TAMARI LATTICES AND THE SYMMETRIC THOMPSON MONOID

PATRICK DEHORNOY

ABSTRACT. We investigate the connection between Tamari lattices and the Thompson group F, summarized in the fact that F is a group of fractions for a certain monoid F_{sym}^+ whose Cayley graph includes all Tamari lattices. Under this correspondence, the Tamari lattice operations are the counterparts of the least common multiple and greatest common divisor operations in F_{sym}^+. As an application, we show that, for every n, there exists a length ℓ chain in the nth Tamari lattice whose endpoints are at distance at most $12\ell/n$.

1. INTRODUCTION

The aim of this text is to show the interest of using monoid techniques to investigate Tamari lattices. More precisely, we shall describe the very close connection existing between Tamari lattices and a certain submonoid F_{sym}^+ of Richard Thompson's group F: equipped with the left-divisibility relation, the monoid F_{sym}^+ is a lattice that includes all Tamari lattices. Roughly speaking, the principle is to attribute to the edges of the Tamari lattices names that live in the monoid F_{sym}^+. By using the subword reversing method, a general technique from the theory of monoids, we then obtain a very simple way of reproving the existence of the lattice operations, computing them, and establishing further properties.

The existence of a connection between Tamari lattices, associativity, and the Thompson group F has been known for decades and belongs to folklore. What is specific here is the role of the monoid F_{sym}^+, which is especially suitable for formalizing the connection. Some of the results already appeared, implicitly in [6] and explicitly in [11]. Several new results are established in the current text, in particular the construction of a unique normal form in the monoid F_{sym}^+ and the group F (Subsection 4.4) and the (surprising) result that the embedding of the monoid F_{sym}^+ in the Thompson group F is not a quasi-isometry (Proposition 5.8). In the language of binary trees, this implies that, for every constant C, there exist chains of length ℓ whose endpoints can be connected by a path of length at most ℓ/C (Corollary 5.13).

Let us mention that a connection between the Tamari lattices and the group F is described in [27]. However both the objects and the technical methods are disjoint from those developed below. In particular, the approach of [27] does not involve the symmetric monoid F_{sym}^+, which is central here, but it uses instead the standard Thompson monoid F^+, which is not directly connected with the Tamari ordering.

The text is organized as follows. In Section 2, we recall the definition of Tamari lattices and Thompson's group F, and we establish a presentation of F in terms of some specific, non-standard generators a_α indexed by binary addresses. In Section 3, we investigate the submonoid F_{sym}^+ of F generated by the elements a_α, we prove that F_{sym}^+ equipped with

divisibility has the structure of a lattice, and we describe the (close) connection between this lattice and Tamari lattices. Then, in Section 4, we use the Polish encoding of trees to construct an algorithm that computes common upper bounds for trees in the Tamari ordering and we deduce a unique normal form for the elements of F and F_{sym}^+. Finally, in Section 5, we gather results about the length of the elements of F with respect to the generators a_α or, equivalently, about the distance in Tamari lattices, with a specific interest on lower bounds.

2. The framework

The aim of this section is to set our notation and basic definitions. In Subsections 2.1 and 2.2, we briefly recall the definition of the Tamari lattices in terms of parenthesized expressions and of binary trees, whereas Subsections 2.3 and 2.4 contain an introduction to Richard Thompson's group F and its action by rotation on trees. This leads us naturally to introducing in Subsection 2.5 a new family of generators of F indexed by binary addresses, and giving in Subsection 2.6 a presentation of F in terms of these generators.

2.1. **Parenthesized expressions and associativity.** Introduced by Dov Tamari in his 1951 PhD thesis, and appearing in the 1962 article [28]—also see [16] and [18]—the nth Tamari lattice, here denoted by $\mathcal{T}_n$, is, for every positive integer n, the poset (partially ordered set) obtained by considering all well-formed parenthesized expressions involving $n+1$ fixed variables and declaring that an expression E is smaller than another one E', written $E \leqslant_\mathcal{T} E'$, if E' may be obtained from E by applying the associative law $x(yz) = (xy)z$ in the left-to-right direction. As established in [28], the poset $(\mathcal{T}_n, \leqslant_\mathcal{T})$ is a lattice, that is, any two elements admit a least upper bound and a greatest lower bound. Moreover, $(\mathcal{T}_n, \leqslant_\mathcal{T})$ admits a top element, namely the expression in which all left parentheses are gathered on the left, and a bottom element, namely the expression in with all right parentheses are gathered on the right.

As associativity does not change the order of variables, we may forget about their names, and use $\bullet$ everywhere. So, for instance, there exist five parenthesized expressions involving four variables, namely $\bullet(\bullet(\bullet\bullet))$, $\bullet((\bullet\bullet)\bullet)$, $(\bullet(\bullet\bullet))\bullet$, $(\bullet\bullet)(\bullet\bullet)$, and $((\bullet\bullet)\bullet)\bullet$, and we have $\bullet((\bullet\bullet)\bullet) \leqslant_\mathcal{T} (\bullet(\bullet\bullet))\bullet$ in the Tamari order as one goes from the first expression to the second by applying the associativity law with $x = \bullet$, $y = \bullet\bullet$, and $z = \bullet$. The Hasse diagrams of the lattices $\mathcal{T}_3$ and $\mathcal{T}_4$ respectively are the pentagon and the 14 vertex polyhedron displayed in Figures 1 and 4 below. As is well known, the number of elements of $\mathcal{T}_n$ is the nth Catalan number $\frac{1}{n+1}\binom{2n}{n}$.
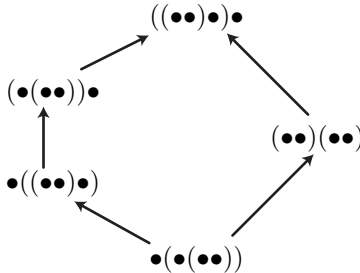

FIGURE 1. The Tamari lattice $\mathcal{T}_3$ made by the five ways of bracketing a four variable parenthesized expression.

The Tamari lattice $\mathcal{T}_n$ is connected with a number of usual objects. For instance, its Hasse diagram is the 1-skeleton—that is, the graph made of the 0- and 1-cells—of the

nth Mac Lane–Stasheff associahedron [20, 26]. Also $\mathcal{T}_n$ embeds in the lattice made by the symmetric group $\mathfrak{S}_n$ equipped with the weak order: $\mathcal{T}_n$ identifies with the sub-poset made by all 312-avoiding permutations (Björner & Wachs [2]).

For every n, replacing in a parenthesized expression the last (rightmost) symbol $\bullet$ with $\bullet\bullet$ defines an embedding ι_n of $\mathcal{T}_n$ into $\mathcal{T}_{n+1}$. We denote by $\mathcal{T}_\infty$ the limit of the direct system $(\mathcal{T}_n, \iota_n)$ so obtained. Note that $\mathcal{T}_\infty$ has a bottom element, namely the class of $\bullet$, which is also that of $\bullet\bullet$, $\bullet(\bullet\bullet)$, $\bullet(\bullet(\bullet\bullet))$, etc., but no top element.

2.2. Trees and rotations.
There exists an obvious one-to-one correspondence between parenthesized expressions involving $n+1$ variables and size n binary rooted trees, that is, trees with n interior nodes and $n+1$ leaves, see Figure 2. In this text, we shall use both frameworks equivalently. We denote by $T_0{}^\wedge T_1$ the tree whose left-subtree is T_0 and whose right-subtree is T_1, but skip the symbol $^\wedge$ in concrete examples involving $\bullet$.

$$\circ \qquad \wedge \qquad \wedge \qquad \wedge \qquad \wedge$$
$$\bullet \qquad \bullet\bullet \qquad (\bullet\bullet)\bullet \qquad \bullet(\bullet\bullet) \qquad \bullet((\bullet\bullet)\bullet)$$

FIGURE 2. Correspondence between parenthesized expressions and trees.

When translated in terms of trees, the operation of applying associativity in the left-to-right direction corresponds to performing one left-rotation, namely replacing some subtree of the form $T_0{}^\wedge(T_1{}^\wedge T_2)$ with the corresponding tree $(T_0{}^\wedge T_1)^\wedge T_2$, see Figure 3. So the Tamari lattice $\mathcal{T}_n$ is also the poset of size n trees ordered by the transitive closure of left-rotation. We naturally use $\leqslant_\mathcal{T}$ for the latter partial ordering.
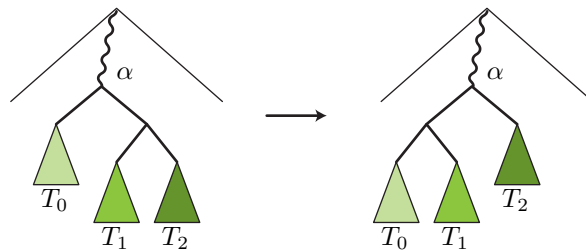

FIGURE 3. Applying a left rotation in a tree: replacing some subtree of the form $T_0{}^\wedge(T_1{}^\wedge T_2)$ with the corresponding tree $(T_0{}^\wedge T_1)^\wedge T_2$.

In terms of trees, the bottom element of the Tamari lattice $\mathcal{T}_n$ is the size n *right-comb* (or right-vine) C_n recursively defined by $C_0 = \bullet$ and $C_n = \bullet^\wedge C_{n-1}$ for $n \geqslant 1$, whereas the top element is the size n *left-comb* (or left-vine) $\widetilde{C}_n$ recursively defined by $\widetilde{C}_0 = \bullet$ and $\widetilde{C}_n = \widetilde{C}_{n-1}{}^\wedge\bullet$ for $n \geqslant 1$.

2.3. Richard Thompson's group F.
Introduced by Richard Thompson in 1965, the group F appeared in print only later, in [21] and [29]. The most common approach is to define F as a group of piecewise linear self-homeomorphisms of the unit interval $[0, 1]$.

Definition 1. The *Thompson group F* is the group of all dyadic order-preserving self-homeomorphisms of $[0, 1]$, where a homeomorphism f is called *dyadic* if it is piecewise linear with only finitely many breakpoints, every breakpoint of f has dyadic rational coordinates, and every slope of f is an integral power of 2.
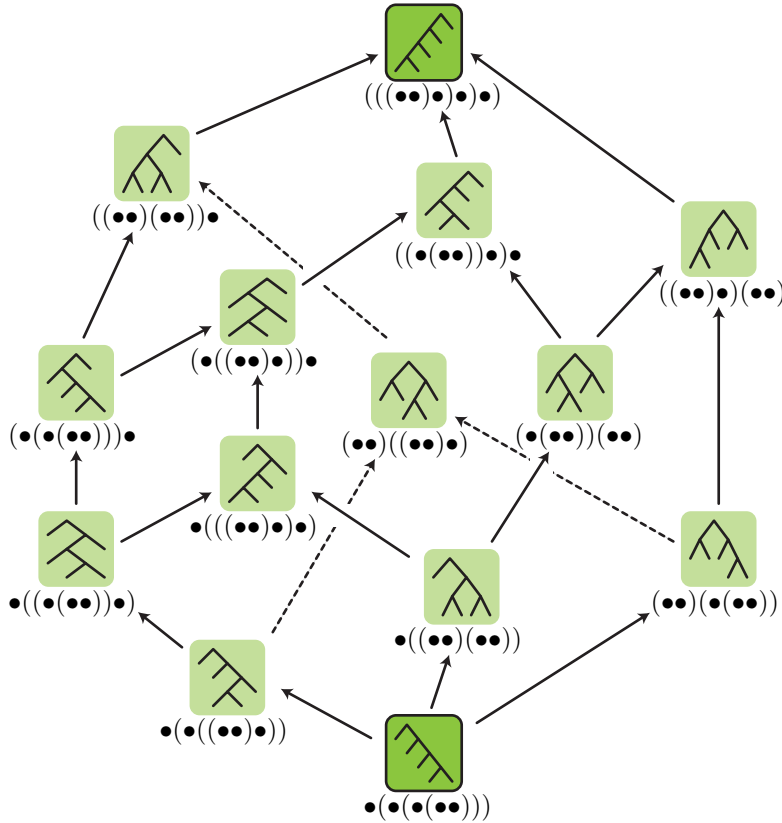
FIGURE 4. The Tamari lattice $\mathcal{T}_4$, both in terms of parenthesized expressions and binary trees.

Typical elements of F are displayed in Figure 5. In this paper, it is convenient to equip F with reversed composition, that is, fg stands for f followed by g—using the other convention simply amounts to reversing all expressions. The notation x_0 is traditional for the element of F defined by

$$x_0(t) = \begin{cases} \frac{t}{2} & \text{for } 0 \leqslant t \leqslant \frac{1}{2}, \\ t - \frac{1}{4} & \text{for } \frac{1}{2} \leqslant t \leqslant \frac{3}{4}, \\ 2t - 1 & \text{for } \frac{3}{4} \leqslant t \leqslant 1, \end{cases}$$

and x_i is used for the element that is the identity on $[0, 1 - \frac{1}{2^i}]$ and is a rescaled copy of x_0 on $[1 - \frac{1}{2^i}, 1]$—see Figure 5 again. It is easy to check that F is generated by the sequence of all elements x_i, with the presentation

(2.1) $\langle x_0, x_1, ... \mid x_{n+1}x_i = x_i x_n \text{ for } i < n \rangle.$

One deduces that F is also generated by x_0 and x_1, with the (finite) presentation

(2.2) $\langle x_0, x_1 \mid [x_0^{-1}x_1, x_0 x_1 x_0^{-1}], [x_0^{-1}x_1, x_0^2 x_1 x_0^{-2}] \rangle,$

where $[x, y]$ denotes the commutator $xyx^{-1}y^{-1}$.

The group F has many interesting algebraic and geometric properties, see [4]. Its center is trivial, the derived group $[F, F]$ is a simple group, F includes no free group of rank more than 1 (Brin–Squier [3]), its Dehn function is quadratic (Guba [17]). It is
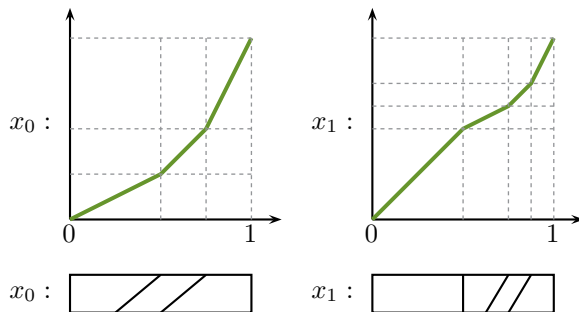
FIGURE 5. Two representations of the elements x_0 and x_1 of the Thompson group F: above, the usual graph of a function of $[0, 1]$ into itself, below, a diagram displaying the two involved dyadic decompositions of $[0, 1]$, with the source above and the target below: this simplified diagram specifies the function entirely.

not known whether F is automatic, nor whether F is amenable. The latter question has received lot of attention as F seems to lie very close to the border between amenability and non-amenability.

Owing to the developments of Section 3 below, we mention one more (simple) algebraic result, namely that F is a group of (left)-fractions, that is, there exists a submonoid of F such that every element of F can be expressed as $f^{-1}g$ with f, g in the considered submonoid.

Proposition 2.1. [4] *Define the* Thompson monoid F^+ *to be the submonoid of F generated by the elements x_i with $i \geqslant 1$. Then, as a monoid, F^+ admits the presentation* (2.1), *and F is a group of left-fractions for F^+.*

Thus F^+ consists of the elements of F that admit at least one expression in terms of the elements x_i in which no factor x_i^{-1} occurs. Although easy, Proposition 2.1 is technically significant as its leads to a unique normal form for the elements of F.

2.4. **The action of F on trees.** An element of F is determined by a pair of dyadic decompositions of the interval $[0, 1]$ specifying the intervals on which the slope has a certain value, and, from there, by a pair of trees.

To make the description precise, define a *dyadic decomposition* of $[0, 1]$ to be an increasing sequence $(t_0, ..., t_n)$ of dyadic numbers with $t_0 = 0$ and $t_n = 1$, such that no interval $[t_i, t_{i+1}]$ may contain a dyadic number with denominator less that those of t_i and t_{i+1}: for instance, $(0, \frac{1}{2}, \frac{3}{4}, 1)$ is legal, but $(0, \frac{3}{4}, 1)$ is not. Then dyadic decompositions are in one-to-one correspondence with binary rooted trees: the decomposition associated with $\bullet$ is $(0, 1)$, whereas the one associated with $T_0{}^\wedge T_1$ is the concatenation of those associated with T_0 and T_1 rescaled to fit in $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$.

As the diagram representation of Figure 5 shows, every element of the group F is entirely specified by a pair of dyadic decompositions, hence by a pair of trees. Provided adjacent intervals are gathered, this pair of decompositions (hence of trees) is unique. We shall denote by (f_-, f_+) the pair of trees associated with f. For instance, we have $1_- = 1_+ = \bullet$, and, as illustrated in Figure 6, $(x_0)_- = \bullet(\bullet\bullet)$, $(x_0)_+ = (\bullet\bullet)\bullet$, $(x_1)_- = \bullet(\bullet(\bullet\bullet))$, and $(x_1)_+ = \bullet((\bullet\bullet)\bullet)$. By construction, the trees f_- and f_+ have the same size. Moreover, we have $(f^{-1})_- = f_+$ and $(f^{-1})_+ = f_-$ as taking the inverse amounts to exchanging source and target in the diagram.

We now define a partial action of the group F on finite trees. Hereafter, we denote by $\mathcal{B}$ the family of all finite, binary, rooted trees, and by $\mathcal{B}^{\#}$ the family of all (finite, binary,
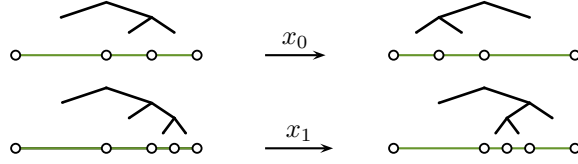
FIGURE 6. Canonical pair of trees associated with an element of the group F.

rooted) labeled trees whose leaves wear labels in $\mathbb{N}$. Thus $\mathcal{B}$ identifies with the family of all parenthesized expressions involving the single variable $\bullet$, and $\mathcal{B}^{\#}$ with the family of all parenthesized expressions involving variables from the list $\{\bullet_0, \bullet_1, ...\}$. Forgetting the labels (or the indices of variables) defines a projection of $\mathcal{B}^{\#}$ onto $\mathcal{B}$; by identifying $\bullet$ with $\bullet_0$, we can see $\mathcal{B}$ as a subset of $\mathcal{B}^{\#}$. If T is a tree of $\mathcal{B}$, we denote by $T^{\#}$ the tree of $\mathcal{B}^{\#}$ obtained by attaching to the leaves of T labels $0, 1, ...$ starting from the left.

Definition 2. A *substitution* is a map from $\mathbb{N}$ to $\mathcal{B}^{\#}$. If σ is a substitution and T is a tree in $\mathcal{B}^{\#}$, we define T^{σ} to be the tree obtained from T by replacing every i-labeled leaf of T by the tree $\sigma(i)$.

Formally, T^{σ} is recursively defined by the rules

$$(\bullet_i)^{\sigma} = \sigma(i), \qquad (T_0{}^{\wedge}T_1)^{\sigma} = T_0^{\sigma}{}^{\wedge}T_1^{\sigma}.$$

For instance, if T is $\bullet_3(\bullet_0\bullet_2)$ and we have $\sigma(0) = \bullet\bullet$ and $\sigma(2) = \sigma(3) = \bullet$, then T^{σ} is $\bullet((\bullet\bullet)\bullet)$.

Definition 3. If T, T' are labeled trees and f is an element of the Thompson group F, we say that $T * f = T'$ holds if we have $T = (f_-^{\#})^{\sigma}$ and $T' = (f_+^{\#})^{\sigma}$ for some substitution σ.

Example 2.2. First consider $f = 1$. Then we have $1_- = 1_+ = \bullet$, whence $1_-^{\#} = 1_+^{\#} = \bullet_0$. For every tree T, we have $T = (1_-^{\#})^{\sigma}$ for any substitution satisfying $\sigma(0) = T$, and, in this case, we have $(1_-^{\#})^{\sigma} = T$. So $T * 1$ is always defined and it is equal to T.

Consider now x_0. Then we have $x_{0-} = \bullet(\bullet\bullet)$, whence $x_0{}^{\#} = \bullet_0(\bullet_1\bullet_2)$. For a tree T, there exists a substitution satisfying $T = (\bullet_0(\bullet_1\bullet_2))^{\sigma}$ if and only if T can be expressed as $T_0{}^{\wedge}(T_1{}^{\wedge}T_2)$. In this case, the tree $((\bullet_0\bullet_1)\bullet_2)^{\sigma}$ is $(T_0{}^{\wedge}T_1)^{\wedge}T_2$. So $T * x_0$ is defined if and only if T is eligible for a left-rotation and, in this case, $T * x_0$ is the tree obtained from T by that left-rotation, see Figure 3.

Consider finally x_1. Arguing as above, we see that $T * x_1$ is defined if and only if T can be expressed as $T_0{}^{\wedge}(T_1{}^{\wedge}(T_2{}^{\wedge}T_3))$, in which case $T * f$ is the tree $T_0{}^{\wedge}((T_1{}^{\wedge}T_2)^{\wedge}T_3)$, that is, the tree obtained from T by a left-rotation at the right-child of the root.
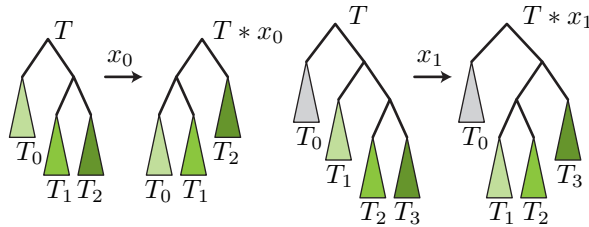
FIGURE 7. Action of x_0 and x_1 on a tree: respectively applying a left-rotation at the root, and at the right-child of the root.

The above definition specifies what can naturally be called a partial action of the group F on (labeled) trees—labels are not important here as rotations do not change their order or repeat them, but they are needed for a clean definition of substitutions.

Proposition 2.3. (i) *For every (labeled) tree T and and every element f of F, there exists at most one T' satisfying $T' = T * f$.*

(ii) *For every (labeled) tree T, we have $T * 1 = T$.*

(iii) *For every (labeled) tree T and all f, g in F, we have $(T * f) * g = T * fg$, this meaning that either both terms are defined and they are equal, or neither is defined. Moreover, for all $f_1, ..., f_n$ in F, there exists T such that $T * f_i$ is defined for each i.*

Sketch, see [11] *for details.* (i) For f in F, a given tree T can be expressed in at most one way as $(f_-^{\#})^{\sigma}$ and, as the same variables occur on both sides of the associativity law, there is in turn at most one corresponding tree $(f_+^{\#})^{\sigma}$.

Point (ii) has been established in Example 2.2. For (iii), the point is that there exists a simple rule for determining the pair of trees associated with fg. Indeed, we have $(fg)_- = f_-^{\sigma}$ and $(fg)_+ = g_+^{\tau}$, where σ and τ are minimal substitutions satisfying $f_+^{\sigma} = g_-^{\tau}$—that is, (σ, τ) is a minimal identifier for f_+ and g_-.

As for the final point, it comes from the fact that, by construction, every tree $f_-^{\#}$ has pairwise distinct labels and, therefore, a tree T can be expressed as $(f_-^{\#})^{\sigma}$ if and only if the skeleton of T (as defined in Definition 4 below) includes the skeleton of $f_-^{\#}$. Then, for $f_1, ..., f_n$ in F, one can always find a tree T whose skeleton includes those of $(f_1)_-^{\#}, ..., (f_n)_-^{\#}$. □

Proposition 2.4. *For all (labeled) trees T, T', the following are equivalent:*

(i) *One can go from T to T' using a finite sequence of rotations—that is, by applying associativity;*

(ii) *The trees T and T' have the same size, and the left-to-right enumerations of the labels in T and T' coincide;*

(iii) *There exists f in F satisfying $T' = T * f$.*

In this case, the element f involved in (iii) *is unique.*

Sketch, see [11] *for details.* The equivalence of (i) and (ii) follows from the syntactic properties of the terms occurring in the associativity law, namely that the same variables occur on both sides, in the same order.

Next, assume that T and T' are equal size trees. Then T and T' determine dyadic decompositions of $[0, 1]$, and there exists a dyadic homeomorphism f, hence an element of F, that maps the first onto the second. Provided the enumerations of the labels in T and T' coincide, we have $T' = T * f$. So (ii) implies (iii).

Conversely, we saw in Example 2.2 that the action of x_0 and x_1 is a rotation. On the other hand, we know that x_0 and x_1 generate F. Therefore, the action of an arbitrary element of f is a finite product of rotations. So (iii) implies (ii).

Finally, the uniqueness of the element f possibly satisfying $T' = T * f$ follows from the fact that the pair (T, T') determines a unique pair of dyadic decompositions of $[0, 1]$, so it directly determines the graph of the dyadic homeomorphism f. □

Proposition 2.4 states that F is the *geometry group* of associativity in the sense of [11]. A similar approach can be developed for every algebraic law, and more generally every family of algebraic laws, leading to a similar geometry monoid (a group in good cases). In the case of associativity together with commutativity, the geometry group happens to be the Thompson group V, whereas, in the case of the left self-distributivity law $x(yz) = (xy)(xz)$, the geometry group is a certain ramified extension of Artin's braid group B_{∞} [8]—also see the case of $x(yz) = (xy)(yz)$ in [9]. In the latter cases, the situation is more complicated than with associativity as, in particular, the counterparts of (i) and (ii) in Proposition 2.4 fail to be equivalent.

2.5. **The generators a_α.** Considering the action of the group F on trees invites us to introducing, beside the standard generators x_i, a new, more symmetric family of generators for F.

In order to define these elements, we need an index system for the subtrees of a tree. A common solution consists in describing the path connecting the root of the tree to the root of the considered subtree using (for instance) 0 for "forking to the left" and 1 for "forking to the right".

Definition 4. A finite sequence of 0's and 1's is called an *address*; the empty address is denoted by $\emptyset$. For T a tree and α a short enough address, the α-*subtree* of T is the part of T that lies below α. The set of all α's for which the α-subtree of T exists is called the *skeleton* of T.

Formally, the α-subtree is defined by the following recursive rules: the $\emptyset$-subtree of T is T, and, for $\alpha = 0\beta$ (*resp.* 1β), the α-subtree of T is the β-subtree of T_0 (*resp.* T_1) if T is $T_0{}^\wedge T_1$, and it is undefined otherwise.

Example 2.5. For $T = \bullet((\bullet\bullet)\bullet)$ (the rightmost example in Figure 2), the 10-subtree of T is $\bullet\bullet$, while the 01- and 111-subtrees are undefined. The skeleton of T consists of the seven addresses $\emptyset$, 0, 1, 10, 100, 101, and 11.

By definition, applying associativity in a parenthesized expression or, equivalently, applying a rotation in a tree T consists in choosing an address α in the skeleton of T and either replacing the α-subtree of T, supposed to have the form $T_0{}^\wedge(T_1{}^\wedge T_2)$, by the corresponding $(T_0{}^\wedge T_1){}^\wedge T_2$, or vice versa, see Figure 3 again. By Proposition 2.4, this rotation corresponds to a (unique) element of F.

Definition 5. For every address α, we denote by a_α the element of F whose action is a left-rotation at α. We denote by $\boldsymbol{A}$ the family of all elements a_α for α an address.

According to Example 2.2 and Figure 7, the action of x_0 is a left-rotation at the root of the tree, and, therefore, we have $x_0 = a_\emptyset$. Similarly, x_1 is left-rotation at the right-child of the root, that is, at the node with address 1, and, therefore, we have $x_1 = a_1$. More generally, all elements a_α can be expressed in terms of the generators x_i, as will be done in Subsection 2.6 below. For the moment, we simply note that iterating the argument for x_1 gives for every $i \geqslant 1$ the equality $x_i = a_{1^{i-1}}$, where 1^{i-1} denotes 11...1, $i - 1$ times 1.

The trees T such that $T * a_\alpha$ is defined are easily characterized. Indeed, a necessary and sufficient for $T * a_\alpha$ to exist is that the α-subtree of T is defined and a left-rotation can be applied to that subtree, that is, it can be expressed as $T_0{}^\wedge(T_1{}^\wedge T_2)$. This is true if and only if the addresses $\alpha 0$, $\alpha 10$, and $\alpha 11$ lie in the skeleton of T, hence actually if and only if $\alpha 10$ lies in the skeleton of T since $\beta 0$ may lie in the skeleton of a tree only if $\beta 1$ and β do. Symmetrically, $T * a_\alpha^{-1}$ is defined if and only if $\alpha 01$ lies in the skeleton of T. As a tree has a finite skeleton, there exist for every tree T finitely many addresses α such that $T * a_\alpha^{\pm 1}$ is defined, see Figure 8.

Before proceeding, we note that the forking nature of the family $\boldsymbol{A}$ naturally gives rise to a large family of shift endomorphisms of the group F.

Lemma 2.6. *For every address α, there exists a (unique) shift endomorphism sh_α of F that maps a_β to $a_{\alpha\beta}$ for every β.*

Proof. For f in F, let $\mathrm{sh}_1(f)$ denote the homeomorphism obtained by rescaling f, applying it in the interval $[\frac{1}{2}, 1]$, and completing with the identity on $[0, \frac{1}{2}]$. Then sh_1 is an endomorphism of F, and it maps x_i to x_{i+1} for every i. Moreover, for every β, the
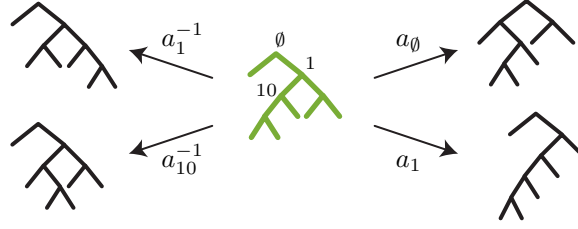
FIGURE 8. The four elements $a_\alpha^{\pm 1}$ such that $T * a_\alpha^{\pm 1}$ is defined in the case $T = \bullet(((\bullet\bullet)\bullet)(\bullet\bullet))$.

element $\mathrm{sh}_1(a_\beta)$ is the rescaled version of a_β applied in the interval $[\frac{1}{2}, 1]$. By definition, this is $a_{1\beta}$.

Symmetrically, for f in F, let $\mathrm{sh}_0(f)$ denote the homeomorphism obtained by rescaling f, applying it in the interval $[0, \frac{1}{2}]$, and completing with the identity on $[\frac{1}{2}, 1]$. Then sh_0 is an endomorphism of F, and, for every β, the element $\mathrm{sh}_0(a_\beta)$ is the rescaled version of a_β applied in the interval $[0, \frac{1}{2}]$, hence it is $a_{0\beta}$.

Finally, we recursively define sh_α for every α by $\mathrm{sh}_\emptyset = \mathrm{id}_F$ and, for $i = 0, 1$, $\mathrm{sh}_{i\alpha}(f) = \mathrm{sh}_i(\mathrm{sh}_\alpha(f))$. By construction, $\mathrm{sh}_\alpha(a_\beta) = a_{\alpha\beta}$ holds for all α, β. $\qquad \square$

2.6. Presentation of F in terms of the elements a_α.

As the family $\{x_0, x_1\}$, which is $\{a_\emptyset, a_1\}$, generates the group F, the family $\boldsymbol{A}$ generates F as well. By using the presentation (2.1) or (2.2), we could easily deduce a presentation of F in terms of the elements a_α. However, we can obtain a more natural and symmetric presentation by coming back to trees and associativity, and exploiting the geometric meaning of the elements of $\boldsymbol{A}$.

Lemma 2.7. *Say that two addresses α, β are* orthogonal, *written $\alpha \perp \beta$, if there exists γ such that α begins with $\gamma 0$ and β begins with $\gamma 1$, or vice versa. Then all relations of the following family $\boldsymbol{R}$ are satisfied in F:*

$$(2.3) \qquad\qquad a_\alpha\, a_\beta = a_\beta\, a_\alpha \qquad \text{for } \alpha \perp \beta,$$

$$(2.4) \qquad a_{\alpha 11\beta}\, a_\alpha = a_\alpha\, a_{\alpha 1\beta}, \quad a_{\alpha 10\beta}\, a_\alpha = a_\alpha\, a_{\alpha 01\beta}, \quad a_{\alpha 0\beta}\, a_\alpha = a_\alpha\, a_{\alpha 00\beta},$$

$$(2.5) \qquad\qquad a_\alpha^2 = a_{\alpha 1}\, a_\alpha\, a_{\alpha 0}.$$

Proof. By Proposition 2.4, in order to prove that two elements f, f' of F coincide, it is enough to exhibit a tree T such that $T * f$ and $T * f'$ are defined and equal.

The commutation relations of type (2.3) are trivial. If α and β are orthogonal, the α- and β-subtrees are disjoint, and the result of applying rotations (as well as any transformations) in each of these subtrees does not depend on the order. So we have

$$\mathrm{sh}_\alpha(f)\, \mathrm{sh}_\beta(g) = \mathrm{sh}_\beta(g)\, \mathrm{sh}_\alpha(f)$$

for all transformations f, g and, in particular, $a_\alpha\, a_\beta = a_\beta\, a_\alpha$.

The quasi-commutation relations of type (2.4) are more interesting. Assume that T, T' are trees and $a_\emptyset$ maps T to T'. Then, by definition, the 1-subtree of T' is a copy of the 11-subtree of T. Now, assume that f is a (partial) mapping of $\mathcal{B}$ to itself. Then, starting from T, first applying $a_\emptyset$ and then applying f to the 11-subtree leads to the same result as first applying f to the 1-subtree and then applying $a_\emptyset$, see Figure 9. Moreover, if f is a partial mapping, the result of one operation is defined if and only if the result of the other is. So, in all cases, we have

$$a_\emptyset\, \mathrm{sh}_{11}(f) = \mathrm{sh}_1(f)\, a_\emptyset.$$

Applying this to $f = a_\beta$ then gives $a_\emptyset\, a_{11\beta} = a_{1\beta}\, a_\emptyset$. Shifting by α this relation, we obtain $a_\alpha\, a_{\alpha 11\beta} = a_{\alpha 1\beta}\, a_\alpha$, the first relation of (2.4). Arguing similarly with the 0- and 10-subtrees in place of the 11-subtree, one obtains the other relations of (2.4).

Finally, the relations of (2.5) stem from the pentagon of Figure 1. As Figure 10 shows, the relation $a_\emptyset^2 = a_1\, a_\emptyset\, a_0$ is satisfied in F and, therefore, so is its shifted version $a_\alpha^2 = a_{\alpha 1}\, a_\alpha\, a_{\alpha 0}$ for every address α. $\qquad\square$
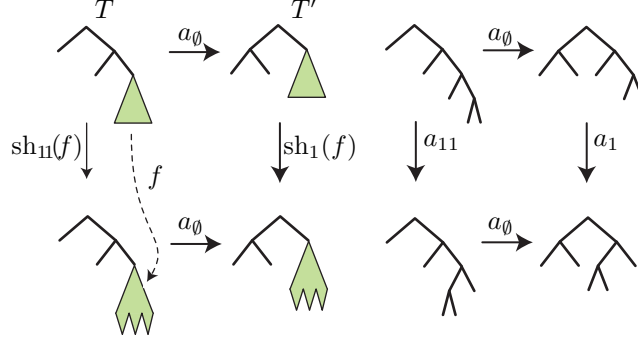

FIGURE 9. Quasi-commutation relation in F: the general scheme and one example.
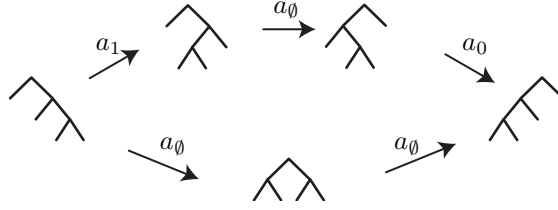

FIGURE 10. Pentagon relation in the group F.

It is then easy to check that the above relations actually exhaust the relations connecting the elements a_α in the group F.

Proposition 2.8. [6, 11] *The group F admits the presentation $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle$.*

Proof. By Lemma 2.7, the relations of $\boldsymbol{R}$ are valid in F. Conversely, to prove that these relations make a presentation, it is sufficient to show that they include the relations of a previously known presentation. This is what happens as, for $1 \leqslant i < n$, the relation $a_{1^n} a_{1^{i-1}} = a_{1^{i-1}} a_{1^{n-1}}$, which is a reformulation of the relation $x_{n+1} x_i = x_i x_n$ of (2.1), occurs in $\boldsymbol{R}$ as the first relation of (2.4) with $\alpha = 1^{i-1}$ and $\beta = 1^{n-i}$. $\qquad\square$

As an application, we compute the elements a_α in terms of the generators x_i.

Proposition 2.9. *If α is an address containing at least one 0, say $\alpha = 1^i 0^{1+i_0} 10^{i_1}...10^{i_m}$ with $m \geqslant 0$ and $i, i_0, ..., i_m \geqslant 0$, then, putting $g = x_{i+m+1}^{i_m+1} ... x_{i+2}^{i_1+1} x_{i+1}^{i_0+1}$, we have*

$$(2.6) \qquad\qquad a_\alpha = g^{-1} x_{i+m+2}^{-1} x_{i+m+1} g.$$

Proof. It is sufficient to establish the formula in the case $i = 0$ as, then, applying sh_1^i gives the general case. We use induction on (m, i_0) with respect to the lexicographical (well)-order, that is, (m', i_0') is smaller than (m, i_0) if and only if we have either $m' < m$, or $m' = m$ and $i_0' < i_0$.

Assume first $(m, i_0) = (0, 0)$, that is, $\alpha = 0$. Then the pentagon relation at $\emptyset$ gives

$$a_\alpha = a_0 = a_\emptyset^{-1} a_1^{-1} a_\emptyset^2 = x_1^{-1}(x_2^{-1} x_1)x_1,$$

which is the expected instance of (2.6). Assume now $m \geqslant 1$ and $i_0 = 0$, that is $\alpha = 010^{i_1} \ldots 10^{i_m}$. Then the quasi-commutation relation for $\emptyset$ and 01β gives

$$(2.7) \qquad a_\alpha = a_{010^{i_1} \ldots 10^{i_m}} = a_\emptyset^{-1} a_{10^{1+i_1} 10^{i_2} \ldots 10^{i_m}} a_\emptyset = x_1^{-1}(a_{10^{1+i_1} 10^{i_2} \ldots 10^{i_m}})x_1.$$

The number of non-initial symbols 1 in $0^{1+i_1} 10^{i_2} \ldots 10^{i_m}$ is $m - 1$. As $(m - 1, i_1)$ is smaller than $(m, 0)$, the induction hypothesis gives $a_{0^{1+i_1} 10^{i_2} \ldots 10^{i_m}} = g^{-1} x_{m+1}^{-1} x_m g$ with $g = x_m^{i_m+1} \ldots x_2^{i_2+1} x_1^{i_1+1}$. Using sh_1, we get $a_{10^{1+i_1} 10^{i_2} \ldots 10^{i_m}} = h^{-1} x_{m+2}^{-1} x_{m+1} h$ with $h = x_{m+1}^{i_m+1} \ldots x_3^{i_2+1} x_2^{i_1+1}$. Merging with (2.7), we deduce the expected value for a_α.

Assume finally $i_0 \geqslant 1$. Then the quasi-commutation relation for $\emptyset$ and 00β gives

$$a_\alpha = a_{0^{1+i_0} 10^{i_1} \ldots 10^{i_m}} = a_\emptyset^{-1} a_{0^{i_0} 10^{i_1} \ldots 10^{i_m}} a_\emptyset = x_1^{-1} a_{0^{i_0} 10^{i_1} \ldots 10^{i_m}} x_1.$$

The pair $(m, i_0 - 1)$ is smaller than the pair (m, i_0), so the induction hypothesis gives $a_\alpha = x_1^{-1}(g^{-1} x_{m+2}^{-1} x_{m+1} g)x_1$ with $g = x_{m+1}^{i_m+1} \ldots x_2^{i_1+1} x_1^{i_0}$, again the expected instance of (2.6). So the induction is complete. $\square$

Example 2.10. Consider $\alpha = 01100$, which corresponds to $m = 2$, and $i = i_0 = i_1 = 0$, and $i_2 = 2$. Then we find $a_\alpha = g^{-1} x_4^{-1} x_3 g$ with $g = x_3^3 x_2 x_1$, that is, a_{01100} is equal to $x_1^{-1} x_2^{-1} x_3^{-3} x_4^{-1} x_3^4 x_2 x_1$.

3. A LATTICE STRUCTURE ON THE THOMPSON GROUP F

Here comes the core of our study, namely the investigation of the submonoid F_{sym}^+ of F generated by the elements a_α. The main result is that F_{sym}^+ has the structure of a lattice when equipped with its divisibility relation, and that this lattice is closely connected with the Tamari lattices, which occur as initial sublattices.

These results are not trivial, as, in particular, determining a presentation of F_{sym}^+ is not so easy. Our approach relies on using subword reversing, a general method of combinatorial group theory that turns out to be well suited for F_{sym}^+. One of the outcomes is a new proof (one more!) of the fact that Tamari posets are lattices.

The section is organized as follows. The symmetric Thompson monoid F_{sym}^+ is introduced in Subsection 3.1, and it is investigated in Subsection 3.2 using subword reversing. The lattice structure on F_{sym}^+ and its connection with the Tamari lattices are described in Subsection 3.3. Finally, a few results about the algorithmic complexity of the reversing process are gathered in Subsection 3.4.

3.1. **The symmetric Thompson monoid F_{sym}^+.** Once new generators a_α of the Thompson group F have been introduced, it is natural to investigate the submonoid generated by these elements.

Definition 6. The *symmetric Thompson monoid* F_{sym}^+ is the submonoid of F generated by the elements a_α with α a binary address.

The family $\boldsymbol{A}$ of all elements a_α is a sort of closure of the family of standard generators x_i under all local left–right symmetries, so the above terminology is natural. Another option could be to call F_{sym}^+ the *dual Thompson monoid* as the relation of F^+ and F_{sym}^+ is reminiscent of the relation of the standard braid monoids and the dual braid monoids generated by the Birman–Ko–Lee braids.

Although straightforward, the following connection is essential for our purpose:

Lemma 3.1. *For all trees T, T', the following are equivalent*
  (i) *We have $T \leqslant_{\mathcal{T}} T'$ in the Tamari order;*
  (ii) *There exists f in F_{sym}^+ satisfying $T' = T * f$.*

Proof. By definition, $T \leqslant_{\mathcal{T}} T'$ holds if there exists a finite sequence of left-rotations transforming T into T'. Now applying the left-rotation at α is letting a_α act. $\square$

In order to investigate the monoid F_{sym}^+ and its connection with the Tamari lattices, it will be necessary to first know a presentation of F_{sym}^+. Owing to Propositions 2.1 and 2.8, the following result should not be a surprise.

Proposition 3.2. [6, 11] *The monoid F_{sym}^+ admits the presentation $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$, and F is a group of right-fractions for F_{sym}^+ (that is, every element of F can be expressed as fg^{-1} with f, g in F_{sym}^+).*

However, the proof of Proposition 3.2 is more delicate than the proof of Proposition 2.1, and no very simple argument is known.

Sketch of the proof developed in [6, 11]. In order to prove that the relations of $\boldsymbol{R}$ generate all relations connecting the elements a_α in the monoid F^+, one introduces, for every size n tree T, an explicit sequence c_T of elements a_α satisfying $C_n * c_T = T$—as will be made in the proof of Proposition 3.9 below. The point is then to show that, if $T' = T * w$ holds, then the relations of $\boldsymbol{R}$ are sufficient to establish the equivalence of $c_{T'}$ and $c_T w$. Then, if two $\boldsymbol{A}$-words u, v represent the same element of F_{sym}^+, and T is a tree such that both $T * u$ and $T * v$ are defined, the above argument shows that $c_T u$ and $c_T v$ are $\boldsymbol{R}$-equivalent, since both are $\boldsymbol{R}$-equivalent to c_{T*u}. Provided $\boldsymbol{R}$-equivalence is known to allow left-cancellation, one deduces that u and v are $\boldsymbol{R}$-equivalent, as expected. $\square$

Here we shall propose a new proof, which is more lattice-theoretic in that it exclusively relies on the so-called subword reversing method, which we shall see below is directly connected with the Tamari lattice operations. Instead of working with F_{sym}^+, we investigate the abstract monoid $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$ defined by the presentation $(\boldsymbol{A}, \boldsymbol{R})$ of Proposition 2.8. A priori, as F_{sym}^+ is generated by $\boldsymbol{A}$ and satisfies the relations of $\boldsymbol{R}$, we only know that F_{sym}^+ is a quotient of $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$.

Definition 7. Assume that M is a monoid. For f, g in M, we say that f *left-divides* g, or that g is a *right-multiple* of f, written $f \preccurlyeq g$, if $fg' = g$ holds for some g' of M. We use $\mathrm{Div}(f)$ for the family of all left-divisors of f.

It is standard that the left-divisibility relation is a partial pre-ordering. Moreover, if M contains no invertible element except 1, this partial pre-ordering is a partial ordering, that is, the conjunction of $f \preccurlyeq g$ and $g \preccurlyeq f$ implies $f = g$.

Lemma 3.3. *In order to establish Proposition 3.2, it is sufficient to prove that the monoid $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$ is cancellative and any two elements admit a common right-multiple.*

Proof. A classical result of Ore (see for instance [5]) says that, if a monoid M is cancellative and any two elements of M admit a common right-multiple, then M embeds in a group of right-fractions G. Moreover, if M admits the presentation $\langle A \mid R \rangle^+$, then G admits the presentation $\langle A \mid R \rangle$. So, if the hypotheses of the lemma are satisfied, then the monoid $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$ embeds in a group of fractions that admits the presentation $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle$. By Proposition 2.8, the group $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle$ is the group F. Therefore, $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$ is isomorphic to the submonoid of F generated by $\boldsymbol{A}$, that is, to F_{sym}^+. Hence F_{sym}^+ admits the expected presentation, and F is a group of right-fractions for F_{sym}^+. $\square$

3.2. **Subword reversing.** In order to apply the strategy of Lemma 3.3, we have to prove that the presented monoid $\langle \boldsymbol{A} \,|\, \boldsymbol{R} \rangle^+$ is cancellative and any two elements of $\langle \boldsymbol{A} \,|\, \boldsymbol{R} \rangle^+$ admit a common right-multiple. The subword reversing method [7, 10] proves to be relevant. We recall below the basic notions, and refer to [13] or [14, Section II.4] for a more complete description.

Hereafter, words in an alphabet A are called *(positive) A-words*, whereas words in the alphabet $A \cup A^{-1}$, where A^{-1} consists of a copy a^{-1} for each letter a of A, are called *signed A-words*. We say that a group presentation (A, R) is *positive* if all relations in R have the form $u = v$ where u and v are nonempty positive A-words. We denote by $\langle A \,|\, R \rangle^+$ and by $\langle A \,|\, R \rangle$ the monoid and the group presented by (A, R), respectively, and we use $\equiv_R^+$ (*resp.* $\equiv_R$) for the congruence on positive A-words (*resp.* on signed A-words) generated by R. Finally, for w a signed A-word, we denote by $\overline{w}$ the element of $\langle A \,|\, R \rangle$ represented by w, that is, the $\equiv_R$-class of w.

Definition 8. Assume that (A, R) is a positive presentation. If w, w' are signed A-words, we say that w is *right-R-reversible* to w' in one step if w' is obtained from w either by deleting some length 2 subword $a^{-1}a$ or by replacing some length 2 subword $a^{-1}b$ with a word vu^{-1} such that $av = bu$ is a relation of R. We write $w \curvearrowright_R w'$ if w is right-R-reversible to w' in finitely many steps.

The principle of right-R-reversing is to use the relations of R to push the negative letters (those with exponent -1) to the right, and the positive letters (those with exponent $+1$) to the left. The process can be visualized in diagrams as in Figure 11.

Example 3.4. Consider the presentation $(\boldsymbol{A}, \boldsymbol{R})$, which is positive. Let w be the signed $\boldsymbol{A}$-word $a_1^{-1} a_\emptyset a_{00}^{-1} a_1$. Then w contains two negative–positive length 2 subwords, namely $a_1^{-1} a_\emptyset$ and $a_{00}^{-1} a_1$. There exists in $\boldsymbol{R}$ a unique relation of the form $a_1 \ldots = a_\emptyset \ldots$, namely $a_1 a_\emptyset a_0 = a_\emptyset^2$, and a unique relation $a_{00} \ldots = a_1 \ldots$, namely $a_{00} a_1 = a_1 a_{00}$. Therefore, there exists two ways to right-$\boldsymbol{R}$-reverse w, namely replacing $a_1^{-1} a_\emptyset$ with $a_\emptyset a_{00}^{-1} a_\emptyset^{-1}$ and obtaining $w_1 = a_\emptyset a_0 a_\emptyset^{-1} a_1$, or replacing $a_{00}^{-1} a_1$ with $a_1 a_{00}^{-1}$ and obtaining $w_1' = a_1^{-1} a_\emptyset a_1 a_{00}^{-1}$. The words w_1 and w_1' each contain a unique negative–positive length 2 subword, and reversing it leads in both cases to $w_2 = a_\emptyset a_0 a_\emptyset^{-1} a_1 a_{00}^{-1}$. The word w_2 contains a unique negative–positive length two subword and reversing it leads to $w_3 = a_\emptyset a_0 a_\emptyset a_0^{-1} a_\emptyset^{-1} a_{00}^{-1}$. As the latter word contains no negative–positive subword, no further right-reversing is possible. See Figure 11.

It is easy to see that, if (A, R) is a positive presentation and w, w' are signed A-words, then $w \curvearrowright_R w'$ implies $w \equiv_R w'$ and that, if u, v, u', v' are positive A-words, then $u^{-1}v \curvearrowright_R v'u'^{-1}$ implies $uv' \equiv_R^+ vu'$. In particular, using ε for the empty word,

$$(3.1) \qquad u^{-1}v \curvearrowright_R \varepsilon \quad \text{implies} \quad u \equiv_R^+ v.$$

In general, (3.1) need not be an equivalence, but it turns out that this is the interesting situation, in which case the presentation (A, R) is said to be *complete with respect to right-reversing*. Roughly speaking, a presentation is complete with respect to right-reversing if right-reversing always detects equivalence. The important point here is that the presentation $(\boldsymbol{A}, \boldsymbol{R})$ has this property.

Lemma 3.5. [6] *The presentation $(\boldsymbol{A}, \boldsymbol{R})$ is complete with respect to right-reversing.*
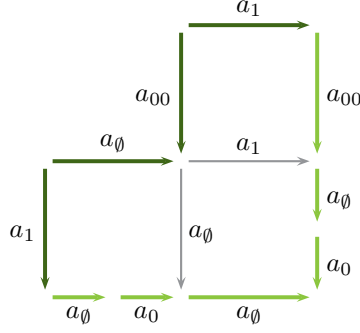
FIGURE 11. Right-$\boldsymbol{R}$-reversing of the signed $\boldsymbol{A}$-word $a_1^{-1}a_\emptyset a_{00}^{-1}a_1$: we draw the initial word as a zigzag path (here in dark green) from SW to NE by associating with every letter a a horizontal arrow labeled a and every letter a^{-1} a vertical arrow labeled a (crossed in the wrong direction); then reversing $a^{-1}b$ to vu^{-1} corresponds to closing the open pattern made by a vertical a-arrow and a horizontal b-arrow with the same source by adding horizontal arrows labeled v and vertical arrows labeled u; the final word corresponds to the rightmost path from the SW corner to the NE corner, here $a_\emptyset a_0 a_\emptyset a_0^{-1}a_\emptyset^{-1}a_{00}^{-1}$ (light green).

Sketch. By [13, Proposition 2.9], a sufficient condition for a positive presentation (A, R) to be complete with respect to right-reversing is that (A, R) satisfies

(3.2)

    (i) There exists a $\equiv_R^+$-invariant map λ from positive A-words to $\mathbb{N}$ satisfying $\lambda(uv) \geqslant \lambda(u) + \lambda(v)$ for all u, v and $\lambda(a) \geqslant 1$ for a in A, and

    (ii) For all a, b, c in A and all positive A-words u, v, if $a^{-1}cc^{-1}b \curvearrowright_R vu^{-1}$ holds, then $v^{-1}a^{-1}bu \curvearrowright_R \varepsilon$ holds as well.

We claim that $(\boldsymbol{A}, \boldsymbol{R})$ satisfies (3.2). As for (i), we cannot use for λ the length of words, as it is not $\equiv_R^+$-invariant: in the pentagon relation, the length 2 word $a_\emptyset^2$ is $\equiv_R^+$-equivalent to the length 3 word $a_1 a_\emptyset a_0$. Now, for T a tree, let $\mu(T)$ be the total number of 0's occurring in the addresses of the leaves of T: for instance, we have $\mu(\bullet(\bullet\bullet)) = 2$ and $\mu((\bullet\bullet)\bullet) = 3$, as the leaves of $\bullet(\bullet\bullet)$ have addresses 0, 10, 11, with two 0's, and those of $(\bullet\bullet)\bullet$ have addresses 00, 01, 1, with three 0's. Then put

(3.3) $$\lambda(w) = \mu(\overline{w}_+) - \mu(\overline{w}_-).$$

For instance, if w is $a_\emptyset$, the trees $\overline{w}_-$ and $\overline{w}_+$ are $\bullet(\bullet\bullet)$ and $(\bullet\bullet)\bullet$, and one finds $\lambda(a_\emptyset) = 3 - 2 = 1$. A similar argument gives $\lambda(a_\alpha) = 1$ for every address α. More generally, one easily checks that $T \leqslant_\tau T'$ implies $\mu(T') \geqslant \mu(T)$. Hence the function λ takes values in $\mathbb{N}$. Moreover, a counting argument shows that, in the previous situation, $\mu(T'^\sigma) - \mu(T^\sigma) \geqslant \mu(T') - \mu(T)$ holds for every substitution σ. If u and v are positive $\boldsymbol{A}$-words, then, as seen in the proof of Proposition 2.3, we have $\overline{(uv)}_- = \overline{u}_-^\sigma$ and $\overline{(uv)}_+ = \overline{v}_+^\tau$ for some substitutions σ, τ satisfying $\overline{u}_+^\sigma = \overline{v}_-^\tau$. We deduce

$$\lambda(uv) = \mu(\overline{uv}_+) - \mu(\overline{uv}_-) = \mu(\overline{v}_+^\tau) - \mu(\overline{u}_-^\sigma) = \mu(\overline{v}_+^\tau) - \mu(\overline{v}_-^\tau) + \mu(\overline{u}_+^\sigma) - \mu(\overline{u}_-^\sigma)$$
$$\geqslant \mu(\overline{v}_+) - \mu(\overline{v}_-) + \mu(\overline{u}_+) - \mu(\overline{u}_-) = \lambda(v) + \lambda(u).$$

As for (ii), the problem is to check that, whenever α, β, γ are addresses and the signed word $a_\alpha^{-1}a_\gamma a_\gamma^{-1}a_\beta$ is right-$\boldsymbol{R}$-reversible to some positive–negative word vu^{-1}, then $v^{-1}a_\alpha^{-1}a_\beta u$ is right-$\boldsymbol{R}$-reversible to the empty word. The systematic verification

seems tedious. Actually it is not. First, what matters is the mutual position of the addresses α, β, γ with respect to the prefix ordering, and only finitely many patterns may occur. Next, for every pair of addresses α, β, there exists in $\boldsymbol{R}$ exactly one relation of the form $a_\alpha... = a_\beta...$, which implies that, for every signed $\boldsymbol{A}$-word w, there exists at most one pair of positive $\boldsymbol{A}$-words u, v such that w is right-$\boldsymbol{R}$-reversible to vu^{-1}. Finally, all instances involving quasi-commutation relations turn out to be automatically verified. So, the only critical cases are those corresponding to the triple of addresses $\emptyset, 1, 11$ and its translated and permuted copies, and a direct verification is then easy. For instance, the reader can see on Figure 12 that we have $a_\emptyset^{-1}a_1a_1^{-1}a_{11} \curvearrowright_{\boldsymbol{R}} a_\emptyset^2 a_{00}^{-1}a_0^{-1}a_\emptyset^{-1}a_{01}^{-1}a_1^{-1}$ and $a_\emptyset^{-3}a_{11}a_1a_{10}a_\emptyset a_0 a_{00} \curvearrowright_{\boldsymbol{R}} \varepsilon$. □
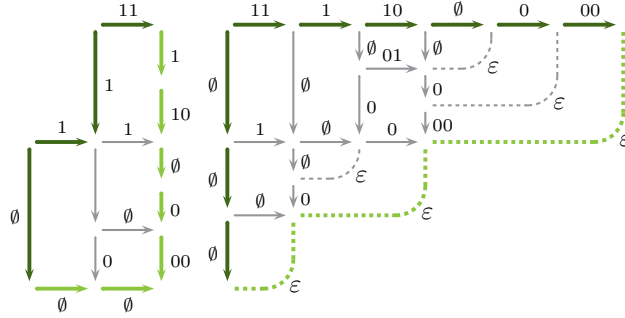

FIGURE 12. Proof of Lemma 3.5: $a_\emptyset^{-1}a_1a_1^{-1}a_{11}$ is right-$\boldsymbol{R}$-reversible to $a_\emptyset^2 a_{00}^{-1}a_0^{-1}a_\emptyset^{-1}a_{01}^{-1}a_1^{-1}$ (left), and $a_\emptyset^{-3}a_{11}a_1a_{10}a_\emptyset a_0 a_{00}$ is right-$\boldsymbol{R}$-reversible to the empty word; dotted lines represent the empty word that appears when a pattern $a_\alpha^{-1}a_\alpha$ is reversed.

Once a positive presentation is known to be complete with respect to right-reversing, it is easy to deduce properties of the associated monoid.

Proposition 3.6. *The monoid $\langle \boldsymbol{A} \,|\, \boldsymbol{R} \rangle^+$ is left-cancellative.*

Proof. By [13, Proposition 3.1], if (A, R) is a positive presentation that is complete with respect to right-reversing, a sufficient condition for the monoid $\langle A \,|\, R \rangle^+$ to be left-cancellative is that

(3.4) R contains no relation of the form $au = av$ with a in A and $u \neq v$.

By definition, $\boldsymbol{R}$ satisfies (3.4). Hence the monoid $\langle \boldsymbol{A} \,|\, \boldsymbol{R} \rangle^+$ is left-cancellative. □

As for right-cancellation, no new computation is needed as we can exploit the symmetries of $\boldsymbol{R}$. First, we introduce a counterpart of right-reversing where the roles of positive and negative letters are exchanged.

Definition 9. Assume that (A, R) is a positive presentation. If w, w' are signed A-words, we say that w is *left-R-reversible* to w' in one step if w' is obtained from w either by deleting some length 2 subword aa^{-1} or by replacing some length 2 subword ab^{-1} with a word $u^{-1}v$ such that $ua = vb$ is a relation of R. We write $w \curvearrowleft_R w'$ if w is left-R-reversible to w' in finitely many steps.

Of course, properties of left-reversing are symmetric to those of right-reversing.

Proposition 3.7. *The monoid $\langle \boldsymbol{A} \,|\, \boldsymbol{R} \rangle^+$ is right-cancellative.*

Proof. The argument is symmetric to the one for Proposition 3.6, and relies on first proving that $(\boldsymbol{A}, \boldsymbol{R})$ is, in an obvious sense, complete with respect to left-reversing. Due to the symmetries of $\boldsymbol{R}$, this is easy. Indeed, for w a signed $\boldsymbol{A}$-word, let $\widetilde{w}$ denote the word obtained by reading the letters of w from right to left, and exchanging 0 and 1 everywhere in the indices of the letters a_α. For instance, $\widetilde{a_{110}a_\emptyset}$ is $a_\emptyset a_{001}$. A direct inspection shows that the family $\widetilde{\boldsymbol{R}}$ of all relations $\widetilde{u} = \widetilde{v}$ for $u = v$ a relation of $\boldsymbol{R}$ is $\boldsymbol{R}$ itself. It follows that, for all signed $\boldsymbol{A}$-words w, w', the relations $w \curvearrowright_{\boldsymbol{R}} w'$ and $\widetilde{w} \curvearrowleft_{\boldsymbol{R}} \widetilde{w'}$ are equivalent. Then, as $w \mapsto \widetilde{w}$ is an alphabetical anti-automorphism, the completeness of $(\boldsymbol{A}, \boldsymbol{R})$ with respect to right-reversing implies the completeness of $(\boldsymbol{A}, \widetilde{\boldsymbol{R}})$, hence of $(\boldsymbol{A}, \boldsymbol{R})$, with respect to left-reversing. As the right counterpart of (3.4) is satisfied, we deduce that the monoid $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$ is right-cancellative. $\square$

In order to complete the proof of Proposition 3.2 using the strategy of Lemma 3.3, we still need to know that any two elements of the monoid $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$ admit a common right-multiple. Using the action on trees, it is easy to prove that result in F_{sym}^+. But this is not sufficient here as we do not know yet that F_{sym}^+ is isomorphic to $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$. We appeal to right-reversing once more.

Proposition 3.8. *Any two elements of $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$ admit a common right-multiple.*

Proof. If (A, R) is a positive presentation, saythat right-R-reversing is *terminating* if, for all positive A-words u, v, there exist positive A-words u', v' satisfying $u^{-1}v \curvearrowright_{\boldsymbol{R}} v'u'^{-1}$. We noted that the latter relation implies $uv' \equiv_R^+ vu'$, thus implying that, in the monoid $\langle A \mid R \rangle^+$, the elements represented by u and v admit a common right-multiple. So, in order to establish the proposition, it is sufficient to prove that right-$\boldsymbol{R}$-reversing is terminating, a non-trivial question as, because of the pentagon relations, the length of the words may increase under right-reversing, and there might exist infinite reversing sequences—try right-reversing of $a^{-1}ba$ in the presentation $(a, b, ab = b^2a)$.

Now, by [13, Proposition 3.11], if (A, R) is a positive presentation, a sufficient condition for right-R-reversing to be terminating is that (A, R) satisfies

(3.5)
> (i) For all a, b in A, there is exactly one relation $a \ldots = b \ldots$ in R, and
>
> (ii) There exists a family $\widehat{A}$ of positive A-words that includes A and is *closed under right-R-reversing*, this meaning that, for all u, v in $\widehat{A}$, there exist u', v' in $A^{\#} \cup \{\varepsilon\}$ satisfying $u^{-1}v \curvearrowright_{\boldsymbol{R}} v'u'^{-1}$.

We claim that $(\boldsymbol{A}, \boldsymbol{R})$ satisfies (3.5). Indeed, (i) follows from an inspection of $\boldsymbol{R}$. As for (ii), let us put

$$\hat{a}_{\alpha,r} = a_\alpha a_{\alpha 0} \ldots a_{\alpha 0^{r-1}}$$

for α an address and $r \geqslant 1$, see Figures 14 and 19 for an illustration of the action of $\hat{a}_{\alpha,r}$ on trees. Then the family $\widehat{\boldsymbol{A}}$ of all words $\hat{a}_{\alpha,r}$ includes $\boldsymbol{A}$ as we have $a_\alpha = \hat{a}_{\alpha,1}$ for every α, and it is closed under right-$\boldsymbol{R}$-reversing as we find

(3.6)
$$\hat{a}_{\beta,s}^{-1}\, \hat{a}_{\alpha,r} \quad \curvearrowright_{\boldsymbol{R}} \quad \begin{cases} \hat{a}_{0^{s-r},s-r} & \text{for } \beta = \alpha \text{ with } r < s, \\ \hat{a}_{\alpha,r}\, \hat{a}_{\beta,s}^{-1} & \text{for } \beta \perp \alpha, \\ \hat{a}_{\alpha,r}\, \hat{a}_{\alpha 0^{r+1}\gamma,s}^{-1} & \text{for } \beta = \alpha 0\gamma, \\ \hat{a}_{\alpha,r}\, \hat{a}_{\alpha 0^r 1\gamma,s}^{-1} & \text{for } \beta = \alpha 10^r\gamma, \\ \hat{a}_{\alpha,r}\, \hat{a}_{\alpha 0^i 1\gamma,s}^{-1} & \text{for } \beta = \alpha 10^i 1\gamma \text{ with } i < r, \\ \hat{a}_{\alpha,r+s}\, \hat{a}_{\alpha 0^i,s}^{-1} & \text{for } \beta = \alpha 10^i \text{ with } i < r, \end{cases}$$

see Figure 13. Note that $\widehat{\boldsymbol{A}}$ is the smallest family that includes $\boldsymbol{A}$ and is closed under right-$\boldsymbol{R}$-reversing as the last type of relation in (3.6) inductively forces any such family to contain $\hat{a}_{\alpha,r}$ for every r. □
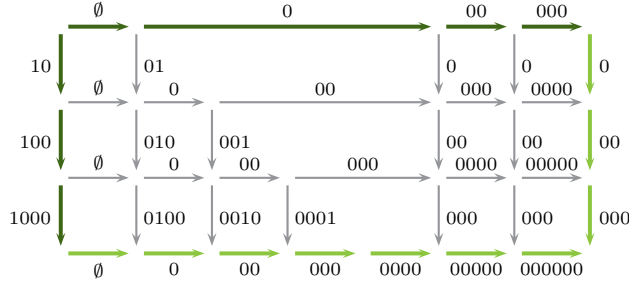

FIGURE 13. Closure of the family $\widehat{\boldsymbol{A}}$ under right-reversing: $\hat{a}_{10,3}^{-1}\hat{a}_{\emptyset,4}$ reverses to $\hat{a}_{\emptyset,7}\hat{a}_{0,3}^{-1}$, which corresponds to the last relation in (3.6) with $\alpha = \emptyset$, $r = 4$, $s = 3$, $i = 1$ (the letter "a" has been skipped everywhere).

In terms of the generators $\hat{a}_{\alpha,r}$, the pentagon relation can be expressed as $a_{\emptyset}^2 = a_1\hat{a}_{\emptyset,2}$, with both sides of length 2. The last type in (3.6) corresponds to an extended pentagon relation $\hat{a}_{\alpha,r}\,\hat{a}_{\alpha 0^i,s} = \hat{a}_{\alpha 10^i,s}\,\hat{a}_{\alpha,r+s}$ for all r,s,i with $i < r$, whose counterpart in terms of tree rotation is displayed in Figure 14.
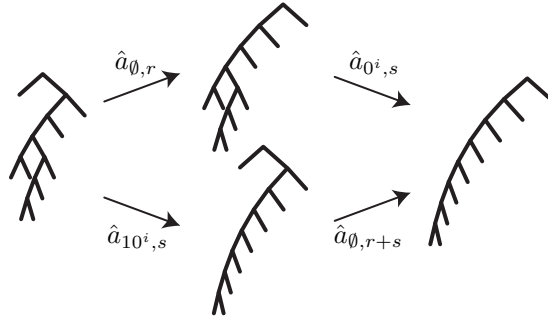

FIGURE 14. Extended pentagon relation $\hat{a}_{\alpha,r}\,\hat{a}_{\alpha 0^i,s} = \hat{a}_{\alpha 10^i,s}\,\hat{a}_{\alpha,r+s}$, here for $\alpha = \emptyset$, $r = 4$, $s = 3$, $i = 2$.

We thus established that the monoid $\langle \boldsymbol{A} \,|\, \boldsymbol{R}\rangle^+$ satisfies the conditions of Lemma 3.3 and, therefore, the proof of Proposition 3.2 is complete.

3.3. **The lattice structure of $F_{\mathbf{sym}}^+$.** Here comes the central point, namely the connection between the right-divisibility relation of the monoid F_{sym}^+, which we now know admits the presentation $\langle \boldsymbol{A} \,|\, \boldsymbol{R}\rangle^+$, and the Tamari posets. We recall that $\mathrm{Div}(f)$ denotes the family of all left-divisors of f.

Proposition 3.9. *For every $n \geqslant 1$, the subposet $(\mathrm{Div}(a_{\emptyset}^{n-1}), \preccurlyeq)$ of $(F_{sym}^+, \preccurlyeq)$ is isomorphic to the Tamari poset $(\mathcal{T}_n, \leqslant_{\mathcal{T}})$. The poset $(\bigcup_n \mathrm{Div}(a_{\emptyset}^n), \preccurlyeq)$ is isomorphic to the Tamari poset $(\mathcal{T}_\infty, \leqslant_{\mathcal{T}})$.*

Proof. An immediate induction gives the equality $C_n * a_{\emptyset}^{n-1} = \widetilde{C}_n$ for every n, that is, the element $a_{\emptyset}^{n-1}$ of F_{sym}^+ maps the right-comb C_n to the left-comb $\widetilde{C}_n$. Hence $C_n * f$ is

defined for every element f of F_{sym}^+ that left-divides $a_\emptyset^{n-1}$. Thus, as C_n belongs to $\mathcal{T}_n$ and the action of F_{sym}^+ preserves the size of the trees, we obtain a well defined map

$$(3.7) \qquad\qquad\qquad\qquad I_n : f \mapsto C_n * f$$

of $\text{Div}(a_\emptyset^{n-1})$ into $\mathcal{T}_n$. By Proposition 2.4, the map I_n is injective. On the other hand, we claim that I_n is surjective. To prove it, it suffices to exhibit, for every size n tree T, an element of F_{sym}^+ that maps the right-comb C_n to T. Now, for every tree T, define two elements c_T, c_T' of F_{sym}^+ by the recursive rules:

$$(3.8) \qquad c_T = \begin{cases} 1 \\ c_{T_0}' \, \text{sh}_1(c_{T_1}') \, a_\emptyset \end{cases} \qquad c_T' = \begin{cases} 1 & \text{for } T \text{ of size } 0, \\ c_{T_0}' \, \text{sh}_1(c_{T_1}') \, a_\emptyset & \text{for } T = T_0{}^\wedge T_1. \end{cases}$$

For every size n tree T and every $p \geq 1$, we have $C_n * c_T = T$ and $C_{n+p} * c_T' = T^\wedge C_p$, as shows an induction on T: everything is obvious for $T = \bullet$, and, for $T = T_0{}^\wedge T_1$, it suffices to follow the diagrams of Figure 15. Note that introducing both c_T and c_T' is necessary for the induction. However, the connection $c_T' = c_T \, a_{1^{i-1}} ... a_1 a_\emptyset$, where i is the length of the rightmost branch in T, is easy to check.

Thus I_n is a bijection of $\text{Div}(a_\emptyset^{n-1})$ onto $\mathcal{T}_n$. Moreover, I_n is compatible with the orderings. Indeed, assume $f \preccurlyeq g$, say $fg' = g$. Then, by Proposition 2.3, we have $(C_n * f) * g' = C_n * g$, whence $C_n * f \leqslant_\mathcal{T} C_n * g$ by Lemma 3.1. This completes the proof that $(\text{Div}(a_\emptyset^{n-1}), \preccurlyeq)$ is isomorphic to the Tamari poset $(\mathcal{T}_n, \leqslant_\mathcal{T})$.

As for $\mathcal{T}_\infty$, we observe that, for every n, we have $C_{n+1} = C_n^{\#\sigma}$ where σ is the substitution that maps $0, ..., n-1$ to $\bullet$ and n to $\bullet\bullet$. On the other hand, by definition, $\text{Div}(a_\emptyset^{n-1})$ is an initial segment of $\text{Div}(a_\emptyset^n)$ and, for every f in $\text{Div}(a_\emptyset^{n-1})$, we have

$$C_{n+1} * f = C_n^{\#\sigma} * f = (C_n * f)^{\#\sigma},$$

hence $I_{n+1}(f) = \iota_n(I_n(f))$. It follows that the family $(I_n)_{n \geqslant 1}$ induces a well defined map I_∞ of $\bigcup_n \text{Div}(a_\emptyset^n)$ into $\mathcal{T}_\infty$. The map I_∞ is injective because I_n is, it is surjective as, by definition, $\mathcal{T}_\infty$ is the limit of the directed system $(\mathcal{T}_n, \iota_n)$, and it preserves the orderings as I_n does. $\qquad\square$
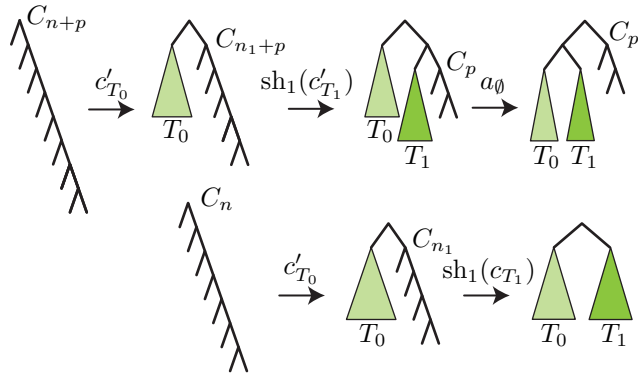

FIGURE 15. For T a size n tree, c_T describes how to construct T from the right-comb C_n, and c_T' describes how to construct $T^\wedge C_p$ from C_{n+p}; the figure illustrates the recursive definition of c_T' (above) and c_T (below) for $T = T_0{}^\wedge T_1$, with n_1 denoting the size of T_1.

Remark 3.10. The subset $\bigcup_n \mathrm{Div}(a_\emptyset^n)$ involved in Proposition 3.9 is a proper subset of F_{sym}^+ as, for instance, it contains no a_α such that 0 occurs in α: indeed, in this case, $C_n * a_\alpha$ is not defined, whereas $C_n * f$ is defined for every f left-dividing $a_\emptyset^{n-1}$.

The connection of Proposition 3.9 can be used in both directions. If we take for granted that the Tamari posets are lattices, we deduce that the subsets $(\mathrm{Div}(a_\emptyset^{n-1}), \preccurlyeq)$ of $(F_{\mathrm{sym}}^+, \preccurlyeq)$ must be lattices as well, that is, with the usual terminology of left-divisibility relation, that any two elements of $\bigcup_n \mathrm{Div}(a_\emptyset^n)$ admit a *least common right-multiple*, or *right-lcm*, and a *greatest common left-divisor*, or *left-gcd*.

On the other hand, if we have a direct proof that $(F_{\mathrm{sym}}^+, \preccurlyeq)$ is a lattice, then the isomorphism of Proposition 3.9 provides a new proof of the lattice property for the Tamari posets. This is what happens.

Proposition 3.11. *The poset $(F_{sym}^+, \preccurlyeq)$ is a lattice.*

Corollary 3.12. *For every n, the Tamari poset $(\mathcal{T}_n, \leqslant)$ is a lattice.*

To establish Proposition 3.11, we once again appeal to subword reversing.

of Proposition 3.11. By [13, Proposition 3.6], if (A, R) is a positive presentation that is complete with respect to right-reversing, a sufficient condition for any two elements of $\langle A \mid R \rangle^+$ that admit a common right-multiple to admit a right-lcm is that (A, R) satisfies Condition (i) of (3.5); moreover, in this case, the right-lcm of the elements represented by two A-words u, v is represented by uv' and vu', where u', v' are the positive A-words for which $u^{-1}v \curvearrowright_R v'u'^{-1}$ holds.

Now, as already noted, $(\boldsymbol{A}, \boldsymbol{R})$ satisfies (3.5). Hence any two elements of F_{sym}^+ that admit a common right-multiple admit a right-lcm. On the other hand, by Proposition 3.8, any two elements of $\langle \boldsymbol{A} \mid \boldsymbol{R} \rangle^+$, that is, of F_{sym}^+, admit a common right-multiple. Hence any two elements of F_{sym}^+ admit a right-lcm. In other words, any two elements in the poset $(F_{\mathrm{sym}}^+, \preccurlyeq)$ admit a least upper bound.

As for left-gcd's, we can argue as follows. Let $\widetilde{\preccurlyeq}$ denote the right-divisibility relation, which is the binary relation such that $f \mathrel{\widetilde{\preccurlyeq}} g$ holds if and only if we have $g'f = g$ for some g' (the difference with $\preccurlyeq$ is that, here, f appears on the right and not on the left). Then we have the derived notions of a left-lcm and a right-gcd. An easy general result says that, if f, g, f', g' are elements of a monoid and satisfy $fg' = gf'$, then f and g admit a left-gcd if and only if f' and g' admit a left-lcm. By Proposition 3.8, any two elements of F_{sym}^+ admit a common right-multiple and so, it suffices to show that any two elements of F_{sym}^+ admit a left-lcm to deduce that they admit a left-gcd. Now, the existence of left-lcm's in F_{sym}^+ follows from the properties of left-$\boldsymbol{R}$-reversing, which we have seen in the proof of Proposition 3.7 are similar to those of right-$\boldsymbol{R}$-reversing. $\square$

Remark 3.13. Another way of deducing the existence of left-gcd's from that of right-lcm's is to use Noetherianity properties. The existence of the function λ of (3.3) implies that a set $\mathrm{Div}(f)$ contains no infinite increasing sequence $f_1 \prec f_2 \prec \dots$ in F_{sym}^+. For all f, g, the family $\mathrm{Div}(f) \cap \mathrm{Div}(g)$ is nonempty as it contains 1, and, by Noetherianity, it contains a $\preccurlyeq$-maximal element, which must be a left-gcd of f and g.

3.4. **Computing the operations.** We conclude this section with results about the algorithmic complexity of subword reversing in F_{sym}^+. Here we concentrate on space complexity, namely bounds on the length of words; it would be easy to state analogous bounds on the number of reversing steps, hence for time complexity.

Proposition 3.14. *If w, w' are signed $\boldsymbol{A}$-words, $w \curvearrowright_{\boldsymbol{R}} w'$ implies $|w'| \leqslant |w|^2/4 + |w|$. More precisely, we have $|w'| \leqslant p + q + pq$ if w contains p positive letters and q negative letters. These bounds are sharp.*

Proof. By construction, the $\boldsymbol{R}$-reversing steps in the right-$\boldsymbol{R}$-reversing of w to w' can be gathered into $\widehat{\boldsymbol{R}}$-reversing steps, which are at most pq in number. Consider the sum of the indices r of the involved generators $\hat{a}_{\alpha,r}$. Each $\boldsymbol{R}$-reversing step increases this sum by 1 at most (in the case of a pentagon relation), so the total sum in the final $p+q$ generators $\hat{a}_{\alpha,r}$ is at most $p+q+pq$. So, when the generators $\hat{a}_{\alpha,r}$ are decomposed as products of a_α's, at most $p+q+pq$ of the latter occur.

The bound is sharp, as an easy induction gives

$$(a_{1^{p-1}}\dots a_1 a_\emptyset)^{-1}\ a_{1^p}^q\ \curvearrowright_{\boldsymbol{R}}\ a_\emptyset^q\ (\hat{a}_{1^{p-1},q}\dots\hat{a}_{1,q}\hat{a}_{\emptyset,q})^{-1},$$

a word of length $p+q$ that is right-$\boldsymbol{R}$-reversible to a word of length $p+q+pq$. $\qquad\square$

Other upper bounds can be obtained by using the action of F_{sym}^+ on trees. To state the result, it is convenient to introduce the following natural terminology.

Definition 10. For every signed $\boldsymbol{A}$-word w, the *right-numerator* $N_r(w)$ and the *right-denominator* $D_r(w)$ of w are the unique $\boldsymbol{A}$-words satisfying $w\curvearrowright_{\boldsymbol{R}} N_r(w)D_r(w)^{-1}$. Symmetrically, the *left-numerator* $N_\ell(w)$ and the *left-denominator* $D_r(w)$ of w are the unique $\boldsymbol{A}$-words satisfying $w\curvearrowleft_{\boldsymbol{R}} D_\ell(w)^{-1}N_\ell(w)$.

As left- and right-$\boldsymbol{R}$-reversings are terminating, the positive $\boldsymbol{A}$-words $N_r(w)$, $D_r(w)$, $N_\ell(w)$, and $D_\ell(w)$ exist for every signed $\boldsymbol{A}$-word w.

Proposition 3.15. *Assume that w is a signed $\boldsymbol{A}$-word and $T*w$ is defined for some size n tree T. Then we have*

$$(3.9)\qquad \max(|N_\ell(w)|+|D_r(w)|,|N_\ell(w)|+|D_r(w)|)\leqslant(n-1)(n-2)/2.$$

In order to establish Proposition 3.15, we need a preliminary result about the action of $\boldsymbol{A}$-words on trees. First, if T is a tree and w is a signed $\boldsymbol{A}$-word, we say that $T*w$ is defined if $T*\overline{u}$ is defined for every prefix u of w. Now, if two signed $\boldsymbol{A}$-words w,w' represent the same element of F, the hypothesis that $T*w$ is defined for some tree T does not guarantee that $T*w'$ is also defined: for instance, $T*\varepsilon$ is always defined, but $T*a_\alpha^{-1}a_\alpha$ is not. However, this cannot happen with reversing.

Lemma 3.16. *Assume that w,w' are signed-$\boldsymbol{A}$-words and w is right- or left-$\boldsymbol{R}$-reversible to w'. Then, for every tree T such that $T*w$ is defined, $T*w'$ is defined as well.*

Proof. The problem with arbitrary equivalences is that new pairs $a_\alpha^{-1}a_\alpha$ or $a_\alpha a_\alpha^{-1}$ may be created. This however is impossible in the case of (right- or left-) reversing, as we can only delete such pairs, but not create them. A complete formal proof requires to check all possible cases: this is easy, and we skip the details. $\qquad\square$

of Proposition 3.15. Let $T'=T*w$. By definition, w is right-$\boldsymbol{R}$-reversible to $N_r(w)D_r(w)^{-1}$, and left-$\boldsymbol{R}$-reversible to $D_\ell(w)^{-1}N_\ell(w)$. By Lemma 3.16, this implies that $T*N_r(w)D_r(w)^{-1}$ and $T*D_\ell(w)^{-1}N_\ell(w)$ are defined. Put $T_\ell=T*D_\ell(w)^{-1}$ and $T_r=T*N_r(w)$. By hypothesis, the terms T,T', T_ℓ, and T_r all have size n. Hence there exists a positive $\boldsymbol{A}$-word u (namely c_{T_ℓ}) mapping the right comb C_n to T_ℓ. By symmetry, there exists a positive $\boldsymbol{A}$-word v mapping T_r to the left comb $\widetilde{C}_n$. Then $uN_\ell(w)D_r(w)v$ and $uD_\ell(w)N_r(w)v$ are $\boldsymbol{R}$-equivalent positive $\boldsymbol{A}$-words, and both map C_n to $\widetilde{C}_n$, see Figure 16. Now $a_\emptyset^{n-2}$ also maps C_n to $\widetilde{C}_n$. Hence, by Proposition 2.4, we must have

$$(3.10)\qquad a_\emptyset^{n-2}\equiv_{\boldsymbol{R}}^+ uN_\ell(w)D_r(w)v\equiv_{\boldsymbol{R}}^+ uD_\ell(w)N_r(w)v.$$

Then the function λ of (3.3) provides an upper bound for the lengths of the words $\boldsymbol{R}$-equivalent to a given word. In the current case, we have $\lambda(a_\emptyset^{n-2})=(n-1)(n-2)/2$, and (3.9) follows. $\qquad\square$
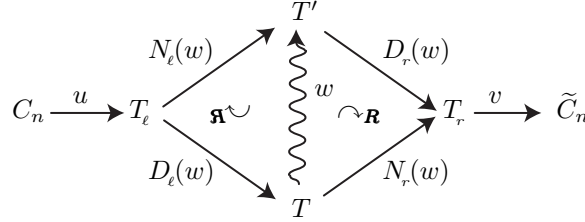
FIGURE 16. Bounding the lengths of the left- and right-numerators and denominators of a signed $\boldsymbol{A}$-word w in terms of the size of a term T such that $T * w$ is defined.

The upper bound of (3.9) is close to sharp: for $w = (a_{1^{p-1}} \dots a_1 a_\emptyset)^{-1} a_{1^p}^q$, the word $D_r(w)$ is $\hat{a}_{1^{p-1},q} \dots \hat{a}_{1,q} \hat{a}_{\emptyset,q}$, which has length pq in the alphabet $\boldsymbol{A}$, so the sum of the lengths of $N_\ell(w)$ and $D_r(w)$ is $p + pq$, while the minimal size of a term T such that $T * w$ is defined is $p + q + 2$.

To conclude with subword reversing, we mention one more result that involves both left- and right-reversing. The example of ε and $a_\emptyset a_\emptyset^{-1}$ shows that $\boldsymbol{R}$-equivalent words need not have $\boldsymbol{R}$-equivalent numerators and denominators: the right-numerator of ε is ε, whereas the right-numerator of $a_\emptyset a_\emptyset^{-1}$ is $a_\emptyset$. This cannot happen when left- and right-numerators are mixed in a double reversing.

Proposition 3.17. *For w a signed $\boldsymbol{A}$-word, define $N_{\ell r}(w)$ to be $N_r(D_\ell(w)^{-1} N_\ell(w))$ and $D_{\ell r}(w)$ to be $D_r(D_\ell(w)^{-1} N_\ell(w))$. Then $w \equiv_{\boldsymbol{R}} w'$ implies $N_{\ell r}(w') \equiv_{\boldsymbol{R}}^+ N_{\ell r}(w)$ and $D_{\ell r}(w') \equiv_{\boldsymbol{R}}^+ D_{\ell r}(w)$.*

We first observe that $N_{\ell r}(w) D_{\ell r}(w)^{-1}$ is a minimal fractionary expression of $\overline{w}$:

Lemma 3.18. *If w, w' are $\boldsymbol{R}$-equivalent signed $\boldsymbol{A}$-words, there exist a positive $\boldsymbol{A}$-word u satisfying*

$$(3.11) \qquad N_r(w') \equiv_{\boldsymbol{R}}^+ N_{\ell r}(w)\, u \quad and \quad D_r(w') \equiv_{\boldsymbol{R}}^+ D_{\ell r}(w)\, u.$$

Proof. By construction, the word w is $\boldsymbol{R}$-equivalent to $D_\ell(w)^{-1} N_\ell(w)$, and the latter word is right-$\boldsymbol{R}$-reversible to $N_{\ell r}(w) D_{\ell r}(w)^{-1}$. Hence we have

$$(3.12) \qquad D_\ell(w) N_{\ell r}(w) \equiv_{\boldsymbol{R}}^+ N_\ell(w) D_{\ell r}(w),$$

and, moreover, as mentioned in the proof of Proposition 3.11, the element of F_{sym}^+ represented by $D_\ell(w) N_{\ell r}(w)$ and $N_\ell(w) D_{\ell r}(w)$ is the right-lcm of $\overline{N_\ell(w)}$ and $\overline{D_\ell(w)}$.

On the other hand, $N_r(w') D_r(w')^{-1}$ is $\boldsymbol{R}$-equivalent to w', hence to w, and therefore to $D_\ell(w)^{-1} N_\ell(w)^{-1}$. We deduce $D_\ell(w) N_r(w') \equiv_{\boldsymbol{R}} N_\ell(w) D_r(w')$, whence

$$(3.13) \qquad D_\ell(w) N_r(w') \equiv_{\boldsymbol{R}}^+ N_\ell(w) D_r(w')$$

since F_{sym}^+ embeds in F. As $\overline{D_\ell(w) N_{\ell r}(w)}$ is the right-lcm of $\overline{N_\ell(w)}$ and $\overline{D_\ell(w)}$, comparing (3.12) and (3.13) implies the existence of u satisfying (3.11). $\square$

of Proposition 3.17. By Lemma 3.18, there exist positive $\boldsymbol{A}$-words u and u' satisfying $N_{\ell r}(w') \equiv_{\boldsymbol{R}}^+ N_{\ell r}(w) u$ and $N_{\ell r}(w) \equiv_{\boldsymbol{R}}^+ N_{\ell r}(w') u'$, whence $N_{\ell r}(w) \equiv_{\boldsymbol{R}}^+ N_{\ell r}(w) u u'$. As F_{sym}^+ is left-cancellative, we deduce $\varepsilon \equiv_{\boldsymbol{R}}^+ u u'$. The only possibility is then that u and u' are empty. $\square$

We now return to Tamari lattices, and show how to use right-reversing to compute lowest upper bounds in the Tamari poset $\mathcal{T}_n$ appealing to the words c_T of (3.8). Of course, left-reversing can be used symmetrically to compute greatest lower bounds.

Proposition 3.19. *Assume that T, T' are size n trees. Then the least upper bound T'' of T and T' in the Tamari lattice $\mathcal{T}_n$ is determined by*

$$T'' = T * N_r(c_T^{-1} c_{T'}) = T' * D_r(c_T^{-1} c_{T'}).$$

Proof. As mentioned in the proof of Proposition 3.11, the words $c_T \, N_r(c_T^{-1} c_{T'})$ and $c_{T'} \, D_r(c_T^{-1} c_{T'})$ both represent the right-lcm of $\overline{c_T}$ and $\overline{c_{T'}}$ in F_{sym}^+. Hence, owing to Proposition 3.9, the image of $\overline{c_T \, N_r(c_T^{-1} c_{T'})}$ under I_n, which, by definition, is $C_n * c_T \, N_r(c_T^{-1} c_{T'})$, that is, $T * N_r(c_T^{-1} c_{T'})$, is the least upper bound in $\mathcal{T}_n$ of $I_n(\overline{c_T})$, that is, of $C_n * c_T$, which is T, and $I_n(\overline{c_{T'}})$, that is, of $C_n * c_{T'}$, which is T'. $\qquad\square$

Example 3.20. Let $T = \bullet(((\bullet(\bullet\bullet))\bullet)\bullet)$ and $T' = (\bullet(\bullet\bullet))(\bullet(\bullet\bullet))$. Using the method explained in Lemma 4.2 below, one obtains $c_T = a_{11} a_1^2$ and $c_{T'} = a_1 a_\emptyset$. Right-reversing $a_1^{-2} a_{11}^{-1} a_1 a_\emptyset$ leads to $a_{100} a_\emptyset a_0 a_{00} a_\emptyset^{-2}$ (see Figure 17), and we deduce that the least upper bound of T and T' in the Tamari poset is the tree $T * a_{100} a_\emptyset a_0 a_{00}$, namely $(((\bullet(\bullet\bullet))\bullet)\bullet)\bullet$ (which is also $T' * a_\emptyset^2$).
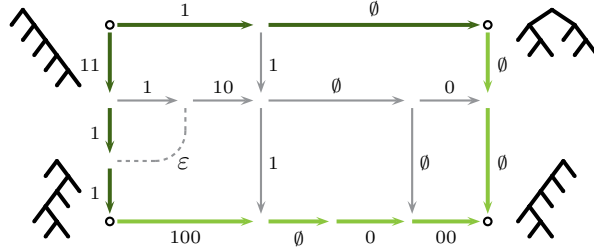

FIGURE 17. Computing the right-lcm of c_T and c'_T by right-reversing determines the least upper bound of T and T' in the Tamari lattice, here for $T = \bullet(((\bullet(\bullet\bullet))\bullet)\bullet)$ and $T' = (\bullet(\bullet\bullet))(\bullet(\bullet\bullet))$.

4. The Polish normal form on F

We now develop another approach for determining least common multiples in the monoid F_{sym}^+, whence, equivalently, least upper bounds in the Tamari lattices, namely using what is known as the Polish algorithm. Initially introduced in the case of the self-distributivity law [8, Chapter IX], the latter is easily adapted to our current context, where it provides a unique normal form for the elements of F_{sym}^+ and a method for determining the upper bound of two trees in the Tamari lattice. The main technical tool here is the covering relation of [12], a variant of the weight sequences of [23]—also see [22] and [1].

The section is organized as follows. In Subsection 4.1 we recall the standard Polish encoding of trees and its connection with the Tamari ordering. In Subsection 4.2 we describe an algorithm that, starting with the Polish encoding of two trees, determines a common upper bound of the latter in the Tamari lattice together with a distinguished way of performing the rotations. Then, in Subsection 4.3, we use the covering relation to control the previous algorithm and, in particular, prove that it always determines the least upper bound of the initial trees. Finally, in Subsection 4.4, we deduce a unique normal form for the elements of F that enjoys a sort of weak rationality property.

4.1. **The Polish encoding of trees.** As is well known, trees or, equivalently, parenthesized expressions can be encoded without parentheses using the Polish notation. Here we consider the right version, and use $\circ$ as the operation symbol.

Definition 11. For T a tree, the *(right)-Polish encoding* of T is the word $\langle T \rangle$ recursively defined by $\langle T \rangle = T$ if T has size 0, and $\langle T \rangle = \langle T_0 \rangle \langle T_1 \rangle \circ$ for $T = T_0{}^\wedge T_1$.

For T of size n, the Polish encoding $\langle T \rangle$ is a word of length $2n + 1$, which we consider as a map of $\{1, ..., 2n + 1\}$ into $\{\bullet, \circ\}$ (when we restrict to unlabeled trees): thus $\langle T \rangle(k)$ refers to the kth letter of the word $\langle T \rangle$. There exists a natural one-to-one *origin* function from the positions of the letters of $\langle T \rangle$ to the addresses of the nodes of T, recursively defined for $T = T_0{}^\wedge T_1$ with T_0 of size n_0 by the rule that the origin of k in T is 0α where α is the origin of k in T_0 for $k \leqslant 2n_0 + 1$, it is 1α where α is the origin of $k - 2n_0 - 1$ in T_1 for $2n_0 + 1 < k \leqslant 2n$, and it is $\emptyset$ for $k = 2n + 1$. For instance, the Polish encoding of the tree $\bullet((\bullet\bullet)\bullet)$ of Figure 2 is $\bullet\bullet\bullet\circ\bullet\circ\circ$, and the corresponding origins are $\bullet_{(0)}\bullet_{(100)}\bullet_{(101)}\circ_{(10)}\bullet_{(11)}\circ_{(1)}\circ_{(\emptyset)}$.

For our current purpose, it is important to note the following connection between the Polish encoding and the Tamari order.

Lemma 4.1. *Let $<^{\mathrm{Lex}}$ denote the lexicographical extension of the ordering $\bullet < \circ$ to $\{\bullet, \circ\}$-words. Then, for all trees T, T', the relation $T \leqslant_{\mathcal{T}} T'$ implies $\langle T \rangle \leqslant^{\mathrm{Lex}} \langle T' \rangle$.*

Proof. When translated to the right Polish notation, applying a left-rotation in a tree amounts to replacing some subword of the form $\langle T_0 \rangle \langle T_1 \rangle \langle T_2 \rangle \circ\circ$ with the corresponding word $\langle T_0 \rangle \langle T_1 \rangle \circ \langle T_2 \rangle \circ$. The latter word is $<^{\mathrm{Lex}}$-larger than the former, as the beginning of the word is preserved, until the first letter $\bullet$ associated with $\langle T_2 \rangle$, which is replaced with $\circ$. $\square$

When the initial letter $\bullet$ is erased, the words that are the Polish encoding of a trees identify with Dyck words, defined as those words in the alphabet $\{\bullet, \circ\}$ such that no initial segment has more $\circ$'s than $\bullet$'s, see for instance [25]. Using the standard correspondence between such words and random walks in $\mathbb{N}^2$, we obtain a simple receipe for determining the elements c_T and c'_T of (3.8) from $\langle T \rangle$.

Lemma 4.2. *(See Figure 18.) Assume that T is a size n tree. For k in $\{1, ..., 2n + 1\}$ recursively define $\nu_T(k)$ by $\nu_T(1) = -1$ and, for $k \geqslant 2$,*

$$(4.1) \qquad \nu_T(k) = \begin{cases} \nu_T(k-1) + 1 & \text{for } \langle T \rangle(k-1) = \langle T \rangle(k) = \bullet, \\ \nu_T(k-1) - 1 & \text{for } \langle T \rangle(p-1) = \langle T \rangle(p) = \circ, \\ \nu_T(k-1) & \text{otherwise.} \end{cases}$$

Then c'_T is obtained from $\langle T \rangle$ by replacing each letter $\langle T \rangle(k)$ with ε if it is $\bullet$ and with a_{1^i} with $i = \nu_T(k)$ if is it $\circ$; the word c_T is obtained similarly after erasing the last block of $\circ$.

We skip the verification, a comparison of the recursive definitions of $\langle T \rangle$ and c'_T.

4.2. **The Polish algorithm.** Assume that T, T' are trees of size n and we look for a (minimal) tree T'' that is an upper bound of T and T' in the Tamari order. If T and T' do not coincide, then one of the words $\langle T \rangle$, $\langle T' \rangle$ is lexicographically smaller than the other, say for instance $\langle T \rangle$. This means means that there exists k such that $\langle T \rangle(k)$ is $\bullet$, whereas $\langle T' \rangle(k)$ is $\circ$. In this case, we shall say that T and T' have a *clash at k*. Here is the point.

Lemma 4.3. *Assume that T is a tree and that the kth letter in $\langle T \rangle$ is $\bullet$. Then there exists at most one pair (α, r) such that $T * \hat{a}_{\alpha,p}$ is defined and T and $T * \hat{a}_{\alpha,r}$ have a clash*
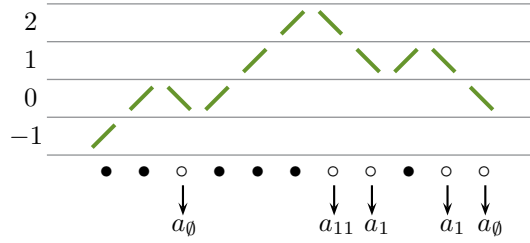
FIGURE 18. Computing c_T and c'_T from the Polish encoding $\langle T \rangle$ of T: write the kth letter of $\langle T \rangle$ at level $\nu_T(k)$; then c'_T is read from the levels of the letters $\circ$. Here, for $(\bullet\bullet)((\bullet(\bullet\bullet))\bullet)$, we read $c'_T = a_\emptyset a_{11} a_1^2 a_\emptyset$, and, discarding the last two symbols $\circ$, $c_T = a_\emptyset a_{11} a_1$.

at k. Moreover, if there exists T'' such that T and T'' have a clash at k, there exists exactly one pair as above.

Proof. As Figure 19 shows, if we have $T' = T * \hat{a}_{\alpha,r}$, then the words $\langle T \rangle$ and $\langle T' \rangle$ coincide up to the first letter coming from the $\alpha 10^{r-1}1$-subtree of T: the latter is $\bullet$ (as is always the first letter of a Polish encoding), whereas, in $\langle T' \rangle$, we have a letter $\circ$ at this position. Thus, the action of $\hat{a}_{\alpha,r}$ on $\langle T \rangle$ is to replace $\bullet$ by $\circ$ at a position whose origin in T has the form $\alpha 10^{r-1}10^i$ for some i.

Consider now the kth letter in $\langle T \rangle$, supposed to be a letter $\bullet$. The origin of k in T is a certain address of leaf in T, say β. By the above argument, a pair (α, r) may result in a clash at k only if we can write $\beta = \alpha 10^{r-1}10^i$ for some $r \geqslant 1$ and $i \geqslant 0$. For every β, this happens for at most one pair (α, r), and this happens if and only if β contains at least two digits 1.

Assume now that T and T'' have a clash at k, and consider the value of $\nu_T(k)$ as defined in (4.1). By construction (and by the standard properties of Dyck words), we have $\nu_{T''}(k) \geqslant 0$ as $\langle T'' \rangle(k)$ is $\circ$. By construction, we have $\nu_T(k) > \nu_{T''}(k)$ since $\langle T \rangle(k)$ is $\bullet$, whence $\nu_T(k) \geqslant 1$. This implies (actually an equivalence) that the address β contains at least two digits 1. Hence there exists a pair (α, r) as above.   □
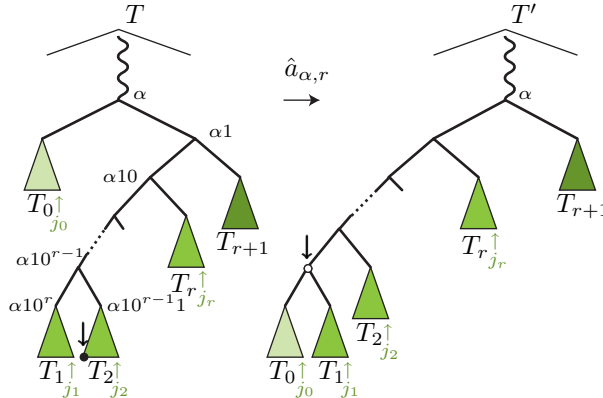

FIGURE 19. Action of $\hat{a}_{\alpha,r}$: the Polish encodings coincide up to the first $\bullet$ corresponding to T_2 in $\langle T \rangle$ (black arrow); the latter is replaced with $\circ$ in $\langle T' \rangle$ because, in T', there is one more right-edge after T_1 than in T, and the clash occurs between the marked letters.

Now the principle of an algorithm should be clear: starting with two trees T, T' such that the Polish encoding $\langle T \rangle$ and $\langle T' \rangle$ coincide up to position $k-1$, we have found a unique way of applying an iterated left-rotation $\hat{a}_{\alpha, r}$ to one of the trees so that the clash is moved further to the right. By iterating the process, we obtain after finitely many steps two trees whose Polish encodings coincide, that is, we obtain a common upper bound for the initial trees T, T'.

Definition 12. Assume that T, T' are trees of equal size.

(i) If $\langle T \rangle <^{\mathrm{Lex}} \langle T' \rangle$ holds, we denote by $s(T, T')$ the unique element $\hat{a}_{\alpha, r}$ such that $T * \hat{a}_{\alpha, r}$ and T' have no clash at the position where T and T' have one.

(ii) We denote by $S(T, T')$ the signed $\widehat{\boldsymbol{A}}$-word recursively defined by the rules

$$(4.2) \qquad S(T, T') = \begin{cases} \varepsilon & \text{for } T = T', \\ s(T, T')\, S(T * s(T, T'), T') & \text{for } \langle T \rangle <^{\mathrm{Lex}} \langle T' \rangle, \\ S(T, T' * s(T', T))\, s(T', T)^{-1} & \text{for } \langle T \rangle >^{\mathrm{Lex}} \langle T' \rangle. \end{cases}$$

Example 4.4. Let us consider the trees of Example 3.20 again, namely $T_0 = \bullet(((\bullet(\bullet\bullet))\bullet)\bullet)$ and $T_0' = (\bullet(\bullet\bullet))(\bullet(\bullet\bullet))$. We find

$\langle T_0 \rangle = \bullet\bullet\bullet\underline{\bullet}\circ\circ\bullet\circ\bullet\circ\circ$,

$\langle T_0' \rangle = \bullet\bullet\bullet\circ\circ\bullet\bullet\bullet\circ\circ\circ$.

Thus we have $\langle T_0 \rangle <^{\mathrm{Lex}} \langle T_0' \rangle$, with a clash at 4 (underlined). The origin of 4 in T_0 is 10011, whence $s(T_0, T_0') = a_{100}$, and $S(T_0, T_0') = a_{100}\, S(T_1, T_1')$ with $T_1 = T_0 * a_{100}$ and $T_1' = T_0'$, corresponding to

$\langle T_1 \rangle = \bullet\bullet\bullet\circ\underline{\bullet}\circ\bullet\circ\bullet\circ\circ$,

$\langle T_1' \rangle = \bullet\bullet\bullet\circ\circ\bullet\bullet\bullet\circ\circ\circ$.

We have now $\langle T_1 \rangle <^{\mathrm{Lex}} \langle T_1' \rangle$, with a clash at 5. The origin of 5 in T_1 is 1001, whence $s(T_1, T_1') = \hat{a}_{\emptyset, 3}$, and $S(T_1, T_1') = \hat{a}_{\emptyset, 3}\, S(T_2, T_2')$ with $T_2 = T_1 * \hat{a}_{\emptyset, 3}$ and $T_2' = T_1'$, hence

$\langle T_2 \rangle = \bullet\bullet\bullet\circ\circ\bullet\circ\bullet\circ\bullet\circ$,

$\langle T_2' \rangle = \bullet\bullet\bullet\circ\circ\bullet\underline{\bullet}\bullet\circ\circ\circ$.

This time, we have $\langle T_2' \rangle <^{\mathrm{Lex}} \langle T_2 \rangle$, with a clash at 7. The origin of 7 in T_2' is 110, so $s(T_2', T_2)$ is $a_\emptyset$, and $S(T_2, T_2') = S(T_3, T_3')\, a_\emptyset^{-1}$ with $T_3 = T_2$ and $T_3' = T_2' * a_\emptyset$, that is,

$\langle T_3 \rangle = \bullet\bullet\bullet\circ\circ\bullet\circ\bullet\circ\bullet\circ$,

$\langle T_3' \rangle = \bullet\bullet\bullet\circ\circ\bullet\circ\bullet\underline{\bullet}\circ\circ$.

We find now $\langle T_3' \rangle <^{\mathrm{Lex}} \langle T_3 \rangle$, with a clash at 9. The origin of 9 in T_3' is 11, whence $s(T_3', T_3) = a_\emptyset$, and $S(T_3, T_3') = S(T_4, T_4')\, a_\emptyset^{-1}$ with $T_4 = T_3$ and $T_4' = T_3' * a_\emptyset$, that is,

$\langle T_4 \rangle = \bullet\bullet\bullet\circ\circ\bullet\circ\bullet\circ\bullet\circ$,

$\langle T_4' \rangle = \bullet\bullet\bullet\circ\circ\bullet\circ\bullet\circ\bullet\circ$.

We have $T_4 = T_4'$, so the algorithm halts. The tree T_4 is a common upper bound of T_0 and T_0', and the word $S(T_0, T_0')$ is $a_{100}\hat{a}_{\emptyset, 3}a_\emptyset^{-2}$.

Thus, for all equal size trees T, T', we obtained a distinguished signed $\widehat{\boldsymbol{A}}$-word $S(T, T')$, and, by construction, the relation $T' = T * S(T, T')$ is satisfied.

Remark 4.5. As mentioned in the beginning of the section, an entirely similar algorithm can be defined with the self-distributivity law $x(yz) = (xy)(xz)$ replacing the associativity law $x(yz) = (xy)z$. Then tree rotations are replaced with distributions, which consist in replacing subtrees $T_0 {}^\wedge (T_1 {}^\wedge T_2)$ with $(T_0 {}^\wedge T_1) {}^\wedge (T_0 {}^\wedge T_2)$. In this case, the size of the trees is changed by the transformations, and termination becomes problematic. Actually, in spite of experimental evidence [15] and positive partial results [8], the question, which seems to be extremely difficult, remains open.

4.3. **The covering relation.** For the moment, we have no connection between the common upper bound of two trees provided by the Polish algorithm of Subsection 4.2 and their least upper bound in the Tamari lattice. In particular, if T, T' are trees satisfying $T \leqslant_{\mathcal{T}} T'$, it is not a priori clear that the Polish algorithm terminates with the pair (T', T'), that is, the clashes always occur on the first of the two current trees. We shall see now that this is actually true. The main tool will be the covering relation, a binary relation that provides a description of the shape of a tree in terms of the addresses of its leaves. We recall that, if T is a size n tree, $T^{\#}$ denotes the labeled tree obtained by attributing to the leaves of T labels 0 to n from left to right. So, for instance, for $T = \bullet((\bullet\bullet)\bullet)$, we have $T^{\#} = \bullet_0((\bullet_1\bullet_2)\bullet_3)$, and $\langle T^{\#} \rangle = \bullet_0\bullet_1\bullet_2\circ\bullet_3\circ\circ$.

Definition 13. (See Figure 20.) Assume that T is a size n tree. For $0 \leqslant i \leqslant n$, we define $\mathrm{add}_T(i)$ to be the origin of $\bullet_i$ in $\langle T^{\#} \rangle$. Then, for $j > i$, we say that j *covers* i in T, written $j \rhd_T i$, if there exists an address γ such that $\mathrm{add}_T(j)$ has the form $\gamma 1^p$ for some positive p and $\mathrm{add}_T(i)$ begins with $\gamma 0$. We write $j \unrhd_T i$ for "$j \rhd_T i$ or $j = i$".
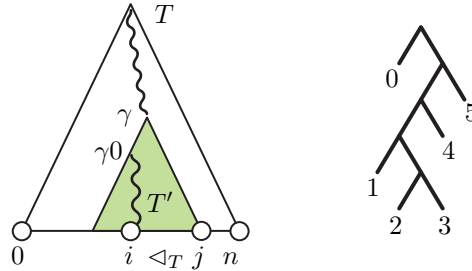

FIGURE 20. Covering relation of T: the leaves are numbered 0 to n, and j covers i in T if there exists a subtree T' such that j is the last (rightmost) label in T', whereas i is a non-final label in T'. For instance, in the right hand tree, 4 covers 1, 2, 3, but does not cover 0, and 3 covers 1 and 2, whereas 2 covers nobody.

It is easily seen [12] that, for every j occurring in a tree T, the set of all i's covered by j is either empty or is an interval ending in $j - 1$: if $j \rhd_T i$ and $j > i' \geqslant i$ hold, then so does $j \rhd_T i'$. Also, the relation $\rhd_T$ is transitive, and it determines T. We shall need the more precise result that every initial fragment of the covering relation determines the corresponding initial fragment of the Polish encoding of T.

Lemma 4.6. *Assume that T is a size n tree. Then, for $1 \leqslant j \leqslant n + 1$, the number of symbols $\circ$ following the jth letter $\bullet$ in $\langle T \rangle$ is the number of i's satisfying $j \rhd_T i$ and $k \not\rhd_T i$ for $j > k > i$.*

Proof. Write $j \rhd_T^{\#} i$ if we have $j \rhd_T i$ and $k \not\rhd_T i$ for $j > k > i$. Then, by definition, $j \rhd_T^{\#} i$ holds if and only we have $\mathrm{add}_T(j) = \alpha 1^q$ and $\mathrm{add}_T(i) = \alpha 0 1^p$ for some α and some $p, q \geqslant 0$. Indeed, if we have $\mathrm{add}_T(i) = \alpha 0 \beta$ with β containing at least one 0, say $\beta = 1^p 0 \gamma$, then we have $k \rhd_T i$ for k satisfying $\mathrm{add}_T(k) = \alpha 0 1^p 0 1^r$.

On the other hand, an induction shows that the jth letter $\bullet$ in $\langle T \rangle$ is followed by r letters $\circ$ if and only if the address $\mathrm{add}_T(j)$ has the form $\gamma 1^r$ for some γ that does not finish with 1, that is, is empty or finishes with 0.

Now, assume that $\mathrm{add}_T(j)$ is $\gamma 1^r$. For $0 \leqslant m < r$, let i_m be the (unique) position whose address in T has the form $\gamma 1^m 0 1^q$ for some q. By the above characterization, we have $j \rhd_T^{\#} i_m$. So the number of i's satisfying $j \rhd_T^{\#} i$ is at least r.

Conversely, assume that there are r different positions $i_0 < ... < i_{r-1}$ satisfying $j \rhd_T^\# i_m$. By the above characterization, there exist $\alpha_0, ..., \alpha_{r-1}$ satisfying $\mathrm{add}_T(i_m) = \alpha_m 01^*$ and $\mathrm{add}_T(j) = \alpha 1^*$ As the numbers i_m are pairwise distinct, so are the addresses α_m and, therefore, we have $\mathrm{add}_T(j) = \alpha_0 1^{r'}$ with $r' \geqslant r$. $\qquad\square$

It directly follows from Lemma 4.6 that the covering relation of a tree T determines the Polish encoding of T, hence T itself. Actually, the lemma shows more.

Lemma 4.7. *Assume that T, T' are equal size trees, and (as a set of pairs) $\rhd_T$ is properly included in $\rhd_{T'}$. Then $\langle T \rangle <^{\mathrm{Lex}} \langle T' \rangle$ holds.*

Proof. Let j be minimal such that there exists i satisfying $j \rhd_{T'} i$ but not $j \rhd_T i$. For $k < j$, the restriction of the covering relations $\rhd_T$ and $\rhd_{T'}$ to the interval $[1, k]$ coincide and, therefore, by Lemma 4.6, the numbers of symbols $\circ$ following $\bullet_{k-1}$ in $\langle T^\# \rangle$ and $\langle T'^\# \rangle$ are equal. So, up to $\bullet_{j-1}$, the words $\langle T \rangle$ and $\langle T' \rangle$ coincide.

Consider now $\bullet_{j-1}$. We claim that the number r' of $\circ$ following $\bullet_{j-1}$ in $\langle T'^\# \rangle$ is larger than its counterpart r in $\langle T^\# \rangle$, resulting in a clash between $\langle T \rangle$ and $\langle T' \rangle$ and in the inequality $\langle T \rangle <^{\mathrm{Lex}} \langle T' \rangle$. To see that $r' > r$ holds, we use Lemma 4.6 again. Using $\rhd_T^\#$ as in the proof of Lemma 4.6, we note that $j \rhd_T^\# i$ implies $j \rhd_{T'}^\# i$ as the restrictions of $\rhd_T$ and $\rhd_{T'}$ to $[1, j-1]$ coincide. By hypothesis, there are r values of i satisfying $j \rhd_T^\# i$, and these values also satisfy $j \rhd_{T'}^\# i$. Now, by hypothesis, there exists i' satisfying $j \not\rhd_T i'$ and $j \rhd_{T'} i'$. If i' is chosen maximal, we have $j \rhd_{T'}^\# i'$. Hence there are strictly more than r values of i satisfying $j \rhd_{T'}^\# i$, as expected. $\qquad\square$

When a left-rotation transforms a tree T into a tree T', the covering relation of T is included in that of T'. The precise relation is as follows. Hereafter we use $\{0, 1\}^*$ (*resp.* $\{1\}^*$) for the set of all addresses (*resp.* all addresses of the form 1^i).

Lemma 4.8. *Assume $T' = T * \hat{a}_{\alpha, r}$. Then $\rhd_{T'}$ is obtained by adding to $\rhd_T$ the pairs (j, i) that satisfy*

$$(4.3) \qquad \exists m \in \{1, ..., r\} \ (\mathrm{add}_T(j) \in \alpha 10^{r+1-m}\{1\}^*) \quad and \quad \mathrm{add}_T(i) \in \alpha 0\{0, 1\}^*.$$

Proof. Consider Figure 19 again. Let $j_1, ..., j_r$ denote the last variable in the subtrees $T_1, ..., T_r$. A direct inspection shows that every covering pair in T is still a covering pair in T', and that the new covering pairs are the pairs (j_m, i) with $1 \leqslant m \leqslant r$ and i occurring in T_0: the action of $\hat{a}_{\alpha, r}$ is to let $j_1, ..., j_m$ cover the variables of T_0. Converted into addresses, this gives (4.3). $\qquad\square$

Lemma 4.8 is important for the Polish algorithm as it bounds possible coverings.

Lemma 4.9. *Assume that T, T' are equal size trees satisfying $\langle T \rangle <^{\mathrm{Lex}} \langle T' \rangle$. Then the covering relation of $T * s(T, T')$ is included in the transitive closure of $\rhd_T$ and $\rhd_{T'}$.*

Proof. Assume $s(T, T') = \hat{a}_{\alpha, r}$ and let $T_1 = T * \hat{a}_{\alpha, r}$. We use the notation of Figure 19 once more, calling j_m the rightmost variable occurring in T_m for $0 \leqslant m \leqslant r$. Let I denote the set of all i's occurring in the subtree T_0. By Lemma 4.8, the pairs that belong to $\rhd_{T_1}$ and not to $\rhd_T$ are the pairs $(j_1, i), ..., (j_r, i)$ with i in I. The hypothesis that $\hat{a}_{\alpha, r}$ is $s(T, T')$ implies that the number of $\circ$ following $\bullet_{j-1}$ in $\langle T'^\# \rangle$ is larger than its counterpart in T, so j_1 must cover strictly more positions in T' than in T. So, necessarily, $j_1 \rhd_{T'} j_0$ holds. On the other hand, $j_0 \unrhd_T i$ holds for every i in I, and $j_m \unrhd_T j_1$ holds for $1 \leqslant m \leqslant r$. It follows that, for all m in $\{1, ..., r\}$ and i in I, the pair (j_m, i) belongs to the the transitive closure of $\rhd_T$ and $\rhd_{T'}$. $\qquad\square$

Lemma 4.10. *Assume that T, T' are equal size trees. Then the Polish algorithm running on (T, T') terminates with a pair (T_∞, T_∞) such that $\rhd_{T_\infty}$ is the transitive closure of $\rhd_T$ and $\rhd_{T'}$.*

Proof. Let (T_t, T_t') denote the pair of trees obtained after t steps of the Polish algorithm running on (T, T'), and N be the total number of steps. By Lemma 4.8, the relations $\rhd_{T_t}$ make a non-decreasing sequence with respect to inclusion, and so do the relations $\rhd_{T_t'}$. So, in particular, the transitive closure of $\rhd_T$ and $\rhd_{T'}$ is included in the transitive closure of $\rhd_{T_N}$ and $\rhd_{T_N'}$. Now, by hypothesis, the latter is $\rhd_{T_\infty}$.

On the other hand, Lemma 4.9 shows that, for every t, the relation $\rhd_{T_{t+1}}$ is included in the transitive closure of $\rhd_{T_t}$ and $\rhd_{T_t'}$, and so is $\rhd_{T_{t+1}'}$. Hence the transitive closure of $\rhd_{T_{t+1}}$ and $\rhd_{T_{t+1}'}$ is the transitive closure of $\rhd_{T_t}$ and $\rhd_{T_t'}$. Hence $\rhd_{T_\infty}$, which is the transitive closure of $\rhd_{T_N}$ and $\rhd_{T_N'}$, is the transitive closure of $\rhd_{T_0}$ and $\rhd_{T_0'}$. □

We are ready to put pieces together and state the main results of this section.

Proposition 4.11. *For T, T', T'' equal size trees, the following are equivalent:*
  (i) *The tree T'' is the least upper bound of T and T' in the Tamari lattice;*
  (ii) *The Polish algorithm running on (T, T') returns (T'', T'');*
  (iii) *The covering relation of T'' is the transitive closure of those of T and T'.*

Proof. Let $T_\vee$ be the least upper bound of T and T' in the Tamari lattice, and T_∞ be the tree such that the Polish algorithm running on (T, T') returns (T_∞, T_∞). By Lemma 4.8, $\rhd_T$ and $\rhd_{T'}$ are included in $\rhd_{T_\vee}$. Hence the transitive closure of $\rhd_T$ and $\rhd_{T'}$, which by Lemma 4.10 is $\rhd_{T_\infty}$, is included in $\rhd_{T_\vee}$.

On the other hand, by definition, we have $T \leqslant_{\mathcal{T}} T_\infty$ and $T' \leqslant_{\mathcal{T}} T_\infty$, whence $T_\vee \leqslant_{\mathcal{T}} T_\infty$. This implies that $\rhd_{T_\vee}$ is included in $\rhd_{T_\infty}$. Hence $\rhd_{T_\vee}$ and $\rhd_{T_\infty}$ coincide, and, therefore, $T_\vee = T_\infty$ holds. So (i) and (ii) are equivalent.

Next, as said above, (ii) implies (iii) by Lemma 4.8. Conversely, if T'' is such that $\rhd_{T''}$ is the transitive closure of $\rhd_T$ and $\rhd_{T'}$, then $\rhd_{T''}$ coincides with $\rhd_{T_\infty}$ and, therefore, we must have $T'' = T_\infty$. So, (ii) and (iii) are equivalent. □

Corollary 4.12. *For T, T' equal size trees, the following are equivalent:*
  (i) *We have $T \leqslant_{\mathcal{T}} T'$ in the Tamari order;*
  (ii) *There exists f in F_{sym}^+ such that $T' = T * f$ holds;*
  (iii) *$S(T, T')$ is a positive **A**-word, that is, the Polish algorithm running on (T, T') finishes with (T', T').*
  (iv) *The relation $\rhd_T$ is included in $\rhd_{T'}$.*

Proof. The equivalence of (i) and (ii) has been established in Lemma 3.1.

Next, it is obvious that (iii) implies (ii) as every element $\hat{a}_{\alpha,r}$ belongs to F_{sym}^+. Conversely, if $T \leqslant_{\mathcal{T}} T'$ holds, then the least upper bound of T and T' is T'. Hence, by Proposition 4.11, the Polish algorithm running on (T, T') finishes with (T', T'). This means that the word $S(T, T')$ contains positive letters $\hat{a}_{\alpha,r}$ only. So (ii) implies (iii).

Finally, as observed above, (iii) is equivalent to saying that the Polish algorithm running on (T, T') finishes with (T', T'), whereas (iv) is equivalent to saying that $\rhd_{T'}$ is the transitive closure of $\rhd_T$ and $\rhd_{T'}$. By Proposition 4.11, the latter properties are equivalent and, therefore, (iii) and (iv) are equivalent. □

It should be noted that the equivalence of (i) and (iv) in Corollary 4.12 already appears as [23, Theorem 2.1].

4.4. **The Polish normal form.** One of the interests of Proposition 4.11 and Corollary 4.12 is that they provide unique distinguished decompositions for every element of F and of F_{sym}^+ in terms of the generators $\hat{a}_{\alpha,r}$. Indeed, we obtained for every pair of equal size trees (T, T') a certain signed $\widehat{A}$-word $S(T, T')$ such that $T * S(T, T')$ is defined and equal to T'. This word $S(T, T')$ does not depend on T.

Lemma 4.13. *Assume that f belongs to F and $T * f$ is defined. Then the signed $\widehat{A}$-word $S(T, T * f)$ is an expression of f, and it does not depend on T.*

Proof. First, we have $T * f = T * S(T, T')$, so, by Proposition 2.4, the word $S(T, T')$ is an expression of f. Next, assume that σ is a substitution, and let us compare the Polish algorithm running on a pair (T, T') and on the pair (T^σ, T'^σ). The word $\langle T^\sigma \rangle$ is obtained from the word $\langle T \rangle$ by replacing every variable $\bullet_i$ with the corresponding word $\langle \sigma(i) \rangle$. As the variables occur in the same order in the words $\langle T \rangle$ and $\langle T' \rangle$, substituting $\bullet_i$ with $\langle \sigma(i) \rangle$ introduces no new clash. Therefore, if (T_t, T_t') are the trees at the tth step of the algorithm running on (T, T'), then $(T_t^\sigma, T_t'^\sigma)$ are the trees at the tth step of the algorithm running on (T^σ, T'^σ), implying $S(T, T') = S(T^\sigma, T'^\sigma)$.

By definition, for every f in F, there exists a unique pair of trees (f_-, f_+) such that every pair $(T, T * f)$ can be expressed as $((f_-^\#)^\sigma, (f_+^\#)^\sigma)$. The above result then shows that $S(T, T * f)$ coincides with $S(f_-, f_+)$, which only depends on f. □

Definition 14. For f in F, the *Polish normal form* of f is the signed $\widehat{A}$-word $S(f_-, f_+)$.

Example 4.14. Let $f = a_\emptyset a_1 = a_{11} a_\emptyset \; (= x_1 x_2 = x_3 x_1)$. Then we have $f_- = \bullet(\bullet(\bullet(\bullet\bullet)))$ and $f_+ = (\bullet\bullet)((\bullet\bullet)\bullet)$. Running the Polish algorithm on these trees returns the (positive) $\widehat{A}$-word $a_\emptyset a_1$: so the latter is the Polish normal form of f. By contrast, $a_{11} a_\emptyset$, which is another $\widehat{A}$-expression of f, is not normal. One verifies similarly that the word $a_\emptyset^2$ is normal, whereas the equivalent words $a_1 \hat{a}_{\emptyset, 2}$ and $a_1 a_\emptyset a_0$ are not.

Corollary 4.12 immediately implies:

Proposition 4.15. *An element of F belongs to the submonoid F_{sym}^+ if and only if its Polish normal form contains no letter $\hat{a}_{\alpha,r}^{-1}$.*

As the family of generators $\widehat{A}$ is infinite, it makes no sense to wonder whether Polish normal words form a rational language or whether the Polish normal form can be connected with an automatic structure. However, let us observe that being Polish normal is a local property that can be characterized in terms of adjacent letters.

Proposition 4.16. *A positive $\widehat{A}$-word $\hat{a}_{\alpha_1, r_1} \dots \hat{a}_{\alpha_\ell, r_\ell}$ is Polish normal if and only if*

$$\alpha_t 0^{r_t} < \alpha_{t+1} 10^{r_{t+1}-1} 1$$

holds for every $t < \ell$, where $<$ denotes the left–right (partial) ordering of addresses.

Proof. Let $w = \hat{a}_{\alpha_1, r_1} \dots \hat{a}_{\alpha_\ell, r_\ell}$ and assume that $T * w$ is defined. For $0 \leqslant t \leqslant \ell$, put $T_t = T * \hat{a}_{\alpha_1, r_1} \dots \hat{a}_{\alpha_t, r_t}$. Then w is normal if, for every $1 \leqslant t \leqslant \ell$, we have $\hat{a}_{\alpha_t, r_t} = s(T_{t-1}, T_\ell)$, that is, $\hat{a}_{\alpha_t, r_t}$ appears at the tth step of the Polish algorithm running on $(T, T * w)$. Now, as shown in Figure 19, the origin in T_t of the letter $\circ$ involved in the clash between T_{t-1} and T_t is $\alpha_t 0^{r_t}$, whereas the origin in T_t of the letter $\bullet$ involved in the clash between T_t and T_{t+1} lies in $\alpha_{t+1} 10^{r_{t+1}-1} 1 \{0\}^*$. The normality condition is then that, in $\langle T_t \rangle$, the former letter lies on the left of the latter. By construction of the Polish encoding, this happens if and only if the first address precedes the second in the "left–right–root" linear ordering of addresses. Due to the form of the second address, this is equivalent to $\alpha_t 0^{r_t} < \alpha_{t+1} 10^{r_{t+1}-1} 1$. □

For instance, the word $a_\emptyset a_\emptyset$ is normal, as we have $\emptyset 0^1 = 0 < \emptyset 10^{1-1}1 = 11$, but $a_1 \hat{a}_{\emptyset,2}$ is not, as we do not have $10^1 = 10 < \emptyset 10^{2-1}1 = 101$.

5. DISTANCE IN TAMARI LATTICES

We conclude this description of the connections between the Tamari lattice and the Thompson group F with a few observations about distances in $\mathcal{T}_n$. The general principle is that it is easy to obtain upper bounds, but difficult to prove lower bounds and many questions remain open in this area. Our main observation here is that the embedding of the monoid F_{sym}^+ into the group F is not an isometry, and not even a quasi-isometry (Definition 16): for every positive constant C, there exist elements of F_{sym}^+ whose length in F is smaller than their length in F_{sym}^+ by a factor at least C. In terms of Tamari lattices, this implies that chains are not geodesic (Corollary 5.13).

The plan of the section is as follows. In Subsection 5.1, we quickly survey the known results about the diameter of Tamari lattices. Then, we show in Subsection 5.2 how to use the syntactic relations of $\boldsymbol{R}$ to obtain (rather weak) distance lower bounds. Finally, in Subsection 5.3, we use the covering relation to establish (stronger) lower bounds.

5.1. **The diameter of $\mathcal{T}_n$.** Surprisingly, the diameter of the Tamari lattice $\mathcal{T}_n$ is not known for every n.

Definition 15. For T, T' in $\mathcal{T}_n$, the *distance* between T and T', denoted by $\mathrm{dist}(T, T')$ is the minimal number of rotations needed to transform T into T'. The *diameter* of $\mathcal{T}_n$ is the maximum of $\mathrm{dist}(T, T')$ for T, T' in $\mathcal{T}_n$.

Theorem 5.1 (Sleator, Tarjan, Thurston [24]). *For $n \geqslant 11$, the diameter of $\mathcal{T}_n$ is at most $2n - 6$; for n large enough, it is exactly $2n - 6$.*

The argument uses the fact that the maximal distance between two size n trees is also the maximal number of flips needed to transform two triangulations of an $(n+2)$-gon one into the other. A lower bound for the latter is obtained by putting the considered triangulations on the two halves of a sphere and bounding the hyperbolic volume of the resulting tiled polyhedron. It is conjectured that the value $2n - 6$ is correct for every $n \geqslant 11$. However, due to its geometric nature, the argument of [24] works only for $n \geqslant n_0$, with no estimation of n_0.

By contrast, combinatorial arguments involving the covering relation of Subsection 4.3 lead to (weaker) results that are valid for every n.

Theorem 5.2. [12] *For $n = 2p^2$, the diameter of $\mathcal{T}_n$ is at least $2n - 2\sqrt{2n} + 1$ and, for every n, it is at least $2n - \sqrt{70n}$.*

Although no theoretical obstruction seems to exist, the covering arguments have not yet been developed enough to lead to an exact value of the diameter. However some candidates for realizing the maximal distance are known.

Conjecture 5.3. [12] *For α an address, let $\langle \alpha \rangle$ denote the tree recursively specified by the rules $\langle \emptyset \rangle = \bullet$, $\langle 0\alpha \rangle = \langle \alpha \rangle^\wedge \bullet$, and $\langle 1\alpha \rangle = \bullet^\wedge \langle \alpha \rangle$. Define*

$$Z_n = \begin{cases} \langle 111(01)^{p-2} \rangle \\ \langle 111(01)^{p-2}0 \rangle \end{cases} \quad Z_n' = \begin{cases} \langle 000(10)^{p-2} \rangle & \text{for } n = 2p+3, \\ \langle 000(10)^{p-2}1 \rangle & \text{for } n = 2p+4, \end{cases}$$

see Figure 21. Then one has $\mathrm{dist}(Z_n, Z_n') = 2n - 6$ for $n \geqslant 9$.

Conjecture 5.3 has been checked up to size 19 (sizes below 9 are special, because the trees are then too small for the generic scheme to start; by the way, the value $2n - 6$ is valid for $n = 5, 6, 7$, but not for $n \leq 4$ and for $n = 8$).
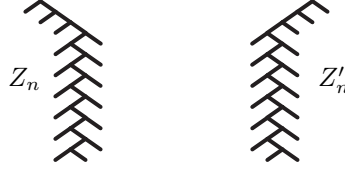
FIGURE 21. The zigzag trees of Conjecture 5.3, here for $n = 15$; the distance is 24, as predicted.

5.2. **Syntactic invariants.** A natural way to investigate distances in the Tamari lattices is to use the action of F on trees and to study the length of the elements of F with respect to the generating family $\boldsymbol{A}$. Indeed, Proposition 2.4 directly implies

Lemma 5.4. *For all trees T, T', we have $\mathrm{dist}(T, T') = \|S(T, T')\|_{\boldsymbol{A}}$, where $\|f\|_{\boldsymbol{A}}$ is the $\boldsymbol{A}$-length of f, that is, the length of the shortest signed $\boldsymbol{A}$-word representing f.*

In order to establish (lower) bounds on $\|f\|_{\boldsymbol{A}}$, a natural approach is to use the syntatic properties of the relations of $\boldsymbol{R}$.

Lemma 5.5. *For w a signed $\boldsymbol{A}$-word, denote by $|w|_1$ the number of letters $a_{1i}^{\pm 1}$ in w.*
  (i) *If u, u' are $\boldsymbol{R}$-equivalent positive $\boldsymbol{A}$-words, then $|u|_1 = |u'|_1$ holds.*
  (ii) *If w, w' are signed $\boldsymbol{A}$-words, then $w \curvearrowright_{\boldsymbol{R}} w'$ implies $|w|_1 \geqslant |w'|_1$ holds.*

Proof. In both cases, it suffices to inspect the relations of $\boldsymbol{R}$. In the case of the pentagon relations, we have $|a_\alpha^2|_1 = |a_{\alpha 0} a_\alpha a_{\alpha 1}|_1$, both being 2 for α in $\{1\}^*$, and 0 otherwise. Similarly, for (ii), we find $|a_\alpha^{-1} a_{\alpha 1}|_1 = |a_\alpha a_{\alpha 0}^{-1} a_\alpha^{-1}|_1$, both being 2 for α in $\{1\}^*$, and 0 otherwise. The inequality comes from $|a_\alpha^{-1} a_\alpha|_1 = 2 > 0 = |\varepsilon|_1$ for α in $\{1\}^*$. $\square$

Note that the counterpart of Lemma 5.5(ii) involving left-reversing is false: $a_\emptyset a_0^{-1}$ is left-$\boldsymbol{R}$-reversible to $a_\emptyset^{-1} a_1 a_\emptyset$, and we have $|a_\emptyset a_0^{-1}|_1 = 1 < 3 = |a_\emptyset^{-1} a_1 a_\emptyset|_1$.

Proposition 5.6. *For every f of F and every $\boldsymbol{A}$-word w representing f, we have*

$$(5.1) \qquad\qquad \|f\|_{\boldsymbol{A}} \geqslant |D_{\ell r}(w)|_1 + |N_{\ell r}(w)|_1.$$

Proof. Put $\ell = |D_{\ell r}(w)|_1 + |N_{\ell r}(w)|_1$. By definition, the word w is right-$\boldsymbol{R}$-reversible to the word $N_r(w) D_r(w)^{-1}$, so, by Lemma 5.5(ii), we have

$$|w|_1 \geqslant |N_r(w) D_r(w)^{-1}|_1 = |N_r(w)|_1 + |D_r(w)|_1.$$

Next, it follows from Proposition 3.17 that there exist a positive $\boldsymbol{A}$-word u satisfying $N_r(w) \equiv_{\boldsymbol{R}}^+ N_{\ell r}(w) u$ and $D_r(w) \equiv_{\boldsymbol{R}}^+ D_{\ell r}(w) u$. Then, by Lemma 5.5(i), we deduce $|N_r(w)|_1 + |D_r(w)|_1 \geqslant \ell$, whence $|w| \geqslant |w|_1 \geqslant \ell$.

Now assume $w' \equiv_{\boldsymbol{R}} w$. By the result above, we have $|w'| \geqslant |N_{\ell r}(w')|_1 + |D_{\ell r}(w')|_1$. By Proposition 3.17, we have $N_{\ell r}(w') \equiv_{\boldsymbol{R}}^+ N_{\ell r}(w)$ and $D_{\ell r}(w') \equiv_{\boldsymbol{R}}^+ D_{\ell r}(w)$, whence, by Lemma 5.5(i), $|N_{\ell r}(w')|_1 + |D_{\ell r}(w')|_1 = \ell$. Thus $|w'| \geqslant \ell$ holds for every word w' that represents $\overline{w}$. By definition, this means that $\|f\|_{\boldsymbol{A}} \geqslant \ell$ is true. $\square$

Of course, a symmetric criterion involving $\{0\}^*$ instead of $\{1\}^*$ may be stated.

Example 5.7. Let $f = a_1 a_{11}^{-1} a_\alpha a_{11}^{-1}$. Left-reversing the word $a_1 a_{11}^{-1} a_\alpha a_{11}^{-1}$ yields the word $a_{111}^{-1} a_{1111}^{-1} a_1 a_\emptyset$, which in turn is right-reversible to $a_1 a_\emptyset a_1^{-1} a_{11}^{-1}$. We conclude that $N_{\ell r}(w)$ is $a_1 a_\emptyset$ and $D_{\ell r}(w)$ is $a_{11} a_1$. Then Proposition 5.6 gives $\|f\|_{\boldsymbol{A}} \geqslant 4$, that is, the word $a_1 a_{11}^{-1} a_\alpha a_{11}^{-1}$ is geodesic.

By construction, the elements c_T involved in proof of Proposition 3.9 are represented by $\boldsymbol{A}$-words all letters of which are of the form a_α with α in $\{1\}^*$, and, therefore, these words are geodesic. However, elements of this type are quite special, and the criterion of Proposition 5.6 is rarely useful. In particular, it follows from the construction that every element of F can be represented by a word of the form $c_T^{-1} c_{T'}$ but, even when the fraction is irreducible, that is, when the elements represented by c_T and $c_{T'}$ admit no common left-divisor in F_{sym}^+, it need not be geodesic, as shows the example of $a_\emptyset^{-p} a_{1^p} a_{1^{p-1}} \dots a_1$, an irreducible fraction of length $2p$ which is $\boldsymbol{R}$-equivalent to the positive–negative word $a_\emptyset a_0^{-p} a_\emptyset^{-1}$ of length $p+2$.

5.3. **The embedding of $F_{\mathbf{sym}}^+$ into F.** More powerful results can be obtained using the covering relation of Subsection 4.3. As an example, we shall now establish that the embedding of the monoid F_{sym}^+ into the group F provided by Proposition 3.2 is not an isometry, that is, there exist elements of F_{sym}^+ whose length as elements of F is smaller than their length as elements of F_{sym}^+. This result is slightly surprising: clearly, fractions need not be geodesic in general, but we might expect that, when an element of F belongs to F_{sym}^+, then its length inside F_{sym}^+ equals its length inside F.

Definition 16. If $(X, d), (X', d')$ are metric spaces, a map $f : X \to X'$ is a *quasi-isometry* if there exist $C \geqslant 1$ and $C' \geqslant 0$ such that $\frac{1}{C} d(f(x, y)) - C' \leqslant d'(f(x, y)) \leqslant C d(x, y) + C'$ holds for all x, y in X.

The result we shall prove is as follows.

Proposition 5.8. *For f in F_{sym}^+, let $\|f\|_{\boldsymbol{A}}^+$ denotes the $\boldsymbol{A}$-length of f in F_{sym}^+, that is, the length of a shortest positive $\boldsymbol{A}$-word representing f. Then the embedding of F_{sym}^+ into F is not a quasi-isometry of $(F_{sym}^+, \|\|_{\boldsymbol{A}}^+)$ into $(F, \|\|_{\boldsymbol{A}})$.*

In order to establish Proposition 5.8, it is enough to exhibit a sequence of elements f_p of F_{sym}^+ satisfying $\|f_p\|_{\boldsymbol{A}} = o(\|f_p\|_{\boldsymbol{A}}^+)$. This is what the next result provides.

Lemma 5.9. *For every $p \geqslant 1$, let u_p be the $\widehat{\boldsymbol{A}}$-word*

$$\hat{a}_{(10)^p,1}\, \hat{a}_{(10)^{p-1},2} \dots \hat{a}_{10,p}\, \hat{a}_{\emptyset,p+1}.$$

Then, for every p, we have $\|\overline{u_p}\|_{\boldsymbol{A}} \leqslant 3p+1$ and $\|\overline{u_p}\|_{\boldsymbol{A}}^+ = (p+1)(p+2)/2$.

Establishing Lemma 5.9 requires to prove two inequalities, namely an upper bound on $\|\overline{u_p}\|_{\boldsymbol{A}}$ and a lower bound on $\|\overline{u_p}\|_{\boldsymbol{A}}^+$. As always, the first task is easier than the second.

Lemma 5.10. *For every $p \geqslant 1$, we have $\|\overline{u_p}\|_{\boldsymbol{A}} \leqslant 3p+1$.*

Proof. For $p \geqslant 1$, let $w_p = a_{(10)^p} a_{(10)^{p-1}1}^{-1} a_{(10)^{p-1}} a_{(10)^{p-2}1}^{-1} a_{(10)^{p-2}} \dots a_{101}^{-1} a_{10} a_1^{-1} a_\emptyset$. Then w_p is a signed $\boldsymbol{A}$-word of length $2p+1$. An easy induction using the formulas of (3.6) shows that w_p is right-reversible to the positive–negative word $u_p \hat{a}_{\emptyset,p}^{-1}$, see Figure 22. As the latter word has $\boldsymbol{A}$-length $3p+1$, the result follows. $\qquad\square$

In order to complete the proof of Lemma 5.9, it remains to prove that $\|\overline{u_p}\|_{\boldsymbol{A}}^+$ is $p(p+3)/2$. As the length of u_p is $p(p+3)/2$, the point is to prove:

Lemma 5.11. *For every $p \geqslant 1$, the word u_p is geodesic in F_{sym}^+.*

Proof. With the notation of Conjecture 5.3, let T_p be the size $2p+2$ tree $\langle (10)^p 1 \rangle$. Then $T_p * u_p$ is defined and equal to $T'_p = \langle 0^{p+1} 1^p \rangle$, see Figure 23.

We look at the covering relations that are satisfied in T'_p but not in T_p. First, $2p+1$ covers 0 in T'_p but not in T_p. We deduce that every $\boldsymbol{A}$-word transforming T_p to T'_p contains
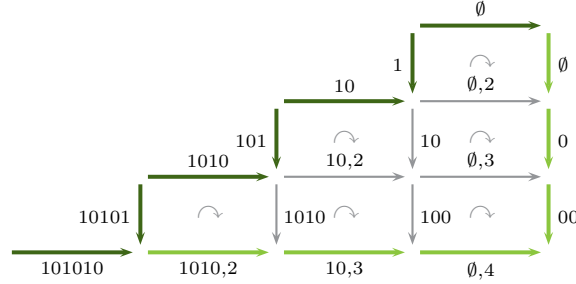
FIGURE 22. Right-reversing the signed A-word w_p, here with $p = 3$. The numerator is the word u_p of Lemma 5.9, whereas the denominator is the length p word $\hat{a}_{\emptyset,p}$ (as usual, the letters "a" have been skipped).

at least one step with critical index $2p + 1$, the critical index of a_α being defined as the unique j such that applying a_α adds at least one relation $j \rhd i$, that is, the unique j such that $\mathrm{add}_T(j)$ lies in $\alpha 10\{1\}^*$.

Now, here is the point: $2p$ covers 0 and 1 in T'_p, but not in T_p. We deduce that every A-word transforming T_p to T'_p contains at least one step with critical index $2p$. But we claim that every such word must actually contain at least *two* such steps, that is, one step cannot be responsible for the two new covering relations. Indeed, $2p$ covers 2 in T_p, but does not cover 1. The only situation when a step adding $2p \rhd 1$ in a tree T can simultaneously add $2p \rhd 0$ is when 1 covers 0 in T. But this cannot happen here, because we are considering positive A-words only, so any possible covering satisfied at an intermediate step must remain in the final tree T'_p. As 1 does not cover 0 in T'_p, it is impossible that 1 covers 0 at any intermediate step. Thus two steps are needed to ensure $2p \rhd 1$ and $2p \rhd 0$.

The sequel is similar. In T'_p, the number $2p - 1$ covers $0, 1$, and 2, whereas, in T_p, it covers only 3. Then at least three steps are needed to ensure $2p - 1 \rhd 2$, $2p - 1 \rhd 1$, and $2p - 1 \rhd 0$. Indeed, a step can cause $2p - 1$ to simultaneously cover 1 and 2 only if 2 covers 1 at the involved step, which cannot happen as 2 does not cover 1 in T'_p; the argument is then the same for 1 and 0 once we know that 2 does not cover 1.

Similarly, 4 steps are needed to force $2p - 2$ to cover 3 to 0, then 5 steps to force $2p - 3$ to cover 4 to 0, etc., and $p + 1$ steps to force $p + 1$ to cover p to 0. We conclude that $1 + 2 + \cdots + (p + 1)$, that is, $(p + 1)(p + 2)/2$, steps at least are needed to go from T_p to T'_p. Hence u_p is geodesic among positive A-words. $\square$
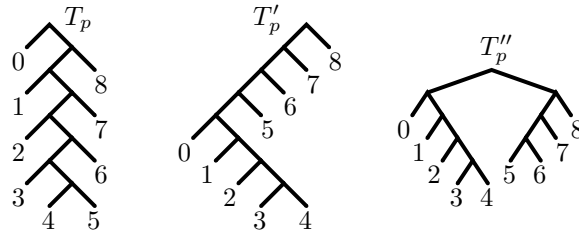

FIGURE 23. The trees of Lemma 5.11 with the leaf labelling involved in the covering relation, here in the case $p = 3$. The tree T''_p is the intermediate tree $T_p * w_p$ obtained after $2p + 1$ steps: applying $\hat{a}_{\emptyset,p}$ to T''_p leads to T'_p in a total of $3p + 1$ rotations.

Thus the proof of Proposition 5.8 is complete.

Finally, we translate the above arguments into the language of Tamari lattices.

Definition 17. For T, T' in $\mathcal{T}_n$ satisfying $T \leqslant_\mathcal{T} T'$, the *positive distance* $\mathrm{dist}^+(T, T')$ from T to T' is the minimal number of left-rotations needed to transform T into T'.

Proposition 5.12. *For every even n, there exist T, T' in $\mathcal{T}_n$ satisfying*

$$(5.2) \qquad\qquad \mathrm{dist}(T, T') \leqslant \frac{12}{n} \, \mathrm{dist}^+(T, T').$$

Proof. Write $n = 2p+2$ and let T_p, T'_p be the size n trees of the proof of Lemma 5.11. Then we have $\mathrm{dist}(T_p, T'_p) \leqslant 3p+1$ and $\mathrm{dist}^+(T_p, T'_p) = (p+1)(p+2)/2$, whence $\mathrm{dist}(T_p, T'_p) \leqslant \frac{12n-16}{n(n+2)} \mathrm{dist}^+(T_p, T'_p)$ in term of n, and, a fortiori, (5.2). $\qquad\square$

Corollary 5.13. *Chains are not geodesic in Tamari lattices; more precisely, for every n, there exists a length ℓ chain of $\mathcal{T}_n$ whose endpoints are at distance less than $12\ell/n$.*

We conclude with a few open questions.

Question 5.14. *For w a positive $\boldsymbol{A}$-word, is the number $\lambda(w)$ of (3.3) a least upper bound for the length of the words that are $\boldsymbol{R}$-equivalent to w?*

The answer is positive in the case of $a_\emptyset^n$. Indeed, we have $\lambda(a_\emptyset^n) = n(n+1)/2$ and $a_\emptyset^n \equiv_{\boldsymbol{R}}^+ \hat{a}_{1^{n-1},1} \, \hat{a}_{1^{n-2},2} \ldots \hat{a}_{1,n-1} \, \hat{a}_{\emptyset,n}$. The general case is not known.

Question 5.15. *Does the Polish normal form of Definition 14 satisfy some Fellow Traveler Property, that is, is the distance between the paths associated with the normal forms of elements f and $f a_\alpha^{\pm 1}$ uniformly bounded?*

A positive solution would provide a sort of infinitary automatic structure on F. The question should be connected with a possible closure of Polish normal words under left- or right-$\boldsymbol{R}$-reversing

Finally, using mapping class groups and cell decompositions, D. Krammer constructed for every size n tree T an exotic lattice structure on $\mathcal{T}_n$ in which T is the bottom element [19].

Question 5.16. *Can the Krammer lattices be associated with submonoids of the Thompson group F?*

More generally, a natural combinatorial description of the Krammer lattices is still missing, but would be highly desirable. Connections with permutations and braids in the line of [2] can be expected.

References

[1] J.L. Baril and J.M.. Pallo, "Efficient lower and upper bounds of the diagonal-flip distance between triangulations", *Inform. Proc. Letters* **100** (2006) 131–136.

[2] A. Björner and M. Wachs, "Shellable nonpure complexes and posets. II", *Trans. Amer. Math. Soc.* **349** (1997) 3945–3975.

[3] M. Brin and C. Squier, "Groups of piecewise linear homeomorphisms of the real line", *Invent. Math.* **79** (1985) 485–498.

[4] J.W. Cannon, W.J. Floyd, and W.R. Parry, "Introductory notes on Richard Thompson's groups", *Enseign. Math.* **42** (1996) 215–257.

[5] A. Clifford and G. Preston, *The algebraic theory of semigroups, volume 1*, Amer. Math. Soc. Surveys, vol. 7, Amer. Math. Soc., 1961.

[6] P. Dehornoy, "The structure group for the associativity identity", *J. Pure Appl. Algebra* **111** (1996) 59–82.

[7] ———, "Groups with a complemented presentation", *J. Pure Appl. Algebra* **116** (1997) 115–137.

[8] ———, *Braids and Self-Distributivity*, Progress in Math., vol. 192, Birkhäuser, 2000.

[9] ———, "Study of an identity", *Algebra Universalis* **48** (2002) 223–248.

[10] ———, "Complete positive group presentations", *J. of Algebra* **268** (2003) 156–197.

[11] _____ , "Geometric presentations of Thompson's groups", *J. Pure Appl. Algebra* **203** (2005) 1–44.

[12] _____ , "On the rotation distance between binary trees", *Advances in Math.* **223** (2010) 1316–1355.

[13] _____ , "The word reversing method", *Intern. J. Alg. and Comput.* **21** (2011) 71–118.

[14] P. Dehornoy, with F. Digne, E. Godelle, D. Krammer, and J. Michel, "Garside Theory", Book in progress, http://www.math.unicaen.fr/∼garside/Garside.pdf.

[15] O. Deiser, "Notes on the Polish Algorithm", http://page.mi.fu-berlin.de/deiser/wwwpublic/psfiles/polish.ps.

[16] H. Friedman and D. Tamari, "Problèmes d'associativité : Une structure de treillis finis induite par une loi demi-associative", *J. Combinat. Th.* **2** (1967) 215–242.

[17] V.S. Guba, "The Dehn function of Richard Thompson's group F is quadratic", *Invent. Math.* **163** (2006) 313–342.

[18] S. Huang and D. Tamari, "Problems of associativity: A simple proof for the lattice property of systems ordered by a semi-associative law", *J. Combinat. Th. Series A* **13** (1972) 7–13.

[19] D. Krammer, "A class of Garside groupoid structures on the pure braid group", *Trans. Amer. Math. Soc.* **360** (2008) 4029–4061.

[20] S. Mac Lane, *Natural associativity and commutativity*, Rice University Studies, vol. 49, 1963.

[21] R. McKenzie and R.J. Thompson, "An elementary construction of unsolvable word problems in group theory", in *Word Problems*, Boone and *al*, eds., Studies in Logic, vol. 71, North Holland, 1973, 457–478.

[22] J.M. Pallo, "Enumerating, ranking and unranking binary trees", *The Computer Journal* **29** (1986) 171–175.

[23] _____ , "An algorithm to compute the Möbius function of the rotation lattice of binary trees", *RAIRO Inform. Théor. Applic.* **27** (1993) 341–348.

[24] D. Sleator, R. Tarjan, and W. Thurston, "Rotation distance, triangulations, and hyperbolic geometry", *J. Amer. Math. Soc.* **1** (1988) 647–681.

[25] R. Stanley, *Enumerative Combinatorics, vol. 2*, Cambridge Studies in Advances Math., no. 62, Cambridge Univ. Press, 2001.

[26] J.D. Stasheff, "Homotopy associativity of H-spaces", *Trans. Amer. Math. Soc.* **108** (1963) 275–292.

[27] Z. Šunić, "Tamari lattices, forests and Thompson monoids", *Europ. J. Combinat.* **28** (2007) 1216–1238.

[28] D. Tamari, "The algebra of bracketings and their enumeration", *Nieuw Archief voor Wiskunde* **10** (1962) 131–146.

[29] R.J. Thompson, "Embeddings into finitely generated simple groups which preserve the word problem", in *Word Problems II*, Adian, Boone, and Higman, eds., Studies in Logic, North Holland, 1980, 401–441.

Laboratoire de Mathématiques Nicolas Oresme, UMR6139, Université de Caen, F-14032 Caen, France, Laboratoire de Mathématiques Nicolas Oresme, UMR6139, CNRS, F-14032 Caen, France

E-mail address: patrick.dehornoy@unicaen.fr