y \in X$, $g \in G$ an element $\ell_{y,k} \in K$ acting as k^g on the subtree yX^* and fixing its complement. In other words, K^X is a subgroup of $\psi(K)$, and we have shown that G is regularly branched. $\qquad\square$

The following lemma will be useful for the proof of Theorem B. It states that, in a regularly branched group, a finite amount of data is sufficient to determine whether an element of $G \wr X$ belongs to $\psi(G)$.

Lemma 1 *Let G be a regularly branched group over its subgroup K, and denote by $\pi : G \to G/K$ the natural map. Let $(g_x)_{x \in X}$ and $a \in \mathrm{Sym}(X)$ be such that there exists $g \in G$ with $\psi_2(g) = a$ and $\pi(\psi_1(g)(x)) = \pi(g_x)$ for all $x \in X$.*

Then there exists an element $g' \in G$ with $\psi_2(g') = a$ and $\psi_1(g')(x) = g_x$ for all $x \in X$.

Proof By hypothesis we have $k_x := g_x \psi_1(g)(x)^{-1} \in K$ for all $x \in X$; so there exists $k \in K$ with $k @ x = k_x$. Write $g' = kg$, and compute $\psi_1(g')(x) = k_x \psi_1(g)(x) = g_x$. $\qquad\square$

Theorem 2 *If G is a contracting regularly branched group, then its set of portraits is a regular language.*

Proof We modify the construction in the proof of Theorem 1 as follows. Let $\mathcal{N}$ be the nucleus of G. As stateset, we use $Q' = Q \sqcup \mathcal{N}$. We set $Q'_i = Q'$ and $Q'_f = 2^{\mathcal{N}}$. We extend the transitions $\delta \subseteq Q \times A \times Q^X$ to $\delta' \subseteq Q' \times A \times (Q')^X$ by allowing transitions from Q' to $\mathcal{N}$. More precisely, we add to δ for all $g \in G$ the transitions $(\pi(g), \sigma_g, (g@x)_{x \in X})$ whenever $g@x \in \mathcal{N}$ for all $x \in X$, and for all $g \in \mathcal{N}$ the transitions $(g, \sigma_g, (g@x)_{x \in X})$. Thus an accepting run must reach a state in $\mathcal{N}$ on every ray, and so by König's lemma may have only finitely many states in $Q' \setminus \mathcal{N}$.

On the one hand, it is clear that every $g \in G$ is accepted, since all states $g@v$ for v sufficiently long belong to $\mathcal{N}$. Conversely, consider an accepting run. It yields by restriction a finite subtree of X^*, with elements of $\mathcal{N}$ at its leaves and elements of $Q \times A$ at its internal vertices. We construct by induction an element of G from it. If the subtree has a single vertex, then it is labelled by an element of $\mathcal{N}$ and therefore directly produces an element of G. Otherwise, consider an internal vertex of maximal height, so all its descendants are labelled by elements of G. We may replace the internal vertex's label by an element of G, and proceed by induction, owing to Lemma 1. □

5 Applications

In this last section, we list some consequences of Theorems A and B. These all pertain to decidability questions.

When given a contracting regularly branched group, the data defining it are assumed to be in the form of a finite generating set S, the restriction of ψ to S, and an oracle promising that the action is faithful, contracting and regularly branched.

Corollary 1 *Let $G, H \leq \text{Isom } \mathcal{T}$ be two finitely generated contracting regularly branched groups. Then there exists an algorithm that determines whether $G \leq H$. As a consequence, it is decidable if $G = H$.*

Proof Suppose that there exists an oracle that recognises whether a group is contracting and branched. Then one can algorithmically determine its nucleus and its branching subgroup. Now given two contracting regular branched groups G and H, by Theorem 2 their languages of portraits are regular, and the inclusion in regular languages is decidable. This completes the proof. □

On the other hand, a closed regularly branched group is determined by the finite data of its branched structure.

Corollary 2 *Let $G \leq \text{Isom } \mathcal{T}$ be a finitely generated, closed self-similar regularly branched group. Then the Hausdorff dimension of G is computable.*

Proof By Theorem 1, there exists a tree automaton that recognizes elements of the portrait of a closed group G. The Hausdorff dimension of G is the entropy of its language of portraits, so is computable from the automaton.

More directly, a linear system of equations can easily be set up, one variable per state of the automaton, whose solution yields the Hausdorff dimension of G, see [23]. □

We finally to turn to the orbit structure of contracting, regularly branched groups. Say the group G acts on the regular tree $\mathcal{T} = X^*$; so it acts on its boundary $\partial\mathcal{T} = X^\omega$. The *orbit relation* is the subset

$$\mathcal{O} = \{(\xi, \eta) \in X^\omega \times X^\omega : \eta \in G\xi\} \subseteq (X \times X)^\omega.$$

Corollary 3 *Let $G \leq \mathrm{Isom}\,\mathcal{T}$ be a finitely generated, contracting, regularly branched group. Then the equivalence relation $\mathcal{O}$ is an ω-regular language, and is computable.*

Proof Let $\mathcal{M}$ be a Müller tree automaton recognizing the portraits of G. Define a (classical) Müller automaton $\mathcal{M}'$ with same stateset Q, same initial and residual states Q_f, Q_i, alphabet $A' = X \times X$, and transitions $\delta' \subset Q \times A' \times Q$, by

$$\delta' = \{(q, (y, z), q_y) : \exists (q, a, (q_x)_{x \in X}) \in \delta \text{ with } a(y) = z\}.$$

Then every run of $\mathcal{M}$, recognizing an element $g \in G$ leads to runs, for all $\xi \in X^\omega$, of $\mathcal{M}'$ recognizing $(\xi, g\xi)$; thus the language of $\mathcal{M}'$ is $\mathcal{O}$.

In particular, for every preperiodic ray ξ, the orbit of ξ is computable as an ω-regular language $\subseteq X^\omega$, by intersecting $\mathcal{O}$ with the ω-regular language $\{\xi\} \times X^\omega$ and projecting to the second coordinate. □

References

1. Abercrombie, A.G.: Subgroups and subrings of profinite rings. Math. Proc. Camb. Philos. Soc. **116**(2), 209–222 (1994). https://doi.org/10.1017/S0305004100072522
2. Anīsīmov, A.V.: The group languages. Kibernetika (Kiev) **4**, 18–24 (1971)
3. Bartholdi, L.: Branch rings, thinned rings, tree enveloping rings. Israel J. Math. **154**, 200410 (2006)
4. Bartholdi, L., Siegenthaler, O.: The twisted twin of the Grigorchuk group. Int. J. Algebra Comput. **20**(04), 465–488 (2010)
5. Bartholdi, L., Grigorchuk, R.I., Sunik, Z.: Branch groups. In: Handbook of Algebra, vol. 3, pp. 989–1112. North-Holland, Amsterdam (2003)
6. Baumslag, G.: Topics in Combinatorial Group Theory. Lectures in mathematics. ETH Zürich, Zürich (1993)
7. Bondarenko, I. V., Samoilovych, I. O.: On finite generation of self-similar groups of finite type. Int. J. Algebra Comput. **23**(1), 69–79 (2013). https://doi.org/10.1142/S0218196712500786
8. Boudet, A., Comon, H.: Diophantine equations, Presburger arithmetic and finite automata. In: Proceedings of the 21st International Colloquium on Trees in Algebra and Programming, CAAP '96, pp. 30–43. Springer, Berlin (1996)
9. Büchi, J.R.: Weak second-order arithmetic and finite automata. Z. Math. Log. Grundl. Math. **6**, 66–92 (1960). https://doi.org/10.1002/malq.19600060105
10. Comon, H., Dauchet, M., Gilleron, R., Jacquemard, F., Lugiez, D., Löding, C., Tison, S., Tommasi, M.: Tree automata techniques and applications (2008). https://hal.inria.fr/hal-03367725

11. de la Harpe, P.: Topics in Geometric Group Theory, Chicago Lectures in Mathematics. University of Chicago Press (2000). https://books.google.es/books?id=cRT01C5ADroC
12. Epstein, D.B.A., Paterson, M.S., Cannon, J.W., Holt, D.F., Levy, S.V., Thurston, W.P.: Word Processing in Groups. A. K. Peters Ltd, USA (1992)
13. Grigorčuk, R.I.: On Burnside's problem on periodic groups. Funkt. Anal. i Prilozhen. **14**(1), 53–54 (1980)
14. Grigorchuk, R.: Solved and unsolved problems around one group. Infinite Groups Geometr. Combin. Dyn. Aspects **248**, 117–218 (2006)
15. Hopcroft, J.E., Motwani, R., Ullman, J.D.: Introduction to Automata Theory, Languages, and Computation. Addison-Wesley Longman Publishing Co., Inc, USA (2006)
16. Kharlampovich, O., Khoussainov, B., Miasnikov, A.: From automatic structures to automatic groups. Groups Geom. Dyn. **8**, 157–198 (2011)
17. Muller, D.E.: Infinite sequences and finite machines. In: Proceedings of the 1963 Proceedings of the Fourth Annual Symposium on Switching Circuit Theory and Logical Design. IEEE Computer Society, USA, pp. 3–16 (1963). https://doi.org/10.1109/SWCT.1963.8
18. Noce, M., Thillaisundaram, A.: Hausdorff dimension of the second Grigorchuk group. Int. J. Algebra Comput. **31**(6), 1037–1047 (2021). https://doi.org/10.1142/S0218196721400038
19. Penland, A., Šunić, Z.: Finitely constrained groups of maximal Hausdorff dimension. J. Aust. Math. Soc. **100**(1), 108–123 (2016). https://doi.org/10.1017/S144678871500018X
20. Penland, A., Šunić, Z.: A language hierarchy and Kitchens-type theorem for self-similar groups. J. Algebra **537**, 173–196 (2019). https://doi.org/10.1016/j.jalgebra.2019.07.017
21. Rabin, M.O.: Decidability of second-order theories and automata on infinite trees. Bull. Am. Math. Soc. **74**, 1025–1029 (1968)
22. Siegenthaler, O.: Hausdorff dimension of some groups acting on the binary tree. J. Group Theory **11**(4), 555–567 (2008). https://doi.org/10.1515/JGT.2008.034
23. Siegenthaler, O.: Discrete and profinite groups acting on regular rooted trees (2010)
24. Siegenthaler, O., Zugadi-Reizabal, A.: The equations satisfied by GGS-groups and the abelian group structure of the Gupta-Sidki group. Eur. J. Combin. **33**(7), 1672–1690 (2012). https://doi.org/10.1016/j.ejc.2012.03.025
25. Šunić, Z.: Hausdorff dimension in a family of self-similar groups. Geometr. Ded. **124**(200706), 213–236 (2007)
26. Šunić, Z.: Pattern closure of groups of tree automorphisms. Bull. Math. Sci. **1**(1), 115–127 (2011). https://doi.org/10.1007/s13373-011-0007-2
27. Šunić, Z., Uria-Albizuri, J.: Portrait growth in contracting, regular branch groups. J. Algebra Appl. (2022). https://doi.org/10.1142/S0219498823501803

🖄 Springer