www.theoryofgroups.ir

International Journal of Group Theory
ISSN (print): 2251-7650, ISSN (on-line): 2251-7669
Vol. x No. x (201x), pp. xx-xx.

www.ui.ac.ir

CATALAN FRAGILE WORDS

DANIELE D'ANGELI, ALFREDO DONNO*, AND EMANUELE RODARO

Communicated by

Abstract. Fragile words have been already considered in the context of automata groups. Here we focus our attention on a special class of strongly fragile words that we call Catalan fragile words. Among other properties, we show that there exists a one-to-one correspondence between the set of Catalan fragile words and the set of full binary trees.

1. Introduction

In this paper we are interested in the description of the so called strongly fragile words that were defined in [4] in the context of automata groups, more precisely about the possibility of generating free groups by means of automata. Examples of automata generating non abelian free groups were constructed by Alëshin [1]. M. Vorobets and Y. Vorobets first proved the correctness of Alëshin's examples and then generalized them to a class of bireversible automata generating free non-abelian groups (see [11, 12, 10]). Fragile words appear in the context of finding examples of non-bireversible automata generating free groups. More precisely, in the rather broad class of (co-accessible) invertible transducers having a sink-state (shortly sink), one wonders if there exists the possibility of generating a non abelian free group. This problem has a simple solution, but for the transitive case it appears more difficult [4] (see also, [3, 7, 8] for more details on automata groups). In this setting (strongly) fragile words correspond to non-trivial relations of minimal length in the group. Although the motivation for the introduction of such words comes from Group Theory, and more specifically from the theory of automata groups, fragile words seem to have interesting combinatorial properties in their own right. Behind their straightforward description, there is no characterization of the language that they generate (see, for instance, [5], for more details on languages). Here, we focus our attention on

MSC(2010): Primary: 20F10; Secondary: 20F05, 68Q70.

Keywords: Automata group, Catalan number, Fragile word, Full binary tree.

Received: dd mmmm yyyy, Accepted: dd mmmm yyyy.

*Corresponding author.

a particular class of strongly fragile words, that comes up to have a nice correspondence with the famous Catalan numbers [9].

2. **Fragile words and rooted trees**

The easiest description of fragile words can be given as follows: let Q be a finite set and let $\widetilde{Q}^* = (Q \cup Q^{-1})^*$ be the associated free monoid, i.e., the set of all reduced finite words in the alphabet $Q \cup Q^{-1}$. We denote by 1 the empty word. Let F_Q be the free group generated by Q. Given $q \in Q$, we denote by ϵ_q the endomorphism of $\widetilde{Q}^*$ that erases the occurrences of q and q^{-1} from $w \in \widetilde{Q}^*$. Moreover, $\overline{w}$ denotes the unique reduced representative in F_Q corresponding to w. A non-trivial word $w \in \widetilde{Q}^*$ is called *fragile* if there exists $q \in Q$ such that $\overline{\epsilon_q(w)} = 1$. A non-trivial word $w \in \widetilde{Q}^*$ is called *strongly fragile* if, for every $q \in Q$, one has $\overline{\epsilon_q(w)} = 1$. A natural question to ask is about the structure of the set of (strongly) fragile words.

Example 2.1. Let $Q = \{a, b, c\}$ and $w = a^{-2}bb^{-1}a^3ba^{-2}ab^{-2}$. Then we have $\overline{w} = aba^{-1}b^{-2}$. Moreover w is fragile since $\overline{\epsilon_b(w)} = 1$ but not strongly fragile since $\overline{\epsilon_a(w)} = b^{-1} \neq 1$ and $\overline{\epsilon_c(w)} = \overline{w} \neq 1$.

In this paper we describe a method to generate a family of strongly fragile words by using rooted trees. This description makes a bridge between strongly fragile words and Catalan numbers, and it leads to the definition of Catalan fragile words.

2.1. **A group theoretical interpretation.**

In this section we give a deeper algebraic motivation for the study of (strongly) fragile words. Let us start by recalling the definition of automaton (see, for instance, [6]).

A (finite) *automaton* (or a transducer) is a quadruple $\mathcal{A} = (Q, A, \mu, \lambda)$, where:

- Q is a finite set, called the set of *states*;
- A is a finite set, called the *alphabet*;
- $\mu : Q \times A \to Q$ is the transition map or *restriction*;
- $\lambda : Q \times A \to A$ is the output map or *action*.

A very convenient way to represent an automaton is through its Moore diagram: this is a directed labelled graph whose vertices are identified with the states of $\mathcal{A}$ (See Fig. 1, where $Q = \{a, b, c, d, e\}$ and $A = \{0, 1\}$). For every state $q \in Q$ and every letter $a \in A$, the diagram has an arrow from q to $\mu(q, a)$ labelled by $a|\lambda(q, a)$. The automaton $\mathcal{A}$ contains a *sink* $e \in Q$, if one has $\lambda(e, a) = a$ and $\mu(e, a) = e$ for any $a \in A$.

The automaton $\mathcal{A}$ is said to be *invertible* if, for all $q \in Q$, the transformation $\lambda(q, \cdot) : A \to A$ is a permutation of A. If $\mathcal{A}$ is invertible, one can define the maps λ and μ on Q^{-1}, the set of formal inverses of Q, by exchanging the input and the output in the automaton, i.e., for every state $q^{-1} \in Q^{-1}$ and every letter $a \in A$, the diagram has an arrow from q^{-1} to $(\mu(q, a))^{-1}$ labelled by $\lambda(q, a)|a$.

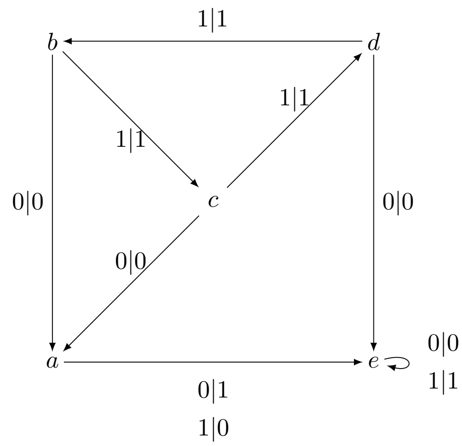

FIGURE 1. The automaton generating the Grigorchuk group.

Finally, the maps λ and μ can be naturally extended to $Q^* \times A^*$ by using the following recursive rules:

$$\lambda(q, aw) = \lambda(q, a)\lambda(\mu(q, a), w), \quad \lambda(qq', a) = \lambda(q, \lambda(q', a)) \tag{2.1}$$

$$\mu(q, aw) = \mu(\mu(q, a), w), \quad \mu(qq', a) = \mu(q, \lambda(q', a))\mu(q', a)$$

for all $q \in Q$, $q' \in Q^*$, and $a \in A$, $w \in A^*$. In this paper, when the automaton has a sink, it is supposed to be reachable from any state, i.e., for every $q \in Q$ there is $u \in A^*$ such that $\mu(q, u) = e$.

Definition 2.2. *The automaton group $G = G(\mathcal{A})$ is the permutation group of A^* generated by the states $\{q : q \in Q\}$ with the operation defined by (2.1), where two elements $g, h \in \widetilde{Q}^*$ represent the same element in G if*

$$\lambda(g, u) = \lambda(h, u)$$

for every $u \in A^*$.

The following proposition shows the crucial role that fragile words play in automata group theory (see [4] for more details). The idea is to extend the alphabet A to $A \cup Q$ and add a sink in case the automaton does not have any (we write $Q \cup \{e\}$). The maps λ and μ keep the original action on A and they are defined on $Q \cup \{e\}$ as follows

$$\lambda(q, p) = p \quad \forall q, p \in Q \cup \{e\},$$

$$\forall q \in Q \quad \mu(q, q) = e \quad \text{and} \quad \mu(q, p) = q \ \forall p \in (Q \cup \{e\}) \setminus \{q\}. \tag{2.2}$$

If the automaton $\mathcal{A}$ is invertible, the new automaton $\mathcal{A}'$ that we get with this construction is still invertible. Moreover, given a subset $Q' \subseteq Q$ and a word $w \in \widetilde{Q'}^*$, we say that w is strongly fragile with respect to Q' if $\overline{\epsilon_q(w)} = 1$ for every $q \in Q'$.

Proposition 2.3. *If $G(\mathcal{A}')$ is not free, then there exist $Q' \subseteq Q$ and a shortest non-trivial relation $w \in \widetilde{Q'}^*$ such that w is a strongly fragile word with respect to Q'.*

Proof. First of all we claim that, if w is a relation of an automata group $G(\mathcal{A})$ then $\mu(w,u)$ is also a relation for every $u \in A^*$. In fact $\lambda(w,v) = v$ for every $v \in A^*$ and so, in particular

$$\lambda(w, uv) = u\lambda(\mu(w,u), v) = uv.$$

This implies $\lambda(\mu(w,u), v) = v$ for every $v \in A^*$, and proves the claim. Now in the group $G(\mathcal{A}')$ take any shortest non-trivial relation $w \in \widetilde{Q'}^*$, for some Q'. By Equation (2.2) we have

$$\mu(w,q) = \epsilon_q(w),$$

and if q or q^{-1} appears in w, the length of $\mu(w,q)$ is shorter than the length of w and so $\overline{\epsilon_q(w)} = 1$ for any $q \in Q'$ because of the minimality of w. □

From Proposition 2.3 we can deduce that a relation w of minimal length in $G(\mathcal{A}')$ is a strongly fragile word that is also a relation in $G(\mathcal{A})$. For this reason, the previous construction has been used in [4] to construct examples of non-bireversible free automata groups.

2.2. **Catalan numbers.** Catalan numbers owe their name to the mathematician E. C. Catalan, who worked in the nineteenth century in several areas of Mathematics, including Combinatorics, Continued Fractions, Number Theory. Catalan showed in 1838 that the n-th Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

counts the number of ways to fully parenthesize a string of $n+1$ letters. The first values of this sequence are

$$C_0 = 1; C_1 = 1; C_2 = 2; C_3 = 5; C_4 = 14; C_5 = 42; C_6 = 132; \ldots.$$

The sequence of Catalan numbers has many other interesting properties and applications, for instance, in counting monotonic paths in an $n \times n$ grid that do not rise above the diagonal or in counting the number of chords in a given polygon satisfying some prescribed intersection conditions (the reader can refer to the monograph [9] for more details and properties).

In this paper, we are interested in the application of Catalan numbers concerning the enumerations of some families of trees. We briefly recall the definition of rooted tree, binary tree and full binary tree.

Rooted trees are trees where a special vertex, called the *root*, has been chosen. Rooted trees are usually represented as in Fig. 2, where R denotes the root, the vertex u is said to be the *parent* of the vertex v, the vertex v is said to be a *child* of the vertex u, and w is said to be a *descendant* of the vertices u and v (observe that each vertex different from the root is a descendant of the root).

In a rooted tree, the degree of the root is the number of its children, and for any other vertex different from the root the degree is equal to the number of its children +1. In particular, a vertex with no children is called a *leaf*, and so each leaf has degree 1. We say that the level i of a rooted tree consists of all vertices having geodesic distance i from the root. In Fig. 2, the root is the unique vertex at level 0; the level 1 contains 3 vertices, the level 2 contains 5 vertices, and the level 3 contains 2 vertices, and we say that the rooted tree has depth 3.

Finally, if in a rooted tree we consider a vertex which is not a leaf, that is, it has at least one descendant, then the subgraph consisting of a single child of that vertex, together with all the descendants of that child, is call a *subtree* of the whole tree.

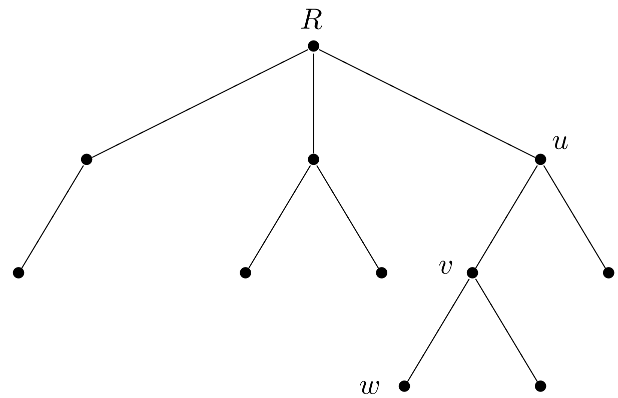

FIGURE 2. An example of a rooted tree.

A *binary tree* is by definition a rooted tree for which every vertex has at most 2 children. In this binary case, one distinguishes between *left* and *right* subtrees. The set of all binary trees with 3 vertices is represented in Fig. 3.

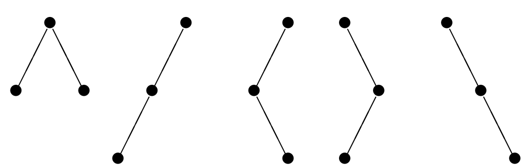

FIGURE 3. The five binary trees with 3 vertices.

Finally, a binary tree where each vertex which is not a leaf has exactly two children is said to be a *full binary tree*. A full binary tree on 3 vertices is called *caret*. Two examples of full binary trees are depicted in Fig. 4.

The proof of the following proposition can be found, for instance, in [9].

Proposition 2.4. *The Catalan number C_n counts the number of binary trees with n vertices, as well as the number of full binary trees with $2n+1$ vertices.*

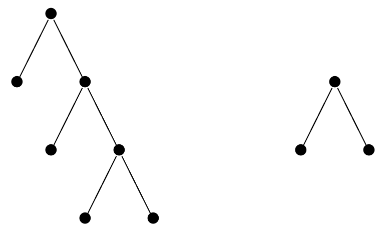

FIGURE 4. A full binary trees with 7 vertices and a caret.

3. An explicit construction of strongly fragile words

Observe that, from what said above, a full binary tree is a graph without cycles where there is a distinguished vertex with degree 2 (the root), a set of vertices with degree 1 (the leaves) and any other vertex has degree 3.

Lemma 3.1. *Let T be a full binary tree with n leaves. Then T has $2n-1$ vertices.*

Proof. Assume that T has n leaves, and let m be the number of the remaining vertices: notice that each leaf has degree 1, whereas $m-1$ vertices have degree 3, and only one vertex (the root) has degree 2. Then the Handshaking Lemma (e.g. [2]) gives:

$$n + 3(m-1) + 2 = 2(n+m-1),$$

so that $m = n-1$ and the total number of vertices is $n + m = 2n-1$. □

Notice that, by virtue of Lemma 3.1 and Proposition 2.4, the number of rooted binary trees with n leaves is exactly C_{n-1}.

In the sequel we want to construct strongly fragile words over an alphabet Q starting from full binary trees with $|Q|$ leaves.

With every full binary binary tree T on $|Q|$ leaves, we associate a strongly fragile word $u(T) \in \widetilde{Q}^*$ defined recursively by an iterated caret reduction process as follows:

- choose a labeling bijection between the set of leaves and Q;
- if q_i and q_j belong to the same caret, remove it from the tree and get a new tree T';
- label the new produced leaf (the root of the removed caret) by $[q_i, q_j] := q_i^{-1}q_j^{-1}q_iq_j$;
- proceed in the same way with the tree T'.

After a finite number of steps, we end up with the empty tree (consisting of just one vertex) and a word which is an iterated nesting of commutators. We denote by D_T the word constructed from T by using the method described above.

Example 3.2. Consider the tree on the left hand side of Fig. 4 and label q_1, q_2, q_3 and q_4 the four leaves (from left to right). After the first step of the above process, we remove the rightmost caret and get three leaves, labelled by q_1, q_2 and $[q_3, q_4]$. After the second step, we get just two leaves labelled by q_1 and $[q_2, [q_3, q_4]]$. Finally, we have

$$D_T = [q_1, [q_2, [q_3, q_4]]] = q_1^{-1} q_4^{-1} q_3^{-1} q_4 q_3 q_2^{-1} q_3^{-1} q_4^{-1} q_3 q_4 q_2 q_1 q_2^{-1} q_4^{-1} q_3^{-1} q_4 q_3 q_2 q_3^{-1} q_4^{-1} q_3 q_4.$$

Another way to get the same word is by starting from the top instead that from the bottom of the tree: let T be a full binary tree containing $|Q|$ leaves and fix a labeling bijection between the leaves and Q. Let T_1, T_2 be the two subtrees rooted at the two vertices of the first level of T; then define

$$u(T) = [u(T_1), u(T_2)],$$

where $u(T_k) = q$ when T_k reduces to the leaf labelled by $q \in Q$. We denote by U_T the word constructed from T by using this second recursive method.

The two methods give rise to the same word, as the following lemma shows.

Lemma 3.3. *Let T be a full binary tree on $|Q|$ leaves. Then $D_T = U_T$.*

Proof. By induction on the depth (number of levels) $d(T)$ of the tree T. If $d(T) = 1$, then $Q = \{q_1, q_2\}$ and the two definitions coincide, giving rise to the commutator $[q_1, q_2]$. Suppose now $d(T) > 1$. Then, at least one of the subtrees T_1, T_2 rooted at the first level does not reduce to a leaf. However, the depth of T_1 and T_2 is at most $d(T) - 1$, so that we can apply the induction. We get $u = D_{T_1} = U_{T_1}$ and $v = D_{T_2} = U_{T_2}$. This implies $D_T = U_T = [u, v]$. □

We call such kind of words *Catalan fragile words* for a reason that will be clear later. The following lemma shows that Catalan fragile words are actually strongly fragile.

Lemma 3.4. *Let T be a full binary binary tree on Q leaves, with $|Q| \geq 2$. Then the Catalan fragile word $u(T)$ is strongly fragile (on Q).*

Proof. Let us prove the statement by induction on the depth $d(T)$ of the tree T. If $d(T) = 1$, then $u(T) = [v, v']$, where v, v' are the leaves of T, that is clearly a strongly fragile word on the alphabet $\{v, v'\}$. Now suppose that $u(T) = [u(T_1), u(T_2)]$, where at least one of the subtrees T_1, T_2 has depth at least 1. We have two cases. If only one among T_1, T_2 has depth at least 1, then we can suppose without loss of generality that T_1 has this property, and T_2 consists of just one vertex v. Therefore

$$u(T) = u(T_1) v u(T_1)^{-1} v^{-1}$$

that is strongly fragile on the leaves of T by induction, since the claim is true for T_1. Otherwise, if both T_1 and T_2 have depth at least 1, one has:

$$u(T) = u(T_1) u(T_2) u(T_1)^{-1} u(T_2)^{-1}$$

where, by induction hypothesis, $u(T_1)$, $u(T_2)$ are strongly fragile. Hence, it is easy to check that also $u(T)$ is strongly fragile on the set of leaves of T. □

In what follows, we denote by $\lambda(T)$ the number of leaves of the full binary tree T and by $|u(T)|$ the length of the strongly fragile word $u(T)$. Moreover, we put:

$$\varphi(m) = \min\{|u(T)| : T \text{ is a full binary tree on } m \text{ leaves}\}.$$

The following lemma holds.

Lemma 3.5. *One has* $\varphi(m) = 2\min_{n+k=m}\{\varphi(n)+\varphi(k)\}$.

Proof. We have:

$$\begin{aligned}\varphi(m) &= \min\{|[u(T_1), u(T_2)]| : \ \lambda(T_1) = n, \lambda(T_2) = k \text{ and } k+n = m\}\\ &= 2\min_{n+k=m}\{|u(T_1)| + |u(T_2)| : \lambda(T_1) = n, \lambda(T_2) = k\})\\ &= 2\min_{n+k=m}\{\varphi(n)+\varphi(k)\}.\end{aligned}$$

□

The following theorem provides an upper bound for $\varphi(m)$.

Theorem 3.6. *We have* $\varphi(m) \leq m^2$*. Moreover, the bound is sharp.*

Proof. Let us prove the statement by induction on m. For $m = 2$, we have $\varphi(2) = 4$. Now, by Lemma 3.5, we have:

$$\varphi(m) = 2\min_{n+k=m}\{\varphi(n)+\varphi(k)\} \leq 2\min_{n+k=m}\{n^2+k^2\}.$$

Minimize $n^2 + k^2$ subject to the condition $n + k = m$ is equivalent to minimize $2n^2 - 2mn + m^2$. A direct computation shows that such a minimum is reached for $n = k = \frac{m}{2}$ and it is equal to $\frac{m^2}{2}$, which gives $\varphi(m) \leq m^2$. Moreover, the previous bound is sharp. Indeed, by taking $m = 2^h$, we may consider the "balance" binary tree B, that is, the tree with 2^h leaves and depth equal to h, in which the simple path connecting the root with any leaf has length h. In this case, by a direct computation by induction on h, one gets $|u(B)| = m^2$. □

In the next theorem we provide an explicit computation of $|u(T)|$ in terms of the length of the paths connecting each leaf of T to the root.

Let $Q = \{q_1, \ldots, q_m\}$. For each $i = 1, \ldots, m$, we denote by d_i the length of the path connecting the root of the full binary tree to its leaf identified with the letter q_i.

Theorem 3.7. *Let T be a full binary tree with m leaves. Then*

$$|u(T)| = \sum_{i=1}^{m} 2^{d_i}.$$

Proof. By induction on the maximal distance d_i. If $d_1 = d_2 = 1$ we have $u(T) = 4$. Otherwise, let $q_1, \ldots, q_n$ be the leaves of the subtree T_1 rooted at the left vertex of the first level of T, and let $q_{n+1}, \ldots, q_m$ be the leaves of the subtree T_2 rooted at the right vertex of the first level. We have

$$|u(T)| = 2(|u(T_1)| + |u(T_2)|) = 2\left(\sum_{i=1}^{n} 2^{d_i-1} + \sum_{i=n+1}^{m} 2^{d_i-1}\right).$$

The last term is exactly $\sum_{i=1}^{m} 2^{d_i}$. □

Corollary 3.8. *Let T be a full binary tree with m leaves. Then*

$$|u(T)| \leq 2^m + \sum_{i=1}^{m-2} 2^i.$$

In particular $|u(T)| \leq |u(T')|$, where T' is the full binary tree with m leaves, with exactly two leaves at depth $m-1$ and one leaf at each depth i, with $i = 1, \ldots, m-2$.

Proof. Consider the full binary tree with m leaves, having two leaves at depth $m-1$ and one leaf at each depth i, with $i = 1, \ldots, m-2$ (the tree on the left hand side of Fig. 4 is an example for $m = 4$). It follows from Theorem 3.7 that the associated strongly fragile word has length

$$l = 2^m + \sum_{i=1}^{m-2} 2^i.$$

We claim that, for any full binary tree T with m leaves, one has $|u(T)| \leq l$. We proceed by induction on the depth of T. Notice that, if one of the two subtrees T_1 and T_2 rooted at the first level reduces to a single vertex, the result easily follows. Hence, we assume that their depth is greater or equal to 1. Suppose that T has m leaves, with n leaves belonging to T_1 and k belonging to T_2, where $m = n + k$. By induction, we have:

$$|u(T_1)| \leq 2^n + \sum_{j=1}^{n-2} 2^j \qquad |u(T_2)| \leq 2^k + \sum_{j=1}^{k-2} 2^j.$$

This implies

$$|u(T)| = 2(|u(T_1)| + |u(T_2)|) \leq 2^{n+1} + 2^{k+1} + \sum_{j=1}^{n-2} 2^{j+1} + \sum_{j=1}^{k-2} 2^{j+1}.$$

We need to show, that

$$2^{n+1} + 2^{k+1} + \sum_{j=1}^{n-2} 2^{j+1} + \sum_{j=1}^{k-2} 2^{j+1} \leq l = 2^{n+k} + \sum_{j=1}^{n+k-2} 2^j.$$

Suppose, without loss of generality, that $n \geq k$. We can rewrite the left side of the inequality as follows

$$2^{n+1} + 2^{k+1} + \left(\sum_{i=2}^{k-1} (2^{n-i+1} + 2^{k-i+1}) + \sum_{i=k}^{n-1} 2^{n-i+1} \right).$$

We have, for $m, n \geq 2$

$$2^{n+1} + 2^{k+1} \leq 2^{n+k},$$

in fact the function $f(x, y) = 2^{x+y} - 2^{x+1} - 2^{y+1}$ on the domain $D = \{(x, y) : \ x, y \geq 2\}$ has minimum in $(x, y) = (2, 2)$ and $f(2, 2) = 0$. Analogously, one can check that

$$2^{n-i+1} + 2^{k-i+1} \leq 2^{n+k-i}, \text{ for each } i = 2, \ldots, k-1,$$

and

$$2^{n-i+1} \leq 2^{n+k-i}, \text{ for each } i = k, \ldots, n-1.$$

The sum over the right hand sides of the previous inequalities is smaller or equal to l (we only have the terms with index j from n to $n + k - 2$). This concludes the proof. □

Let T be a full binary tree, and let v be a leaf of T; if H is another full binary tree, we denote by $T \circ_v H$ the new full binary tree obtained by attaching the root of H at the leaf v of T. The inverse of this operation is the deleting of a subtree: if v is a vertex of the full binary tree T and it is not a leaf of T, then T_v denotes the full binary tree obtained from T by deleting the subtree rooted at v; in this way, v becomes a leaf of the resulting tree T_v.

We want to show that, with each Catalan fragile word, we may uniquely associate a full binary tree. We first need the following lemma.

Lemma 3.9. *A commutator $[a, b]$ is a factor of a Catalan fragile word $u(T) \in \widetilde{Q}^*$ if and only if there is a vertex v of T such that a, b are two leaves that are children of v. Moreover, if we substitute each occurrence of $[a, b]$ or $[a, b]^{-1}$ of $u(T)$ with the symbol v or v^{-1}, respectively, we obtain a word $w \in H^*$ with $H = Q \setminus \{a, b\} \cup \{v\}$ and such that $u(T_v) = w$.*

Proof. Since $u(T) = [u(T_1), u(T_2)]$, where T_1, T_2 are the two subtrees rooted at the two children of the root of T, we claim that $[a, b]$ is a factor of $u(T)$ if and only if $[a, b]$ is a factor of either $u(T_1)$ or $u(T_2)$. Indeed, this is a consequence of the fact that $u(T_1) \in \widetilde{Q}_1^*$ and $u(T_2) \in \widetilde{Q}_2^*$, where Q_1 is the set of leaves of the left subtree T_1, and Q_2 is the set of leaves of the right subtree T_2, so that $Q_1 \cap Q_2 = \emptyset$ and $Q_1 \cup Q_2$ is the set of leaves of the whole tree T.

Thus, we cannot have that $[a, b] = ww'$, where w is a proper suffix of $u(T_1)$ and w' is a proper prefix of $u(T_2)$ (or w is a proper suffix of $u(T_1)^{-1}$ and w' is a proper prefix of $u(T_2)^{-1}$, or the last case where w is a proper suffix of $u(T_2)^{-1}$ and w' is a proper prefix of $u(T_1)$). Now, the first statement of the lemma follows by induction on the depth $d(T)$. Let us prove the second statement of the lemma. Note that, by definition of $u(T)$, the letters a, b appear in $u(T)$ just in the commutator $[a, b]$. Therefore, if we consider T_v, then the associated $u(T_v)$ belongs to H^* with $H = Q \setminus \{a, b\} \cup \{v\}$. Furthermore, by definition of $u(T)$, it is easy to see that $u(T)$ is obtained by substituting each occurrence v, v^{-1} in $u(T_v)$ with $[a, b], [a, b]^{-1}$, respectively. This is equivalent to the fact that the word w in the statement is exactly $u(T_v)$. □

If a, b are two leaves of T having the same father v, we denote by (a, v, b) the caret rooted at v and having a, b as children.

Theorem 3.10. *If $u(T) = u(T')$, then $T = T'$.*

Proof. Let us prove the statement by induction on the cardinality of the set Q such that $u(T) \in \widetilde{Q}^*$. If $Q = \{a, b\}$, then the statement is clear since $u(T) = [a, b] = u(T')$ and $T = (a, r, b) = T'$, where r is the root of T. Suppose $u(T) \in \widetilde{Q}^*$, with $|Q| > 2$. By Lemma 3.9 there exist $a, b \in Q$ such that $[a, b]$ is a factor of $u(T)$, and if we substitute the occurrences of $[a, b], [a, b]^{-1}$ with v, v^{-1}, respectively, then we obtain a new word $w = u(T_v) = u(T'_v) \in Q \setminus \{a, b\} \cup \{v\}$. By the induction hypothesis, we get $T_v = T'_v$. Hence, $T = T_v \circ_v (a, v, b) = T'_v \circ_v (a, v, b) = T'$. □

From the previous result we deduce the following corollary.

Corollary 3.11. *The Catalan fragile words on an alphabet Q are in one-to-one correspondence with the full binary trees with set of leaves Q. In particular, if $|Q| = m$, there exist exactly C_{m-1} Catalan fragile words on Q, where C_n is the n-th Catalan number.*

The problem of giving an explicit and complete characterization of all strongly fragile words is still open. Moreover, from an algebraic point of view, one can take the set Q and the set of all Catalan fragile words on Q. There is no explicit description of the group freely generated by the elements in Q whose relations are generated by the Catalan fragile words on Q. Notice that, since Catalan fragile words are commutators, this group can be mapped onto the free abelian group on $|Q|$ generators.

Acknowledgments

The first author acknowledges support from the FWF Project P 29355-N35.

References

[1] S.V. Alëshin, A free group of finite automata, *Mosc. Univ. Math. Bull.*, **38** (1983), 10–13.

[2] B. Bollobás, Modern graph theory. Graduate Texts in Mathematics, **184**. *Springer-Verlag, New York*, 1998.

[3] L. Bartholdi, R. Grigorchuk, and V. Nekrashevych, From fractal groups to fractal sets. *Fractals in Graz 2001*, 25–118, Trends Math., Birkhäuser Basel, 2003.

[4] D. D'Angeli and E. Rodaro, Fragile words and Cayley type transducers, *Int. J. Group Theory*, **7** no. 3 (2018), 91–109.

[5] S. Eilenberg, *Automata, Languages, and Machines*. Vol. A Pure and Applied Mathematics, Vol. 58 *Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York*, 1974.

[6] R. Grigorchuk, V. Nekrashevych, and V. Sushchanskii, Automata, dynamical systems and infinite groups, *Proc. Steklov Inst. Math.*, **231** (2000), 134–214.

[7] R. Grigorchuk, Some topics of dynamics of group actions on rooted trees, *Proc. Steklov Inst. Math.*, **273** no. 1 (2011), 64–175.

[8] V. Nekrashevych, Self-similar groups, *Mathematical Surveys and Monographs, 117. American Mathematical Society, Providence, RI*, 2005.

[9] S. Roman, An introduction to Catalan numbers, *Compact textbooks in Mathematics, Birkhauser, Basel*, 2005.

[10] B. Steinberg, M. Vorobets, and Y. Vorobets, Automata over a binary alphabet generating free groups of even rank, *Int. J. Algebra Comput.*, **21** 1-2 (2011), 329–354.

[11] M. Vorobets and Y. Vorobets, On a free group of transformations defined by an automaton, *Geom. Dedicata*, **124** (2007), 237–249.

[12] M. Vorobets and Y. Vorobets, On a series of finite automata defining free transformation groups, *Groups Geom. Dyn.*, **4** no. 2 (2010), 377–405.

Daniele D'Angeli

Institut für Diskrete Mathematik, Technische Universität Graz, Steyrergasse 30, 8010, Graz, Austria

Email: `dangeli@math.tugraz.at`

Alfredo Donno

Dipartimento di Ingegneria, Università degli Studi Niccolò Cusano, Via Don Carlo Gnocchi, 3, 00166, Roma, Italy

Email: `alfredo.donno@unicusano.it`

Emanuele Rodaro

Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133, Milano, Italy

Email: emanuele.rodaro@polimi.it