# METRICS AND EMBEDDINGS OF GENERALIZATIONS OF THOMPSON'S GROUP $F$

J. Burillo, S. Cleary and M.I. Stein

Abstract. The distance from the origin in the word metric for generalizations $F(p)$ of Thompson's group $F$ is quasi-isometric to the number of carets in the reduced rooted tree diagrams representing the elements of $F(p)$. This interpretation of the metric is used to prove that several types of embeddings of groups $F(p)$ into each other are quasi-isometric embeddings, and also to study the behavior of the shift maps under these embeddings.

When obtaining the first example of a finitely presented infinite simple group in [12], R. Thompson introduced groups $F$, $T$ and $V$, which have attracted considerable interest since then and have become known as Thompson's groups $F$, $T$ and $V$. Higman [10] generalized the group $V$ to a whole family of groups, a generalization that was extended to the subgroups $F$ and $T$ by Brown [3]. The groups $F(p)$ of this article correspond to the groups $F_{p,\infty}$ in [3]. These families of groups have also been extensively studied by Brin [2].

Thompson's group $F$ has proven to be the most interesting group of all and has emerged in a variety of settings. It can be regarded as the group of piecewise-linear, orientation-preserving homeomorphisms of the interval $[0,1]$ which have breakpoints only at dyadic points and whose slopes, where defined, are powers of two. In 1984 Brown and Geoghegan [4] found it to be the first example of a finitely presented infinite-dimensional torsion-free $FP_\infty$ group. This fact has been extended to all $F(p)$ in [3], and studied by Stein [11], where generalizations to more general groups of homeomorphisms with general rational breakpoint sets are considered. Cleary [7] has studied these properties for groups of piecewise-linear homeomorphisms with irrational breakpoints and slopes.

In [5] Burillo showed the relationship between the word metric of Thompson's group $F$ and an estimate of the distance derived from the unique normal form of the elements. This algebraic estimate is quasi-isometric to the word metric and was used to prove that some subgroups are nondistorted in $F$. In this paper we find a geometric estimate of the word metric in terms of rooted tree diagrams (see [6] and [8]), show that this is quasi-isometric to the word metric, and use this interpretation to prove that some embeddings of groups $F(p)$ into others are quasi-isometric. The interpretations of these embeddings in terms of the tree diagrams also yield insights into the behavior of the shift maps under the embeddings.

After a statement of some results about the groups $F(p)$ in section 1 that will be needed later in the paper, including a brief description of the rooted tree diagrams, section 2 contains the interpretation of the word metric in terms of the normal form and of the number of carets in the tree diagrams. In section 3 the natural embedding of $F(p^k)$ into $F(p)$ is studied and proved to be nondistorted, and in section 4 some different embeddings are found, proving that every $F(p)$ can be quasi-isometrically embedded in $F(q)$ for some $q > p$. The last section is dedicated to study the behavior of the shift maps of $F(p)$ under these embeddings.

---

1991 *Mathematics Subject Classification.* 20F32.

*Key words and phrases.* Thompson's group, word metric, quasi-isometric embedding.

## 1. The groups $F(p)$

The generalizations of Thompson's group which are the subject of this paper are the groups $F(p) = F(1, \mathsf{Z}[\frac{1}{p}], \langle p \rangle)$, the groups of piecewise-linear, orientation-preserving homeomorphisms of the interval $[0,1]$ which have breakpoints only in $\mathsf{Z}[\frac{1}{p}]$, and such that the slopes, where defined, are powers of $p$. For $p = 2$ this group is the well-known Thompson's group $F$, and the groups $F(p)$ are natural generalizations of $F$, and share many of its interesting properties. In this section we outline some of these properties that will be used later in the paper. For a very readable introduction to $F$ see [6], and for generalizations to $F(p)$ see [3] and [11].

The group $F(p)$ admits the following infinite presentation:

$$\mathcal{P}_p = \left\langle x_i,\, i \geq 0 \,|\, x_i^{-1} x_j x_i = x_{j+p-1}, \text{ for } i < j \right\rangle$$

where the maps $x_i \in F(p)$, for $0 \leq i \leq p-2$, can be represented by the homeomorphisms of the unit interval:

$$x_i(t) = \begin{cases} t & \text{if } 0 \leq t \leq \frac{i}{p}, \\ pt - \frac{i(p-1)}{p} & \text{if } \frac{i}{p} \leq t \leq \frac{(p-1)(i+1)}{p^2}, \\ t + \frac{(p-1)(p-i-1)}{p^2} & \text{if } \frac{(p-1)(i+1)}{p^2} \leq t \leq \frac{i+1}{p}, \\ \frac{t+p-1}{p} & \text{if } \frac{i+1}{p} \leq t \leq 1; \end{cases}$$

For $x_i$ with $i \geq p-1$, we let $j = \left\lfloor \frac{i}{p-1} \right\rfloor$, and let $k = i - j(p-1)$. Then, $x_i \in F(p)$ is the identity except in the interval $\left[1 - \frac{1}{p^j}, 1\right]$, where the graph is a scaled-down copy of the graph of $x_k$. The compositions are taken on the right; that is, the element $x_i x_j \in F(p)$ corresponds with the composition $x_j \circ x_i$ as maps in $[0,1]$.

The groups $F(p)$ admit a *shift map* $\phi$, which sends every generator $x_i$ to $x_{i+1}$. The shift map satisfies $x_0^{-1}\phi(x)x_0 = \phi^p(x)$ for all $x \in F(p)$, so it is a conjugacy idempotent. The relations between the shift maps of the different $F(p)$ are studied in section 5.

The infinite presentation $\mathcal{P}_p$ is useful for its symmetry and simplicity, but to discuss the word metric we need to consider a finite presentation. It is clear that $x_0, x_1, \ldots, x_{p-1}$ generate $F(p)$, and it is possible to write a presentation for $F(p)$ with these $p$ generators and $p(p-1)$ relators (see [9] and [11]). For $p = 2$ this is the standard presentation for Thompson's group $F$:

$$\left\langle x_0,\, x_1 \,|\, [x_0 x_1^{-1}, x_2],\, [x_0 x_1^{-1}, x_3] \right\rangle.$$

In the following, when we refer to the word metric, or the length of an element, we will always mean with respect to these finite presentations.

From the relators $x_i x_j x_i^{-1} = x_{j+p-1}$ we can see that every element in $F(p)$ admits an expression of the form

$$x_{i_1}^{r_1} x_{i_2}^{r_2} \ldots x_{i_n}^{r_n} x_{j_m}^{-s_m} \ldots x_{j_2}^{-s_2} x_{j_1}^{-s_1}$$

where $i_1 < i_2 < \ldots < i_n \neq j_m > \ldots > j_2 > j_1$. This expression is unique if we require one additional condition: if both $x_i$ and $x_i^{-1}$ appear, then one of the generators

$$x_{i+1},\, x_{i+1}^{-1},\, x_{i+2},\, x_{i+2}^{-1}, \ldots, x_{i+p-1},\, x_{i+p-1}^{-1}$$

must appear as well. This is required for uniqueness because if none of them appears, there is a subword of the type $x_i\phi^p(x)x_i^{-1}$ which can be replaced by $\phi(x)$. This unique expression of an element will be called its *normal form*. The proof of the uniqueness of the normal form in $F(2)$ in [4] extends easily to every $F(p)$. Given an element $x$, its normal form is the shortest word in the infinite generating set of $\mathcal{P}_p$ representing it.

The groups $F(p)$ can also be realized as groups of homeomorphisms of the real line $\mathsf{R}$. The maps

$$f_i(t) = \begin{cases} t & \text{if } t \le i, \\ pt + i(1-p) & \text{if } i \le t \le i+1, \\ t+p-1 & \text{if } i+1 \le t \end{cases}$$

generate a group of piecewise-linear homeomorphisms of $\mathsf{R}$ which is isomorphic to $F(p)$. As before, compositions are taken on the right. We will use both of these geometric representations for $F(p)$ to deduce different properties of these groups.

Another interpretation for $F(p)$ is based on maps of rooted trees. This interpretation was studied by Higman in [10] and Bieri–Strebel in [1], and it is described with great detail for the case of $F(2)$ in [6]. A rooted $p$-tree is a tree with a distinguished vertex (the *root*) which has $p$ edges, and any other vertex has either valence 1 (*leaves*) or valence $p+1$ (*nodes*). We think of a rooted $p$-tree as a *descending* tree, with the root on top, and different levels of vertices, with the root being the sole vertex of level 0. Every vertex different from the root is connected by an edge to a vertex in the previous level, and it is either a leaf, in which case it is not connected to the next lower level, or a node, which has $p$ *children*, i.e., it is connected to $p$ vertices in the next lower level. A node, together with its $p$ children, and the $p$ edges connecting them, form a *caret*.

A rooted $p$-tree can be thought of as a graphic representation of a subdivision of the interval $[0,1]$ into intervals of lengths $\frac{1}{p^k}$ for different $k$. A vertex in level $k$ corresponds to an interval of length $\frac{1}{p^k}$. If it is a node, the caret represents the subdivision of the interval in $p$ intervals of length $\frac{1}{p^{k+1}}$. The leaves represent the final intervals of the subdivision, and the order of the intervals in $[0,1]$ induces a total order on the leaves of the tree. The endpoints of the intervals are always elements of $\mathsf{Z}[\frac{1}{p}]$.

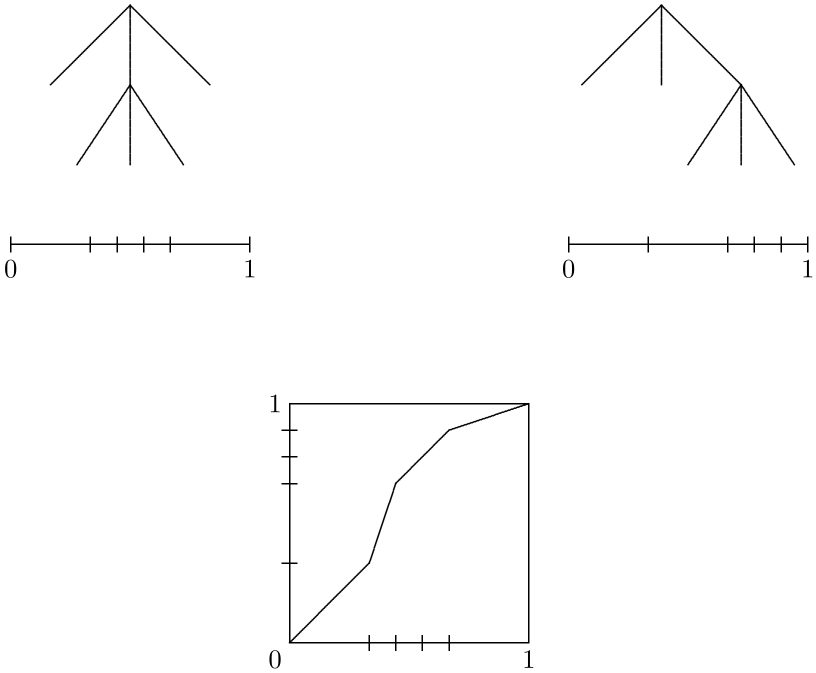


**Fig. 1:** The tree diagram for $x_1 \in F(3)$ with its homeomorphism of $[0,1]$.

An element $x \in F(p)$ can be thought of, then, as a *tree diagram*, i.e., an order-preserving map between the leaves of two rooted $p$-trees with the same number of carets (and thus the same number of leaves). The homeomorphism $x$ of $[0,1]$ represented by this diagram can be seen using the subdivisions represented by the two trees: the intervals of the source tree are sent to the ones in the target tree, preserving the order. The product of two elements $x$ and $y$ can be seen with the tree diagrams using the following process: add carets simultaneously at corresponding leaves to the source and target trees of $x$, and also add carets to the pair of trees of $y$, until the target tree of $x$ and the source tree of $y$ are equal. Then, a tree diagram for $xy$ has for source the source of $x$ and for target the target of $y$, with all the added carets needed to perform the composition.

A diagram is reducible if all the leaves of a caret are sent to all the leaves of another caret in the target, that is, if these carets represent superfluous subdivisions of the corresponding intervals. In the following, we will assume that our diagrams are reduced, meaning that superfluous carets have been eliminated. For a given element, the reduced tree diagram representing it is unique, and there is a close relation between the reduced tree diagram and the normal form of the element ([6], [8]).

Reduced $p$-tree diagrams are a powerful and efficient way to represent elements of $F(p)$, and they will be used several times later in the paper. A complete detailed description of the tree diagram representation for $F(2)$ can also be found in [8], where B. Fordham uses it to obtain an algorithm to compute the exact length of an element of $F(2)$ given its normal form.

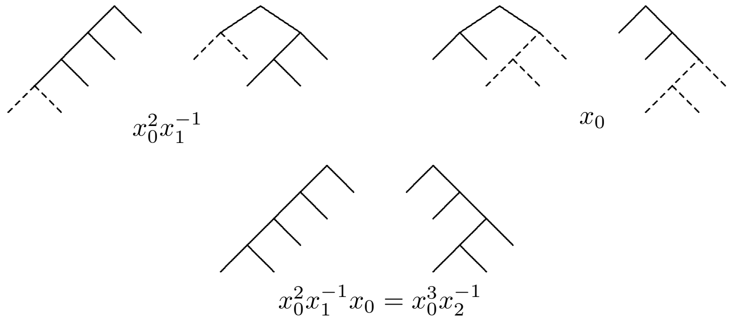


**Fig. 2:** Composing two tree diagrams in $F(2)$, with the added carets in dashes.

Yet a different representation of the groups $F(p)$ can be found in [9], in the context of diagram groups. This representation is essentially the same as the tree diagram representation, where a node with its $p+1$ edges is replaced by a 2-cell with its boundary subdivided in $p+1$ edges, and the cells are attached to each other along the edges according to the same pattern represented by the trees.

## 2. The word metric of $F(p)$

The different interpretations of the groups $F(p)$, both as groups of homeomorphisms and as groups of maps of rooted trees can be used to deduce expressions for the word metric. First we generalize to all $F(p)$ the estimate of the length of an element developed in [5]. This gives a quantity which is equivalent to the length, and can be readily computed from the normal form of an element. Here, we mean equivalent in the sense of quasi-isometry. We will denote by $|x|_p$ the minimal length of an element of $F(p)$ in the word metric with respect to the generators $x_0, x_1, \dots, x_{p-1}$.

**Theorem 1.** *Let* $x \in F(p)$ *have normal form*

$$x_{i_1}^{r_1}x_{i_2}^{r_2}\dots x_{i_n}^{r_n}x_{j_m}^{-s_m}\dots x_{j_2}^{-s_2}x_{j_1}^{-s_1},$$

*and let*

$$D(x) = r_1 + r_2 + \dots + r_n + s_1 + s_2 + \dots + s_m + i_n + j_m.$$

*Then, we have*

$$\frac{D(x)}{3(p-1)} \le |x|_p \le 3D(x).$$

*Proof.* The upper bound is straightforward: replacing each $x_i$ in the normal form by its expression in the generators $x_0, x_1, \dots, x_{p-1}$, the length of the word obtained is less than $3D(x)$, and clearly it is an upper bound for the minimal length.

For the lower bound, observe that since the normal form has the shortest possible length among words representing $x$ in the generators of $\mathcal{P}_p$, we can conclude that $|x|_p \geq r_1 + \ldots + r_n + s_1 + \ldots + s_m$. Finally, let $w$ be the word in the generators $x_0, x_1, \ldots, x_{p-1}$ which has minimal length $|x|_p$. To obtain the unique normal form from $w$, the generators have to be switched around using the relators of $\mathcal{P}_p$, at the price of increasing the index of one of them by $p-1$ per switch. A given generator in $w$ needs to be switched at most $|x|_p - 1$ times, so the highest possible generator appearing in the normal form has index at most

$$p - 1 + (p-1)(|x|_p - 1) = (p-1)|x|_p$$

and then

$$i_n \leq (p-1)|x|_p \qquad \text{and} \qquad j_m \leq (p-1)|x|_p.$$

Combining all these inequalities we obtain the desired lower bound. /

Part of the previous proof is due to V. Guba, who improved on the proof given in [5].

The number $D(x)$ given in the previous result is, then, equivalent to the distance. This readily computable quasi-metric $D(x)$ can be used in the place of the genuine word metric to obtain geometric characterizations for the distance.

**Proposition 2.** *Let $x \in F(2)$ be an element whose normal form is a positive word, i.e.,*

$$x = x_{i_1}^{r_1} x_{i_2}^{r_2} \ldots x_{i_n}^{r_n}.$$

*Let*

$$N(x) = \max \{ i_k + r_k + r_{k+1} + \ldots + r_n + 1, \ \textit{for } k = 1, 2, \ldots, n \}.$$

*Then,*

1. *$N(x)$ is equal to the number of carets in either tree of the reduced 2-tree diagram for $x$,*
2. *$N(x)$ is equal to the $y$-coordinate of the last breakpoint of the graph of $x$ represented as a homeomorphism of* $\mathsf{R}$,
3. *$N(x)$ is quasi-equivalent to the distance. In particular,*

$$\frac{D(x)}{2} \leq N(x) \leq D(x) + 1.$$

*Proof.* The statement that $N(x)$ is equivalent to the distance is clear from the definition: clearly $N(x) \leq D(x) + 1$, and also

$$N(x) \geq i_1 + r_1 + \ldots + r_n + 1 \geq r_1 + \ldots + r_n, \text{ and}$$
$$N(x) \geq i_n + r_n + 1 \geq i_n,$$

from which the lower bound follows.

Both statements (1) and (2) will be proved using the same induction in $n$. If $n = 1$, then $x = x_i^r$, and we have that the tree diagram of $x$ has exactly $i + r + 1$ carets (see Figure 3). The graph for the homeomorphism representing $x$ has the following slopes at points with the given $y$-coordinate:

$$\begin{array}{rll} 1 & \text{for } y \in (-\infty, i), & \\ 2^r & \text{for } y \in (i, i+2), & \\ 2^{r-k} & \text{for } y \in (i+k+1, i+k+2), & k = 1, \dots, r-1, \\ 1 & \text{for } y \in (i+r+1, \infty). & \end{array}$$

In particular, the last breakpoint is $(i+1, i+r+1)$.

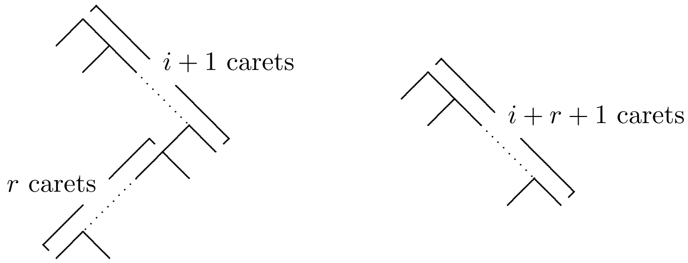


**Fig. 3:** The tree diagram for $x_i^r$.

For $n > 1$, let $x = x_{i_1}^{r_1} x_{i_2}^{r_2} \dots x_{i_n}^{r_n}$, and take $y = x x_i^r$. When composing the two tree diagrams of $x$ and $x_i^r$, we need to add carets to the middle two trees to make them match. We need to study two cases.

*Case 1:* Suppose that $i + 1 \leq N(x)$.

By induction hypothesis, the trees for the reduced tree diagram for $x$ have $N(x)$ carets, and we need to add $r$ carets to the target tree of $x$, and $N(x) - i - 1$ to the source tree for $x_i^r$ to perform the composition. In any case the resulting trees have $N(x) + r$ carets, and observe that if $i + 1 \leq N(x)$, then $N(y) = N(x) + r$. It is important to realize that the tree diagram obtained for $y$ is reduced. This fact can be observed using the relations between the reduced tree diagram and the exponents which appear in the unique normal form (see [8]).

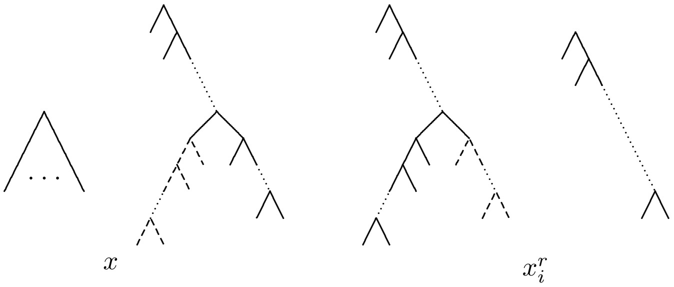


**Fig. 4:** The composition of $x$ with $x_i^r$ when $i + 1 \leq N(x)$.

For the graphs of the corresponding homeomorphisms, again from $i + 1 \leq N(x)$, we see that when composing $x$ with $x_i^r$ the only interval that gets modified is $[i, i+1]$, which gets stretched into $[i, i+r+1]$. The last breakpoint of $y$ is now $N(x) + r$, which as before is equal to $N(y)$.

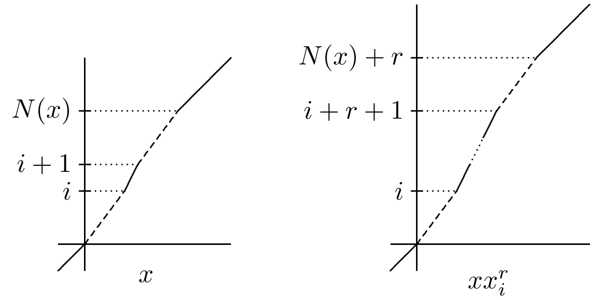


**Fig. 5:** The graphs of $x$ and $xx_i^r$ when $i+1 \leq N(x)$.

*Case 2:* If $i+1 > N(x)$, then in the composition of the two tree diagrams we need to add $i+r+1-N(x)$ carets to the target tree of $x$ to match the source of $x_i^r$:

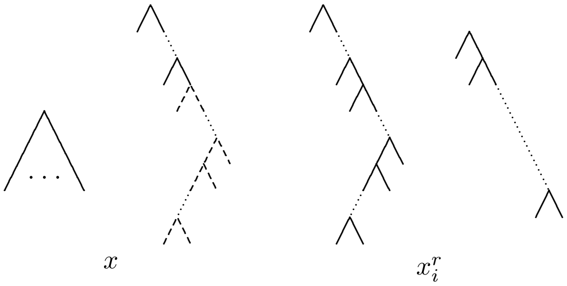


**Fig. 6:** The composition of $x$ with $x_i^r$ when $i+1 > N(x)$.

The resulting trees have $i+r+1$ carets, and if $i+1 > N(x)$, it is clear that $N(y) = i+r+1$ as well. For the homeomorphisms, if $i+1 > N(x)$, all the modifications to the graph of $x$ occur above the last breakpoint, so the last breakpoint for $y$ is the same than for $x_i^r$, which has $y$-coordinate $i+r+1$. /

For the case of a general word, not necessarily positive, there is no relation between the $y$-coordinate of the last breakpoint and the distance. The elements

$$x_0 x_1 \dots x_{k-1} x_k^2 x_{k+1}^{-1} x_k^{-1} \dots x_1^{-1} x_0^{-1}$$

have all breakpoints in the square $[0,1] \times [0,1]$, whereas their norm is linear in $k$. But the number of carets in the reduced diagram is still equivalent to the norm.

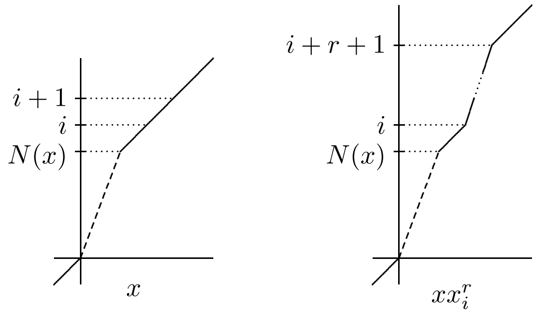


**Fig. 7:** The graphs of $x$ and $xx_i^r$ when $i+1 > N(x)$.

**Theorem 3.** *Let $x \in F(2)$ be an element whose normal form is*

$$x_{i_1}^{r_1} x_{i_2}^{r_2} \dots x_{i_n}^{r_n} x_{j_m}^{-s_m} \dots x_{j_2}^{-s_2} x_{j_1}^{-s_1},$$

*and let*

$$\begin{aligned} y_1 &= x_{i_1}^{r_1} x_{i_2}^{r_2} \dots x_{i_n}^{r_n} \\ y_2 &= x_{j_1}^{s_1} x_{j_2}^{s_2} \dots x_{j_m}^{s_m} \end{aligned}$$

*be the two positive words involved in the normal form for $x = y_1 y_2^{-1}$. Then the number $N(x)$ of carets for any tree in the reduced tree diagram for $x$ is equal to the highest number of carets in the diagrams for $y_1$ and $y_2$. This number of carets is equal to*

$$\begin{aligned} N(x) &= \max\{N(y_1), N(y_2)\} \\ &= \max\{i_k + r_k + r_{k+1} + \dots + r_n + 1, \ \textit{for } k = 1, 2, \dots, n, \\ &\qquad\quad j_l + s_l + s_{l+1} + \dots + s_m + 1, \ \textit{for } l = 1, 2, \dots, m\}, \end{aligned}$$

*and it is equivalent to the distance.*

*Proof.* The equivalence with the distance proceeds as in Proposition 2, except the inequalities are now:

$$\frac{D(x)}{4} \le N(x) \le D(x) + 1.$$

For the number of carets, we only need to realize that to obtain the diagram for $x$ we need to put the two diagrams for $y_1$ and $y_2$ next to each other with the diagram for $y_2$ reversed, and add carets to the one with fewer of them.

The tree diagram obtained will be reduced because of the uniqueness of the normal form, again by the results in [8]. /

Note that these results only apply to $F(2)$, where the proofs are simple and for a positive word in $F(2)$ the number of carets and the $y$-coordinate of the last breakpoint coincide. This fact is not true for $F(p)$ if $p > 2$, but even though those two numbers are different, both are equivalent to the distance. For general $F(p)$ we have the following results with analogous proofs:

**Proposition 4.** *Let*

$$x = x_{i_1}^{r_1} x_{i_2}^{r_2} \dots x_{i_n}^{r_n}$$

*be a positive word in* $F(p)$*. Then the number*

$$N_1(x) = \max \left\{ \left\lfloor \frac{i_k}{p-1} \right\rfloor + r_k + r_{k+1} + \dots + r_n + 1, \textit{ for } k = 1, 2, \dots, n \right\}$$

*is equal to the number of carets in either tree of the reduced* $p$*-tree diagram for* $x$*. This number satisfies the inequality*

$$\frac{D(x)}{2(p-1)} \leq N_1(x) \leq D(x) + 1.$$

/

**Proposition 5.** *Let*

$$x = x_{i_1}^{r_1} x_{i_2}^{r_2} \dots x_{i_n}^{r_n}$$

*be a positive word in* $F(p)$*. Then the number*

$$N_2(x) = \max \{ i_k + r_k(p-1) + r_{k+1}(p-1) + \dots + r_n(p-1) + 1, \\ \textit{for } k = 1, 2, \dots, n \}$$

*is equal to the* $y$*-coordinate of the last breakpoint of the graph for* $x$*. This number satisfies the inequality*

$$\frac{D(x)}{2} \leq N_2(x) \leq D(x)(p-1) + 1.$$

/

And for connection between the number of carets with the word metric for a general, not necessarily positive word, we have:

**Theorem 6.** *Let*

$$x = x_{i_1}^{r_1} x_{i_2}^{r_2} \dots x_{i_n}^{r_n} x_{j_m}^{-s_m} \dots x_{j_2}^{-s_2} x_{j_1}^{-s_1}$$

*be the unique normal form of an element* $x \in F(p)$*. Then the number*

$$N(x) = \max \left\{ \left\lfloor \frac{i_k}{p-1} \right\rfloor + r_k + r_{k+1} + \dots + r_n + 1, \textit{ for } k = 1, 2, \dots, n, \right. \\ \left. \left\lfloor \frac{j_l}{p-1} \right\rfloor + s_l + s_{l+1} + \dots + s_m + 1, \textit{ for } l = 1, 2, \dots, m \right\}$$

*is equal to the number of carets in either tree for the reduced diagram for* $x$*. This number is equivalent to the distance. In particular, it satisfies the inequality*

$$\frac{D(x)}{4(p-1)} \leq N(x) \leq D(x) + 1.$$

/

This estimate of the distance in terms of the number of carets in the trees of the diagram will be used extensively in the next sections, to prove that several embeddings of a group $F(p)$ in another group $F(q)$ are nondistorted.

## 3. THE EMBEDDING OF $F(p^k)$ IN $F(p)$

There are several types of embeddings of groups $F(q)$ as subgroups of groups $F(p)$. The most natural one is when $q$ is a power of $p$, since then $\mathsf{Z}[\frac{1}{q}] = \mathsf{Z}[\frac{1}{p}]$. It is easier to understand these embeddings in terms of carets than in terms of homeomorphisms of $[0,1]$, as can be seen in the example $F(4) \subset F(2)$, which we will describe below in detail.

The embedding $i : F(4) \to F(2)$ can be understood using the tree diagrams. Let $T$ be a rooted 4-tree. As seen in section 2, $T$ can be understood as a subdivision of $[0,1]$ in intervals of length $\frac{1}{4^k}$. But clearly, subdividing an interval in four equal parts corresponds to subdivide the interval first in two parts and then each of these parts in two more. So given a rooted 4-tree $T$, there is a (unique) rooted 2-tree $i(T)$ which yields the same subdivision of $[0,1]$. The tree $i(T)$ can be obtained from $T$ replacing each 4-caret of $T$ by a set of three 2-carets in the obvious manner:

**Fig. 8:** A 4-caret and the set of 2-carets for the embedding of $F(4)$ in $F(2)$.

Now, given an element $x \in F(4)$, represented by the reduced tree diagram $(S,T)$, the element $i(x) \in F(2)$ is represented by the tree diagram $(i(S), i(T))$. Clearly the diagram $(i(S), i(T))$ will not be reduced in general, so it is necessary to reduce it. A table with the four generators of $F(4)$ with their corresponding (reduced) images in $F(2)$ is detailed in Figure 9.

Observe that in the process of reducing the diagram $(i(S), i(T))$, some 2-carets will be eliminated. But for a set of three 2-carets which corresponds to the image of a 4-caret will never be completely erased in the reduction, because if it were, that would mean that the 4-caret they correspond to would already be superfluous. So, in the trees $i(S)$ and $i(T)$ there are at least as many carets as there were in $S$ and in $T$. This provides the following inequality:

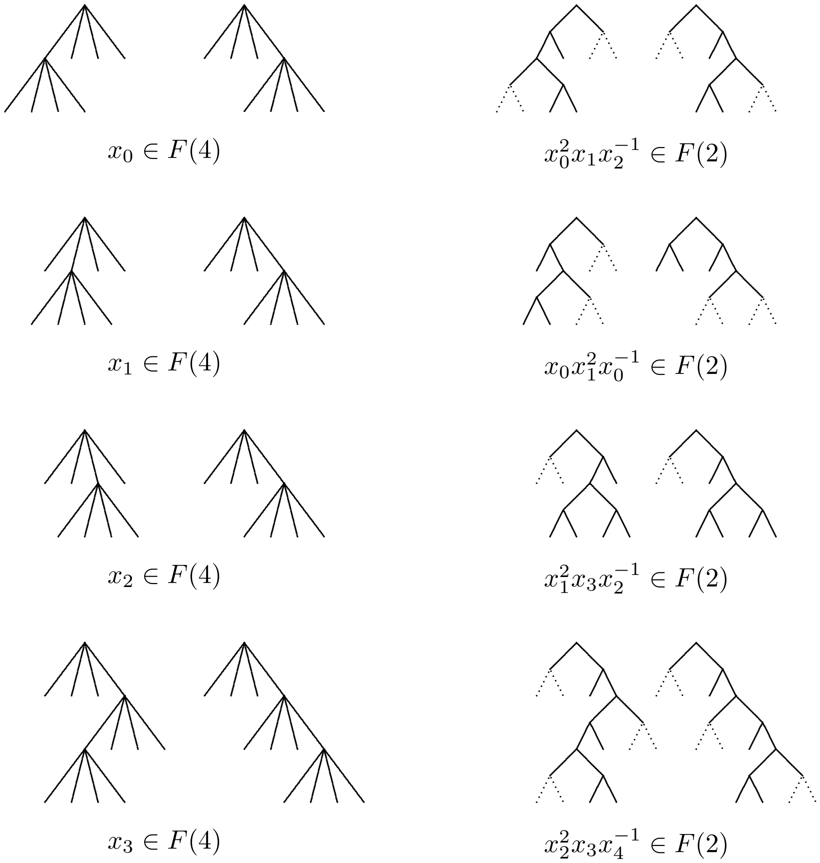
$$N(x) \leq N(i(x)) \leq 3N(x).$$

By virtue of the results in section 2, the number of carets in the reduced diagram of an element is equivalent to the distance, we obtain that the norm of an element in $F(4)$ and of its image in $F(2)$ are equivalent, so the embedding of $F(4)$ in $F(2)$ is a quasi-isometric embedding.

In the general case, with $i : F(p^k) \longrightarrow F(p)$, the embedding can also be seen with trees in the exact same way. A $p^k$-caret is now replaced by

$$1 + p + p^2 + \ldots + p^{k-1} = \frac{p^k - 1}{p - 1}$$

$p$-carets, and after the reductions, at least one of the $p$-carets survives per each $p^k$-caret. This gives the following result:



$x_0 \in F(4)$ $x_0^2 x_1 x_2^{-1} \in F(2)$

$x_1 \in F(4)$ $x_0 x_1^2 x_0^{-1} \in F(2)$

$x_2 \in F(4)$ $x_1^2 x_3 x_2^{-1} \in F(2)$

$x_3 \in F(4)$ $x_2^2 x_3 x_4^{-1} \in F(2)$

**Fig. 9:** The images of the generators of $F(4)$ in $F(2)$.

**Theorem 7.** *The natural embedding of $F(p^k)$ in $F(p)$ is a quasi-isometric embedding.* /

The inequalities here are

$$N(x) \leq N(i(x)) \leq \frac{p^k - 1}{p - 1} N(x).$$

Note that this embedding is not quasi-onto, since the image is nowhere near being $\epsilon$-dense. For a general element in $F(4)$, its image in $F(2)$ will have norm about three times as large.

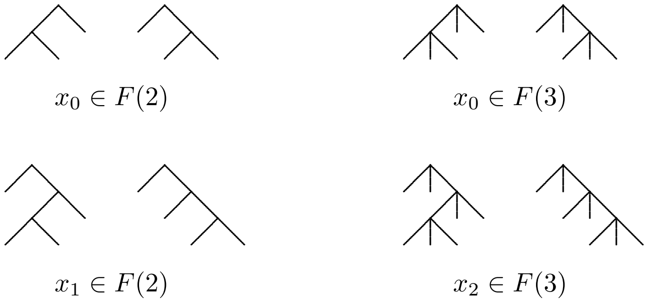
## 4. EMBEDDINGS OF $F(p)$ IN $F(q)$ FOR $p < q$

Given a group $F(p)$, there always exists another group $F(q)$ which contains it as a subgroup. The simplest example is the subgroup of $F(3)$ generated by all the generators with *even* subindex in the presentation $\mathcal{P}_3$. These generators satisfy the relations

$$x_{2i}^{-1}x_{2j}x_{2i} = x_{2j+2}, \qquad \text{for } i < j,$$

and hence generate a subgroup of $F(3)$ isomorphic to $F(2)$. To see this embedding in terms of the tree diagrams, observe that the generators $x_0$ and $x_1$ of $F(2)$ and $x_0$ and $x_2$ of $F(3)$ have the same number of carets, and in the same disposition, the only difference is that the carets for $F(2)$ have two edges, whereas the carets for $F(3)$ have three (see Figure 10).

So the embedding of $F(2)$ into $F(3)$ can be realized in terms of the tree diagrams by, given a rooted 2-tree, just add a new edge in the middle of every caret to transform it into a 3-caret. Given an element of $F(2)$ with its tree diagram, add an edge to every caret in both trees to obtain a 3-tree diagram for the image of the given element in $F(3)$. And the resulting 3-tree diagram is reduced if and only if the starting 2-tree diagram is reduced, since the carets are arranged in exactly the same pattern. Hence both diagrams have the same number of carets, and since the number of carets is equivalent to the distance, the embedding is a quasi-isometric embedding.



$x_0 \in F(2)$ $x_0 \in F(3)$

$x_1 \in F(2)$ $x_2 \in F(3)$

**Fig. 10:** The two generators of $F(2)$ and their images in $F(3)$.

We can generalize this embedding to many more cases:

**Theorem 8.** *The group $F(p)$ embeds in $F(q)$ for all such pairs $p, q$ such that $q-1$ is a multiple of $p-1$.*

*Proof.* Let $q-1 = n(p-1)$. Then take all generators in $F(q)$ which are multiple of $n$. These elements generate a copy of $F(p)$. To see this, observe that these generators satisfy the relations

$$x_{ni}^{-1}x_{nj}x_{ni} = x_{nj+q-1} = x_{nj+n(p-1)} = x_{n(j+p-1)}.$$

/

This embedding is the restriction to the groups $F(p)$ of the embeddings described algebraically by Higman [10,Theorem 7.2] for the larger groups $G_{n,r}$.

This embedding can also be seen in terms of tree diagrams. Take a $p$-caret and insert $n-1$ extra edges between every two of the original edges. The resulting caret has $q$ edges, so by doing this to any rooted $p$-tree we obtain a rooted $q$-tree with the same number of carets. For any reduced $p$-tree diagram we obtain the reduced $q$-diagram for the image of the element it represents, and both diagrams have the same number of carets. Hence:

**Theorem 9.** *If $q-1$ is a multiple of $p-1$, the resulting embedding of $F(p)$ in $F(q)$ is a quasi-isometric embedding.* /

As before, these embeddings are not quasi-onto. In particular, even though we have embeddings of $F(4)$ into $F(2)$ and of $F(2)$ into $F(4)$, the compositions of these embeddings are not quasi-isometries.

## 5. The shift maps

One of the most interesting features of the groups $F(p)$ are the shift maps. The shift map $\phi_p$ for $F(p)$ is defined as the map sending every generator $x_i$ in the presentation $\mathcal{P}_p$ to $x_{i+1}$. As we have seen before, $\phi_p$ satisfies $x_0^{-1}\phi_p(x)x_0 = \phi_p^p(x)$. Of special relevance is the map $\phi_p^{p-1}$, because the relators of $\mathcal{P}_p$ can be written as

$$x_i^{-1}x_jx_i = \phi_p^{p-1}(x_j), \qquad \text{for } i<j;$$

so if we take an HNN extension of $F(p)$ by the map $\phi_p^{p-1}$, we obtain another copy of $F(p)$.

The map $\phi_p^{p-1}$ also has significance in the homeomorphisms of $[0,1]$. The image of an element $x \in F(p)$ by $\phi_p^{p-1}$ is the identity in the interval $\left[0, 1-\frac{1}{p}\right]$ and has a copy of the graph of $x$ in the interval $\left[1-\frac{1}{p}, 1\right]$, scaled down by a factor of $p$.

Also, it is easy to interpret the maps $\phi_p^{p-1}$ in terms of the rooted $p$-tree diagrams. Given a rooted $p$-tree $T$, consider another tree $T'$ obtained by taking one single $p$-caret and attaching $T$ by the root to the rightmost vertex of the caret.

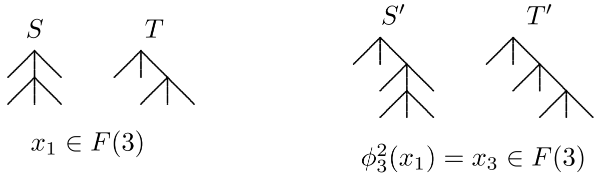


**Fig. 11:** The image of $x_1 \in F(3)$ under $\phi_3^2$.

Using this construction, it is easy to see that if $(S,T)$ is the reduced tree diagram for $x \in F(p)$, then $(S',T')$ is the reduced tree diagram for $\phi_p^{p-1}(x)$. We can use this interpretation of the shift maps to see that they behave nicely under the embeddings studied in sections 3 and 4.

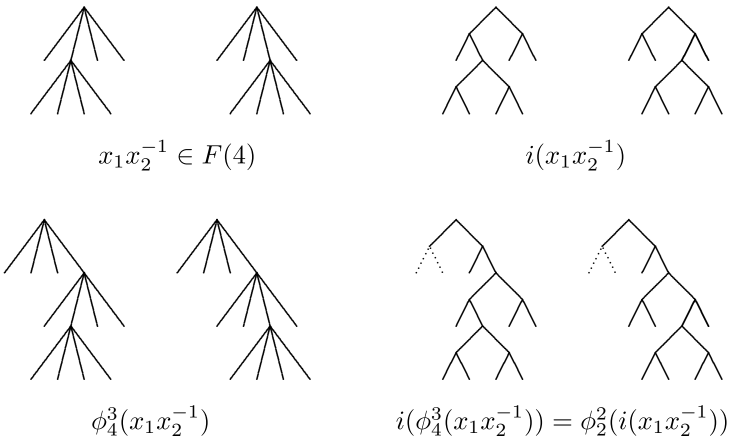
**Proposition 10.** *Let*

$$i : F(p^k) \longrightarrow F(p)$$

*be the natural embedding. Then the shift maps satisfy:*

$$i \circ \phi_{p^k}^{p^k-1} = \phi_p^{k(p-1)} \circ i.$$

*Proof.* Attaching an extra $p^k$-caret to the top of a rooted tree corresponds by the embedding $i$ to attaching $\frac{p^k-1}{p-1}$ $p$-carets to the top of a rooted $p$-tree. But on a reduced $p$-tree diagram, after attaching these carets to the top of each tree, we see that all the attached carets except the $k$ rightmost ones will be reduced.



$x_1x_2^{-1} \in F(4)$ $\quad$ $i(x_1x_2^{-1})$

$\phi_4^3(x_1x_2^{-1})$ $\quad$ $i(\phi_4^3(x_1x_2^{-1})) = \phi_2^2(i(x_1x_2^{-1}))$

**Fig. 12:** An example of the invariance of the shift maps.

So the resulting tree diagram in $F(p)$ is the diagram obtained by attaching the $k$ carets that are necessary to apply $\phi_p^{k(p-1)}$. /

**Proposition 11.** *Let*

$$j : F(p) \longrightarrow F(q)$$

*be an embedding with* $q-1 = n(p-1)$*. Then the shift maps satisfy:*

$$j \circ \phi_p^{p-1} = \phi_q^{q-1} \circ j.$$

*Proof.* Both shift maps act adding a single caret on top of the trees, and this operation is not affected by adding extra edges to an existing caret, which is what the embedding $j$ demands (see section 4). Clearly the same result is obtained by adding a $p$-caret and then adding $n-1$ edges between each two than by adding first the sets of $n-1$ edges and later adding a $q$-caret on top. /

## References

1. Bieri, R., Strebel, R., *On groups of PL-homeomorphisms of the real line*, Notes, Math. Sem. der Univ. Frankfurt, 1985.
2. Brin, M.G., *The Chameleon Groups of Richard J. Thompson: Automorphisms and Dynamics*, Publ. Math. IHES **84** (1996), 5–33.
3. Brown, K.S., *Finiteness properties of groups*, J. Pure App. Algebra **44** (1987), 45–75.
4. Brown, K.S., Geoghegan. R., *An infinite-dimensional torsion-free $FP_\infty$ group*, Invent. Math. **77** (1984), 367–381.
5. Burillo, J., *Quasi-isometrically embedded subgroups of Thompson's group $F$*, preprint.
6. Cannon, J.W., Floyd, W.J., Parry, W.R., *Introductory notes on Richard Thompson's groups*, L'Ens. Math. **42** (1996), 215–256.

7. Cleary, S., *Groups of piecewise-linear homeomorphisms with irrational slopes*, Rocky Mountain J. Math. **25** (1995), 935-955.
8. Fordham, S.B., *Minimal Length Elements of Thompson's group F*, thesis, Brigham Young University, 1995.
9. Guba, V., Sapir, M., *Diagram Groups*, Mem. Amer. Math. Soc. **130** (1997).
10. Higman, G., *Finitely presented infinite simple groups*, Notes on Pure Math., vol. 8, Australian National University, Canberra, 1974.
11. Stein, M., *Groups of piecewise linear homeomorphisms*, Trans. Amer. Math. Soc. **332** (1992), no. 2, 477–514.
12. Thompson, R. J., McKenzie, R., *An elementary construction of unsolvable word problems in group theory*, Word problems, Conference at University of California, Irvine, 1969 (1973), North Holland.

Dept. of Mathematics, Tufts University, Medford, MA 02155, U.S.A.
*E-mail address*: `jburillo@emerald.tufts.edu`

Dept. of Mathematics, California State University, Fresno, CA 93710, U.S.A.
*E-mail address*: `sean_cleary@csufresno.edu`

Dept. of Mathematics, Trinity College, Hartford, CT 06106, U.S.A.
*E-mail address*: `mstein@mail.trincoll.edu`