SOLVABILITY OF EQUATIONS IN GRAPH GROUPS IS DECIDABLE

VOLKER DIEKERT
Institut für Formale Methoden der Informatik (FMI),
Universität Stuttgart,
Universitätsstr. 38
D-70569 Stuttgart, Germany
diekert@fmi.uni-stuttgart.de

ANCA MUSCHOLL
LIAFA, Université Paris VII,
2, place Jussieu, case 7014,
F-75251 Paris Cedex 05, France
anca@liafa.jussieu.fr

Received June 2005
Accepted October 2005

Communicated by Olga Kharlampovich

We show that the existential theory of free partially commutative monoids with involution is decidable. As a consequence the existential theory of graph groups is also decidable. If the underlying alphabet of generators is fixed, we obtain a PSPACE-completeness result, otherwise (in the uniform setting) our decision procedure is in EXPSPACE. Our proof is a reduction to the main result of [6].

Keywords: Equations, graph groups, free partially commutative monoids.

1. Introduction

Solving equations in algebraic structures is a fundamental task in mathematics. Here we tackle this problem for *free partially commutative monoids with involution* and for *graph groups*, which are free groups with a partial commutation relation between generators.

Basic algebraic structures involving partial commutations are *free partially commutative monoids* (also known as *trace monoids*). They were considered in combinatorics by Cartier and Foata [5] and in computer science by Keller [14] and Mazurkiewicz [20,21]. Trace monoids serve as an algebraic tool for investigating concurrent systems. Atomic actions are represented by letters and independency of actions is reflected by a partial commutation relation. If each atomic action a has an inverse $\overline{a}$ such that $a\overline{a} = \overline{a}a = 1$, then, on the algebraic level, we switch from monoids to groups. Without the cancellation law $a\overline{a} = \overline{a}a = 1$, we obtain *free partially commutative monoids with involution.* It turns out that our decidability result on graph groups follows from the corresponding result on free partially commutative monoids with involution. Therefore we focus on the latter objects.

We show that the existential theory of equations with recognizable constraints in free partially commutative monoids with involution is decidable. If the underlying alphabet of generators is fixed, then we obtain a PSPACE-completeness result, otherwise (in the uniform setting) our decision procedure is in EXPSPACE. In the conference version [10] we gave a non-elementary uniform decision procedure, hence we obtain a significant improvement here.

The relation of our work to previous results on existential theories of equations is as follows. In the simplest setting we ask whether a single word equation with constants is solvable. This problem is easily seen to be NP-hard. It becomes PSPACE-hard, as soon as we add regular constraints for the unknowns, simply because the intersection problem for regular languages is PSPACE-complete, [16]. Makanin proved the decidability of word equations [17] and Schulz extended this decidability result in order to include regular constraints [27]. By standard methods this means that the existential theory of word equations (with regular constraints) is decidable. The situation in free groups turned out to be much more complicated.

Makanin also proved that the existential theory in free groups is decidable [18]. However, in that case the scheme of Makanin has been shown to be not-primitive recursive, see [15]. Only when Plandowski invented a new method for solving word equations by some polynomial space bounded algorithm [25], the corresponding problem for free groups was reconsidered; and Gutiérrez succeeded in extending Plandowski's polynomial space algorithm to free groups [13].

The situation in trace monoids is more complicated due to the partial commutation which cannot be expressed simply by equations. To be more precise, over traces an equation like $XY = YX$ is implied by independency, so the equation may have many non-trivial solutions, in contrast to the situation in free monoids or free groups. To overcome this difficulty one is led to work with recognizable constraints over trace monoids, which in the reduction to the free case become regular constraints. So, when Matiyasevich showed in 1996 that the existential theory of free partially commutative monoids is decidable [19,9], Schulz's generalization of Makanin's result was used.

In the present paper we show decidability for graph groups, which is an extension of the result on the existential theory of equations with rational constraints in free groups. However, for graph groups rational constraints are too powerful, in general: They lead to undecidability (this follows from Theorem 22). The good notion turned out to be *normalized regular constraints*, which are introduced here. Our decidability proof does not reduce equations in graph groups to equations free groups, but rather to equations with regular constraints over *free monoids with involution.* Then one can apply [6].

Based on the conference version [10], the results of the present paper have been extended together with Lohrey in order to obtain general transfer results for the existential and positive theories with respect to graph products, [7,8]. Formally, this leads to stronger statements than the results here, but the proofs in [7,8] rely on reductions to the main result of the present work, Theorem 20.

Solving equations in partially commutative structures also found a quite unexpected application in combinatorial topology. In [26] the string graph recognition problem for a compact surface of higher genus is reduced to (quadratic) equations in free partially commutative monoids with involution.

2. Preliminaries

2.1. *Free partially commutative monoids with involution*

Throughout the paper Γ means a finite alphabet which is equipped with an involution $\overline{} : \Gamma \to \Gamma$. An involution is a mapping such that $\overline{\overline{a}} = a$ for all $a \in \Gamma$. The involution $\overline{} : \Gamma \to \Gamma$ is extended to the free monoid Γ^* by $\overline{a_1 \cdots a_n} = \overline{a_n} \cdots \overline{a_1}$. In particular, the involution reverses the order. A symmetric and irreflexive relation $I \subseteq \Gamma \times \Gamma$ (such that $(a, b) \in I$ implies $(\overline{a}, b) \in I$) is called an *independence relation* (which is *compatible with the involution*). Its complement $D = (\Gamma \times \Gamma) \setminus I$ is called a *dependence* relation. The relation D is reflexive and symmetric (and we have $(a, b) \in D$ implies $(\overline{a}, b) \in D$ for all $a, b \in \Gamma$). In the following all (in-)dependence relations are compatible with the involution. In particular, we have $(a, \overline{a}) \in D$ for all $a \in \Gamma$. The *free partially commutative* monoid $M(\Gamma, I)$ is defined by the quotient monoid $\Gamma^*/\{ab = ba \mid (a, b) \in I\}$. According to Mazurkiewicz [20] it is also called a *trace monoid* and its elements (which are congruence classes) are called *traces*. For an overview of trace theory we refer to [11].

If the reference to (Γ, I) is clear, we also write $\mathbb{M}$ instead of $M(\Gamma, I)$. The length $|x|$ of a trace x is the length of any representing word. A letter $a \in \Gamma$ is called *minimal* (*maximal* resp.) in x, if we can write $x = ay$ ($x = ya$ resp.) for some $y \in \mathbb{M}$. The set of minimal (maximal resp.) elements consists of pairwise independent letters. For $a \in \Gamma$ let $I(a) = \{b \in \Gamma \mid (a, b) \in I\}$. The set of letters occurring in $x \in \Gamma^*$ or in $x \in \mathbb{M}$ is denoted $\mathrm{alph}(x)$ and by $I(x)$ we mean $I(x) = \bigcap_{a \in \mathrm{alph}(x)} I(a)$.

By 1 we denote the empty word, the empty trace, and the unit element in a group.

We shall use node-labeled directed acyclic graphs $[V, E, \lambda]$ in order to represent traces. Here V is the set of vertices, E is the edge set, and $\lambda : V \to \Gamma$ is the labeling. Such a graph induces a labeled partial order $[V, E^*, \lambda]$, and a labeled partial order is also called a *partially ordered multi set* or *pomset* for short.

In our setting we assume that V is finite and that $(\lambda(v), \lambda(v')) \in D$ implies either $(v, v') \in E^*$ or $(v', v) \in E^*$, so all dependent vertices are ordered. Thus, $[V, E, \lambda]$ defines a unique trace $x = [V, E, \lambda] \in \mathbb{M}$ in a canonical way by taking the congruence class of any linearization. If we start with a finite word $a_1 \cdots a_n$ for representing a trace, then we may take $V = \{1, \ldots, n\}$. Each i is viewed as a node with label $\lambda(i) = a_i$. We define an arc from a_i to a_j if and only if both, $i < j$ and $(a_i, a_j) \in D$. In this way we obtain a node-labeled directed acyclic graph $[V, E, \lambda]$, which is called the *dependence graph* of the trace $[a_1 \cdots a_n]$. A dependence graph $[V, E, \lambda]$ represents a trace $x \in \mathbb{M}$ and, in addition, $(\lambda(v), \lambda(v')) \in D$ is equivalent to $(v, v') \in \mathrm{id}_\Gamma \cup E \cup E^{-1}$. Up to isomorphism, the dependence graph of x is unique, and so is its induced pomset $[V, E^*, \lambda]$ which is also denoted by $[V, \leq, \lambda]$.

Since the independence relation I is supposed to be compatible with the involution, the involution transfers to traces. If we have $x = [a_1 \cdots a_n] \in \mathbb{M}$, then $\overline{x} = [\overline{a_n} \cdots \overline{a_1}]$, $\overline{[V, E, \lambda]} = [V, E^{-1}, \overline{\lambda}]$ and $\overline{[V, \leq, \lambda]} = [V, \geq, \overline{\lambda}]$ respectively, where $\overline{\lambda}(v) = \overline{\lambda(v)}$ for all $v \in V$. Hence, $(\mathbb{M}, \overline{})$ is a trace monoid with involution. A monoid with involution satisfies $\overline{1} = 1$, $\overline{\overline{x}} = x$, and $\overline{xy} = \overline{y}\,\overline{x}$. (In particular, every group is a monoid with the involution defined by taking inverses.)

2.2. *Factors*

Let $x = [V, \leq, \lambda] \in \mathbb{M}$ be a trace where $\leq$ is the partial order induced by the dependence graph. A *factor* or *factor trace* is a trace $f \in \mathbb{M}$ such that we can write $x = pfq$. Given a factorization $x = pfq$, there is some $F \subseteq V$ such that the induced pomset of F in $[V, \leq, \lambda]$ represents f. Moreover, F has the property that whenever $v \leq v' \leq v''$ with $v, v'' \in F$, then $v' \in F$, too. Conversely, let $F \subseteq V$ be a subset with this property: $v \leq v' \leq v''$ with $v, v'' \in F$ implies $v' \in F$. Then we can factorize $x = pfq$ such that F represents the factor f, but due to partial commutation the factorization is not unique (in contrast to the case of words). There is another difference between words and traces. Assume that $V = F \cup G$ is a disjoint union where F represents a factor f and G represents a factor g. Then this does not mean that x is a product of f and g, in general. Indeed, let $\mathbb{M} = \{a, b\}^* \times \{c, d\}^*$ and $x = abcd$. Then we have $abcd = cadb$, so both $f = bc$ and $g = ad$ are factors. But x is not a product of f and g.

2.3. *Clans and the parameter τ*

In order to have a convenient complexity bound below we define an equivalence relation on Γ such that a and b are in one class, if $D(a) = D(b)$. The number of equivalence classes is denoted by $c(\Gamma, D)$ and each class $[a] = \{b \in \Gamma \mid D(a) = D(b)\}$ is called a *complete clan*, or in the following simply *clan* for short. Thus, a clan in (Γ, D) is a maximal subset $A \subseteq \Gamma$ such that $(a, c) \in D \Leftrightarrow (b, c) \in D$ for all $a, b \in A$ and $c \in \Gamma$. Note that a clan is indeed a complete subgraph of (Γ, D), since D is reflexive. A clan A is called *thin*, if there are $a \in A$, $b \in \Gamma \setminus A$ such that $(a, b) \in I$, otherwise it is called *thick*. There is at most one thick clan due to maximality. The number of thin clans is denoted by $\tau(\Gamma, D)$ or τ for short. It is either $c(\Gamma, D)$ or $c(\Gamma, D) - 1$, it is never 1. If $\mathbb{M}$ is a direct product of d free monoids, then the number τ of thin clans is d for $d > 1$, and it is 0 for $d = 1$. In the following we pick a thin clan and we make it thick by removing independency. It might be that the parameter $c(\Gamma, D)$ does not change, but the parameter τ decreases. This is the reason why the induction below is based on the parameter τ instead of the number of clans.

3. Normal forms

In this section (and in this section only) we assume that the involution $\overline{} : \Gamma \to \Gamma$ is without fixed points, i.e., $\overline{a} \neq a$ for all $a \in \Gamma$. We fix some thin clan of (Γ, D), which we write in the form $A \cup \overline{A}$ with $A \cap \overline{A} = \emptyset$. We define

$$\widehat{D} = D \cup \Gamma \times (A \cup \overline{A}) \cup (\overline{A} \cup A) \times \Gamma.$$

The subset $A \cap \overline{A}$ is not a thin clan anymore with respect to $\widehat{D}$ since it is now included in some thick clan. Let $\widehat{I} = (\Gamma \times \Gamma) \setminus \widehat{D}$, then $\widehat{\mathbb{M}} = M(\Gamma, \widehat{I})$ is a trace monoid with involution, and the number of thin clans in $(\Gamma, \widehat{D})$ is at most $\tau - 1$.

3.1. *Source, median, and target positions*

By $\hat{\psi}$ we denote the canonical homomorphism $\hat{\psi} : \widehat{\mathbb{M}} \to \mathbb{M}$. The aim is to define a normal form mapping $\mathrm{nf} : \mathbb{M} \to \widehat{\mathbb{M}}$ which is compatible with the involution, i.e., we demand $\hat{\psi}(\mathrm{nf}(x)) = x$ and $\mathrm{nf}(\overline{x}) = \overline{\mathrm{nf}(x)}$. Using a classical normal form such as the representation by the lexicographical first word does not work, as shown in Remark 32. The existence of a suitable normal form relies on the following simple lemma:

Lemma 1. *Let $a \in \Gamma$ such that $a \neq \overline{a}$ and let $w \in \{a, \overline{a}\}^*$ be any word. Then there exists a unique $k \geq 0$ such that $w \in a^*(\overline{a}a^*)^k(\overline{a}^*a)^k\overline{a}^*$.*
For the same k we also have $\overline{w} \in a^*(\overline{a}a^*)^k(\overline{a}^*a)^k\overline{a}^*$.

Proof. Let ℓ be the number of a in w and write $w = uv$ where $|u| = \ell$. Then k is the number of $\overline{a}$ in u and we have $u \in a^*(\overline{a}a^*)^k$ and $v \in (\overline{a}^*a)^k\overline{a}^*$. Moreover, $\overline{w} = \overline{v}\,\overline{u}$, $\overline{v} \in a^*(\overline{a}a^*)^k$ and $\overline{u} \in (\overline{a}^*a)^k\overline{a}^*$. □

Recall that we view a trace x as a labeled pomset $[V, \leq, \lambda]$, $\lambda : V \to \Gamma$. Often, we write $v \in x$ instead of $v \in V$ and we write $u \parallel v$ whenever $u, v \in V$ are incomparable with respect to $\leq$. We also write $(u, v) \in I$ when $(\lambda(u), \lambda(v)) \in I$, and analogously for $D = (\Gamma \times \Gamma) \setminus I$. Of course, $u \parallel v$ implies $(u, v) \in I$. Let $a_1 < \cdots < a_q$ be the linearly ordered subset of $(V, \leq)$ containing all vertices with label in the clan $A \cup \overline{A}$. We might have $q = 0$ meaning that there are no vertices with label in $A \cup \overline{A}$. We read $a_1 < \cdots < a_q$ as a word $a_1 \cdots a_q$ in the free monoid $(A \cup \overline{A})^*$. With each vertex $v \in V$ we associate the maximal factor of $a_1 \cdots a_q$ consisting of the vertices w with label in $A \cup \overline{A}$ which are incomparable with v, i.e., $w \parallel v$. For $v \in V$ the *source* $s(v)$ and the *target point* $t(v)$ of v are defined as follows.

$$s(v) = \sup \{i \mid a_i \leq v\}, \quad t(v) = \inf \{i \mid v \leq a_i\}.$$

By convention, $\sup \emptyset = 0$ and $\inf \emptyset = q + 1$. Thus, $0 \leq s(v) \leq q$, $1 \leq t(v) \leq q + 1$ and $s(v) \leq t(v)$ for all $v \in V$. Note that we have $s(v) = t(v)$ if and only if the label of v belongs to $A \cup \overline{A}$.

For $0 \leq s \leq t \leq q+1$ we define the *median position* $m(s,t)$: For $s = t$ we let $m(s,t) = s$. For $s < t$ we choose by Lemma 1 the unique c with $s \leq c < t$ and $k \geq 0$ such that

$$a_{s+1} \cdots a_c \in A^*(\overline{A}A^*)^k,$$
$$a_{c+1} \cdots a_{t-1} \in (\overline{A}^* A)^k \overline{A}^*.$$

Then we define $m(s,t) = c + \frac{1}{2}$ and we call $m(s,t)$ the *median position*. The median position $m(s(v), t(v))$ is called the *global position* of a vertex $v \in V$, it is denoted by $g(v)$, i.e., $g(v) = m(s(v), t(v))$.

Lemma 2. *Let $x = [V, \leq, \lambda]$ and $v, w \in V$ be vertices such that $v \leq w$. Then we have $s(v) \leq s(w)$, $t(v) \leq t(w)$, and $g(v) \leq g(w)$.*

Proof. Obviously, $s(v) \leq s(w)$ and $t(v) \leq t(w)$. The lemma now follows from the fact $m(s-1,t) \leq m(s,t) \leq m(s,t+1)$, which in turn is easy to verify. □

We are ready to define the normal form $\mathrm{nf}(x) \in \widehat{\mathbb{M}}$ of a trace $x \in \mathbb{M}$. We do so by introducing *new arcs* into the dependence graph $[V, E, \lambda]$ of x: Let $v, w \in V$ such that $\lambda(w) \in A \cup \overline{A}$ and $v \parallel w$. (In particular, $\lambda(v) \notin A \cup \overline{A}$ and $g(v) \neq g(w)$ because $g(v) \in \mathbb{N} + \frac{1}{2}$, while $g(w) \in \mathbb{N}$.) We define a new arc from v to w, if $g(v) < g(w)$, otherwise we define a new arc from w to v. The arcs being already present in the dependence graph of x are called *old arcs*. The union $\hat{E}$ of old and new arcs defines a labeled directed graph $[V, \hat{E}, \lambda]$.

Lemma 3. *Let $x = [V, \hat{E}, \lambda] \in \mathbb{M}$ and consider vertices $u, v \in V$. For $u \leq v$ in $[V, \hat{E}, \lambda]$ we have $g(u) \leq g(v)$. Moreover, we have $g(u) < g(v)$ whenever there is a path from u to v in $[V, \hat{E}, \lambda]$ using at least one new arc.*

Proof. Let $(u, v) \in \hat{E}$. Then $g(u) \leq g(v)$ by Lemma 2. Moreover, if (u, v) is a new arc then $g(v) - g(u) \geq 1/2$. □

The lemma says that following a directed path in $\hat{E}$ never decreases the global position. It is however increased as soon as we use a new arc. Therefore the graph $[V, \hat{E}, \lambda]$ is acyclic since $[V, E, \lambda]$ has been acyclic. Thus, $[V, \hat{E}, \lambda]$ defines a unique trace $\mathrm{nf}(x) = [V, \hat{E}, \lambda]$ of $\widehat{\mathbb{M}}$. The important property of the normal form is $\mathrm{nf}(\overline{x}) = \overline{\mathrm{nf}(x)}$. We state this as a lemma.

Lemma 4. *Let the involution $\overline{} : \Gamma \to \Gamma$ be without fixed points and let the normal form be defined as above. Then we have $\mathrm{nf}(\overline{x}) = \overline{\mathrm{nf}(x)}$ for all $x \in \mathbb{M}$.*

Proof. Let $(a, b) \in I$ and $w \in \{a, \overline{a}\}^*$, $a \in A$. By Lemma 1 there exists a unique $k \geq 0$ such that $w \in a^*(\overline{a}a^*)^k(\overline{a}^*a)^k\overline{a}^*$. Write $w = uv$ with $u \in a^*(\overline{a}a^*)^k$ and $v \in (\overline{a}^*a)^k\overline{a}^*$. Then $\mathrm{nf}(wb) = ubv$ and $\mathrm{nf}(\overline{wb}) = \overline{v}\overline{b}\overline{u}$. The claim of the lemma follows easily from this fact, since a trace is uniquely defined by its projections to subalphabets of at most two letters. □

Remark 5. Let $x = [V, E, \lambda]$, $\overline{x} = [V, E^{-1}, \overline{\lambda}]$ be in $\mathbb{M}$. Then the global position of $v \in V$ in $\mathrm{nf}(\overline{x})$ is $q + 1 - g(v)$.

Example 6. Let $x = a\,a\,\overline{a}\,\overline{a}\,a\,\overline{a}\,a\,\overline{a}\,\overline{a}\,a\,a\,b$ and $(a, b) \in I$. Then $k = 3$, $g(b) = 6\frac{1}{2}$ and $\mathrm{nf}(x) = a\,a\,\overline{a}\,\overline{a}\,a\,\overline{a}\,b\,a\,\overline{a}\,\overline{a}\,a\,a$.

Remark 7. If the involution has fixed points, then a normal form satisfying $\mathrm{nf}(\overline{x}) = \overline{\mathrm{nf}(x)}$ for all $x \in \mathbb{M}$ cannot exist, in general. Indeed, assume we were in the situation $a, b \in A$, $a = \overline{a}$, $b = \overline{b}$, and $(a, b) \in I$. Then $\overline{ab} = ab$, but $(a, b) \in \widehat{D}$, so necessarily $ab \neq ba$ and hence $\mathrm{nf}(\overline{ab}) \neq \overline{\mathrm{nf}(ab)}$ in $\widehat{\mathbb{M}}$.

3.2. *Lifting a factorization to normal forms*

The results of this paper are based on the following theorem. We will apply the theorem to equations of the form $x = yz$, reducing a system of equations over $\mathbb{M}$ to one over $\widehat{\mathbb{M}}$. The important point in the statement is that the bound on d depends only on the number of thin clans $\tau(\Gamma, D)$.

Theorem 8. *Let the involution* $^{-} : \Gamma \to \Gamma$ *be without fixed points,* $\tau = \tau(\Gamma, D)$ *and let* $x, x_1, x_2 \in \mathbb{M}$ *be traces. Then the following assertions are equivalent.*

(i) $x = x_1 x_2$.

(ii) There exist $d \leq 3\tau^2 + 8\tau + 7$, *traces* $y_1, \ldots, y_d \in \widehat{\mathbb{M}}$, *an index* $0 \leq c \leq d$, *and a permutation* $\pi \in \mathrm{Perm}(d)$ *with the following properties:*

$$\begin{aligned}
\mathrm{nf}(x) &= y_{\pi(1)} \cdots y_{\pi(d)}, \\
\mathrm{nf}(x_1) &= y_1 \cdots y_c, \\
\mathrm{nf}(x_2) &= y_{c+1} \cdots y_d, \\
\mathrm{alph}(y_i) \times \mathrm{alph}(y_j) &\subseteq I \textit{ for all } i, j, \textit{ where } (i - j)(\pi(i) - \pi(j)) < 0.
\end{aligned}$$

The assertion $(ii) \Rightarrow (i)$ of Theorem 8 is trivial. The proof of the other direction $(i) \Rightarrow (ii)$ covers the rest of this section. Consider $x, x_1, x_2 \in \mathbb{M}$ such that $x = x_1 x_2$. We present x by its pomset $[V, \leq, \lambda]$ and we let $a_1 < \cdots < a_q$ be the linearly ordered subset of all vertices in V which have a label in the thin clan $A \cup \overline{A}$. We define p to be the index such that $a_1, \ldots, a_p \in x_1$ and $a_{p+1}, \ldots, a_q \in x_2$. We have $0 \leq p \leq q$. The values of p and q are fixed until the end of this section. We allow $q = 0$ although this case is trivial. However even if $p = 0$ or $p = q$ then we might have $\mathrm{nf}(x) \neq \mathrm{nf}(x_1)\,\mathrm{nf}(x_2)$, c.f. Example 6 above with $x_2 = b$.

Above we have introduced the notion of global position. In order to determine how $\mathrm{nf}(x)$ can be obtained from $\mathrm{nf}(x_1), \mathrm{nf}(x_2)$, we define the notion of *local position*, too. The local position $\ell(v)$ is the global position of v in x_1, if v belongs to x_1. If v belongs to x_2, then $\ell(v)$ is the global position of v in x_2 plus p, since we define $\ell(v)$ in x. More formally, suppose that v is in x_1. With respect to x_1, the target point $t'(v)$ of v is $\min\{p + 1, t(v)\}$, hence we define:

$$\ell(v) = m(s(v), \min\{p + 1, t(v)\})\,.$$

Similarly, if v belongs to x_2 then the source point $s'(v)$ of v with respect to x_2 is $s'(v) = \max\{s(v), p\}$ and we define:

$$\ell(v) = m(\max\{s(v), p\}, t(v))\,.$$

The next lemma summarizes some direct consequences of the definition of global and local positions. The proof is omitted.

Lemma 9. *Let $x = x_1x_2 \in \mathbb{M}$ with value p as above and let v, v' be vertices in $x = [V, \leq, \lambda]$.*

(1) For $v \leq v'$ we have $s(v) \leq s(v')$, $t(v) \leq t(v')$, $g(v) \leq g(v')$, and $\ell(v) \leq \ell(v')$.
(2) For $v \in x_1$ we have $\ell(v) \leq g(v)$ and $\ell(v) \leq p + 1/2$.
(3) For $v \in x_2$ we have $g(v) \leq \ell(v)$ and $p + 1/2 \leq \ell(v)$.
(4) For $\ell(v) \neq g(v)$ we have $s(v) \leq p < t(v)$.

The next proposition is a crucial technical result. It shows that among the vertices v with $s(v) \leq p < t(v)$, i.e., among the vertices where the local and global positions may differ, there is a few number of different source points. For a given trace $x \in \mathbb{M}$ we define

$$S = \{s(v) \mid s(v) \leq p < t(v), v \in x\}.$$

Proposition 10. *Let $x \in \mathbb{M}$ and S be defined as above. Then we have $|S| \leq \tau + 1$.*

Proof. We may assume that $|S| \geq 2$. Choose a sequence $b_1, \ldots, b_k, b_{k+1} \in x$ with k minimal such that $S = \{s(b_i) \mid 1 \leq i \leq k+1\}$. We may assume that $0 \leq s(b_1) < \cdots < s(b_k) < p$ and $p < t(b_i)$ for $1 \leq i \leq k$. Hence we have $b_i \parallel a_p$ for all $1 \leq i \leq k$ (see also Figure 1 below). We will see that $k \leq \tau$. For each $2 \leq i \leq k$ we choose a path from $a_{s(b_i)}$ to b_i in the dependence graph of x. On this path we pick a last vertex c_i with $c_i \leq a_p$. This vertex is not b_i. Hence, there is a next vertex d_i with $(c_i, d_i) \in D$, $c_i < d_i \leq b_i$, and $d_i \parallel a_p$ for $2 \leq i \leq k$. We choose $d_1 = b_1$. We claim that $\{d_i\} \times \{c_{i+1}, \ldots, c_k\} \subseteq I$ for $1 \leq i \leq k$. Indeed, assume by contradiction that $(d_i, c_j) \in D$ for some $i < j \leq k$. If we would have $c_j \leq d_i$, then we obtain $a_{s(b_j)} \leq c_j \leq d_i \leq b_i$, but this contradicts $s(b_i) < s(b_j)$. Hence $d_i \leq c_j$. But we have $c_j \leq a_p$, hence $d_i \leq a_p$ in contradiction to $d_i \parallel a_p$. This yields the claim.

$$\begin{tikzcd}
a_{s(b_i)} \arrow[r, "*"] & c_i \arrow[r] \arrow[dr, "+"'] & d_i \arrow[r, "*"] & b_i \\
& & a_p &
\end{tikzcd}$$

Fig. 1.

Next, $(d_i, c_i) \in D$ and $\{d_i\} \times \{c_{i+1}, \ldots, c_k\} \subseteq I$ imply that c_i and c_j are in different thin clans for all $2 \leq i < j \leq k$. Moreover, there is another thin clan containing d_1. $\square$

Analogously to Proposition 10 we have $|T| \leq \tau + 1$, where T is the set of target points: $T = \{t(v) \mid s(v) \leq p < t(v), v \in x\}$.

The set of *cutting points* C is defined to be the union $C = C_g \cup C_\ell \cup \{0, p+\frac{1}{2}, q+1\}$, where

$$C_g = \{g(v) \mid s(v) \leq p < t(v), v \in x\},$$
$$C_\ell = \{\ell(v) \mid s(v) \leq p < t(v), v \in x\}.$$

Proposition 11. *The number of cutting points is bounded by* $\tau^2 + 4\tau + 6$.

Proof. We have $C_g \subseteq \{m(s,t) \mid s \in S, t \in T\}$. Hence $|C_g| \leq (\tau+1)^2$ by Proposition 10 and the analogous statement for T. For C_ℓ we can write $C_\ell \subseteq \{m(s, p+1) \mid s \in S\} \cup \{m(p,t) \mid t \in T\}$. Hence $|C_\ell| \leq 2\tau + 2$. $\square$

The cutting points split the real interval $[0, q+1]$ into open intervals of the form (i,j) where $i, j \in C, i < j$, and $(i,j) \cap C = \emptyset$. There are $|C| - 1$ such intervals, so the number is at most $\tau^2 + 4\tau + 5$. With each interval (i,j) we associate a factor trace of x. The factor trace is denoted either $x[0;i,j]$ or $x[3;i,j]$; we call $x[m;i,j]$, $0 \leq m \leq 3$, a *segment*. (Later on we will define segments with index 1 and 2.) Since $p + \frac{1}{2}$ is a cutting point we have either $j \leq p + \frac{1}{2}$ or $p + \frac{1}{2} \leq i$. For $j \leq p + \frac{1}{2}$ we define

$$x[0;i,j] = \{v \in x \mid i < g(v) \leq j \text{ and } t(v) \leq p\}.$$

For $p + \frac{1}{2} \leq i$ we define

$$x[3;i,j] = \{v \in x \mid i \leq g(v) < j \text{ and } p < s(v)\}.$$

Note that we have $\ell(v) = g(v)$ for all $v \in x[0;i,j]$ or $v \in x[3;i,j]$. A segment as defined above is just a set of vertices of x. However, by Lemma 2 it is easy to see that a segment defines a factor trace of x and a factor trace of either x_1 or x_2. In the lemma below, we show that this property is still true for the normal forms $\mathrm{nf}(x), \mathrm{nf}(x_1), \mathrm{nf}(x_2)$.

Lemma 12. *The segments* $x[0;i,j]$ *and* $x[3;i,j]$ *define factor traces* y *of* $\mathrm{nf}(x)$. *If* $j \leq p + \frac{1}{2}$, *then* $x[0;i,j]$ *also defines a factor trace* y_1 *of* $\mathrm{nf}(x_1)$ *and we have* $y = y_1$ *in* $\widehat{\mathbb{M}}$. *If* $p + \frac{1}{2} \leq i$, *then* $x[3;i,j]$ *also defines a factor trace* y_2 *of* $\mathrm{nf}(x_2)$ *and we have* $y = y_2$ *in* $\widehat{\mathbb{M}}$.

Proof. By symmetry we may assume $j \leq p + \frac{1}{2}$. We first show that $x[0;i,j]$ is a factor trace of both $\mathrm{nf}(x), \mathrm{nf}(x_1)$. Assume that $u < v < w$ either in $\mathrm{nf}(x)$ or in $\mathrm{nf}(x_1)$, with $u, w \in x[0;i,j]$. Thus, we have $i < g(u) = \ell(u) \leq j$ and $i < g(w) =$

$\ell(w) \leq j$. By Lemma 9 we obtain $i < g(v) = \ell(v) \leq j$, using that $u < v < w$ in either $\mathrm{nf}(x)$ or $\mathrm{nf}(x_1)$. Thus, if $t(v) \leq p$, then $x[0; i, j]$ is a factor trace.

Assume otherwise that we have $t(v) > p$. In addition, we have also $s(v) \leq p$, because otherwise $p + 1 \leq g(v)$. But $g(v) \leq g(w) \leq p$. Thus, $g(v)$ and $\ell(v)$ are cutting points. Since $t(w) \leq p < t(v)$, the arc from v to w must be a new arc. If $v \to w$ is a new (global) arc in $\mathrm{nf}(x)$, then $i < g(u) \leq g(v) < g(w) \leq j$. If $v \to w$ is a new (local) arc in $\mathrm{nf}(x_1)$, then $i < \ell(u) \leq \ell(v) < \ell(w) \leq j$. In both cases we obtain a contradiction since the open interval (i, j) does not contain any cutting point. Hence $x[0; i, j]$ defines a factor trace y of $\mathrm{nf}(x)$ and a factor trace y_1 of $\mathrm{nf}(x_1)$. We have $y = y_1$ in $\widehat{\mathbb{M}}$ since $\ell(v) = g(v)$ for all v with $t(v) \leq p$. $\square$

The segments $x[0; i, j]$ and $x[3; i, j]$ are pairwise disjoint subsets of x, but they do not cover x, in general. The missing points are those vertices v where $s(v) \leq p < t(v)$. Therefore for each $m = 1, 2$ and $i \leq p < j$ we define an m-segment by:

$$x[m; i, j] = \{v \in x_m \mid s(v) = i \text{ and } t(v) = j\}\,, \quad m = 1, 2\,.$$

Note that segments are non empty, only if $(i, j) \in S \times T$. The total number of non empty 1 and 2 segments is therefore at most $2(\tau + 1)^2 = 2\tau^2 + 4\tau + 2$ by Proposition 10. Moreover, all vertices $v \in x[m; i, j]$ have the same local position $\ell(v)$ and they have the same global position $g(v)$, which are both cutting points. Of course, $\ell(v) \neq g(v)$ is possible.

Lemma 13. *Each m-segment $x[m; i, j]$, $m = 1, 2$ defines a factor trace y of $\mathrm{nf}(x)$ and a factor trace y_m of $\mathrm{nf}(x_m)$. We have $y = y_m$ in $\widehat{\mathbb{M}}$.*

Proof. Consider vertices $u, v, w \in x = [V, \leq, \lambda]$ with $u, w \in x[m; i, j]$. Assume that $\mathrm{nf}(x)$ (or $\mathrm{nf}(x_m)$, resp.) contains a path from u to w via v. Since $\ell(u) = \ell(w)$ and $g(u) = g(w)$, by Lemma 3 we note that the path cannot use any new arc. Hence, the path is already present in $x = [V, \leq, \lambda]$. Since $x[m; i, j]$ is a factor trace of both x and x_m, we see that $x[m; i, j]$ is also a factor trace of both $\mathrm{nf}(x)$ and $\mathrm{nf}(x_m)$. Moreover, $y = y_m$ in $\widehat{\mathbb{M}}$ is also clear, since there are no new arcs in $x[m; i, j]$. $\square$

Let us summarize the notations we have introduced up to this point. For each $m = 0, 1, 2, 3$ we have defined m-segments of the form $x[m; i, j]$ which are pairwise disjoint subsets of x and cover x. Hence $F = \{x[m; i, j] \mid x[m; i, j] \neq \emptyset\}$ is a partition of the set x. Each $x[m; i, j]$ with $m \leq 1$ is a factor trace of both $\mathrm{nf}(x)$ and $\mathrm{nf}(x_1)$ (respectively, each $x[m; i, j]$ with $m > 1$ is a factor trace of both $\mathrm{nf}(x)$ and $\mathrm{nf}(x_2)$). By Propositions 10 and 11 we have $|F| \leq 3\tau^2 + 8\tau + 7$. Thus, we can choose the value d in Theorem 8 such that $d = |F|$.

By Lemma 12 and 13 we denote by $y_f \in \widehat{\mathbb{M}}$ the factor trace of $\mathrm{nf}(x)$ associated with $f \in F$. Since the segments in F cover all of $\mathrm{nf}(x), \mathrm{nf}(x_1)$ and $\mathrm{nf}(x_2)$, it remains to show how to write $\mathrm{nf}(x), \mathrm{nf}(x_1)$ and $\mathrm{nf}(x_2)$ as products of y_f, where f ranges over F. For this we have to compare segments. Every segment is associated with either a pair of consecutive cutting points or a single cutting point. For a non-empty

m-segment $f = x[m; i, j]$ we define a *global weight* $\omega_g(f)$ and a *local weight* $\omega_\ell(f)$ as follows. For $m \in \{0, 3\}$ let $\omega_g(f) = \omega_\ell(f) = \frac{i+j}{2}$, which is the center of the half-open interval associated with f. For $m \in \{1, 2\}$ let $\omega_g(f) = g(v)$ and $\omega_\ell(f) = \ell(v)$ for some $v \in f$. Recall that all $v \in f$ in this case have the same global position and the same local position.

We endow the set F of m-segments, $m \in \{0, 1, 2, 3\}$ with a global total order $\sqsubseteq_g$ and a local total order $\sqsubseteq_\ell$. In the following let $h \in \{g, \ell\}$, so h refers either to the global or to the local situation.

For $f = x[m; i, j]$, $f' = x[m'; i', j']$ we write $f \sqsubseteq_h f'$ if one of the following conditions holds:

(1) $\omega_h(f) < \omega_h(f')$.
(2) $\omega_h(f) = \omega_h(f')$, and $m < m'$.
(3) $\omega_h(f) = \omega_h(f')$, $m = m'$, and $i < i'$.
(4) $\omega_h(f) = \omega_h(f')$, $m = m'$, $i = i'$, and $j \leq j'$.

It is clear that $\sqsubseteq_g$ and $\sqsubseteq_\ell$ are both total orders. The next proposition has several important consequences. For example, it implies that $\sqsubseteq_g$ and $\sqsubseteq_\ell$ are both linearizations of the partial order $\leq$ of x.

Proposition 14. *Let $v, v' \in x$ be vertices such that $v \in f = x[m; i, j]$ and $v' \in f' = x[m'; i', j']$, where $f, f' \in F$. Then we have:*

- *If $v \leq v'$ or $g(v) < g(v')$ holds, then $f \sqsubseteq_g f'$.*
- *If $v \leq v'$ or $\ell(v) < \ell(v')$ holds, then $f \sqsubseteq_\ell f'$.*

Proof. We assume that $v \neq v'$ and $f \neq f'$. We distinguish several cases, depending on the values of m, m'. Recall from Lemma 9 that $v \leq v'$ implies $h(v) \leq h(v')$, where as above, h means either g or ℓ. It is of course possible that $v < v'$ and $h(v) = h(v')$.

(1) Let $m, m' \in \{0, 3\}$. The case $m = 3$, $m' = 0$ means $h(v') < h(v)$, so it cannot occur. Hence $m \leq m'$. We note that $h(v) \leq h(v')$, together with i, j and i', j' being consecutive cutting points, implies that $i \leq i'$ and $j \leq j'$. Hence, $\omega_h(f) = \frac{i+j}{2} \leq \omega_h(f') = \frac{i'+j'}{2}$. If $\omega_h(f) < \omega_h(f')$, then we are done. Otherwise, if we have also $m = m'$, then $i \leq i'$ and $j \leq j'$. Thus, in all cases $f \sqsubseteq_h f'$.
(2) Let $m = 0$ and $m' \in \{1, 2\}$. Since $h(v') \in C$ and i, j are consecutive cutting points we have $i < h(v) \leq j \leq h(v')$, hence $\omega_h(f) = \frac{i+j}{2} < h(v') = \omega_h(f')$.
(3) Let $m, m' \in \{1, 2\}$ and $m \leq m'$. Then we have $\omega_h(f) = h(v) \leq h(v') = \omega_h(f')$. If $h(v) = h(v')$, then we must have $v \leq v'$ by the assumption of the proposition. Thus, Lemma 9 yields $i \leq i'$ and $j \leq j'$.
(4) Let $m = 2$ and $m' = 1$. Clearly, we cannot have $v \leq v'$, thus $h(v) < h(v')$ holds, hence also $\omega_h(f) < \omega_h(f')$.
(5) Let $m \in \{1, 2\}$ and $m' = 3$. Then, dual to the second case, $h(v) \leq i' \leq h(v') < j$, thus $\omega_h(f) = h(v) < \frac{i'+j'}{2} = \omega_h(f')$.

(6) Let $m \in \{1,2\}$ and $m' = 0$. Clearly, we cannot have $v \le v'$, hence we must have $h(v) < h(v')$. Since i', j' are consecutive cutting points we infer that $h(v) \le i' < h(v') \le j'$, hence $\omega_h(f) = h(v) < \omega_h(f') = \frac{i'+j'}{2}$.

(7) Let $m = 3$ and $m' \in \{1,2\}$. As above, we must have $h(v) < h(v')$. Since i, j are consecutive cutting points, we obtain, similarly to the first case, $i \le h(v) < j \le h(v')$. Thus $\omega_h(f) = \frac{i+j}{2} < h(v') = \omega_h(f')$. □

Corollary 15. *Let $F = \{f_1, \ldots, f_d\}$ be sorted such that $f_i \sqsubseteq_g f_{i+1}$ for all i and let $y_i = y_{f_i}$ be the associated factors of* $\mathrm{nf}(x)$*. Then we have* $\mathrm{nf}(x) = y_1 \cdots y_d$.

Proof. Let $v, v' \in \mathrm{nf}(x)$ with $v \in f_i$ and $v' \in f_j$ such that there is an arc from v to v' in the dependence graph of $\mathrm{nf}(x)$. We have to show that $i \le j$. We know that $v < v'$ (if it is an old arc) or $g(v) < g(v')$ (if it is a new arc). By Proposition 14 we obtain $f_i \sqsubseteq_g f_j$, which is equivalent to $i \le j$. □

Corollary 16. *Let $F = \{f_1, \ldots, f_d\}$ be sorted such that $f_i \sqsubseteq_\ell f_{i+1}$ for all i and let $y_i = y_{f_i}$ be the associated factors of* $\mathrm{nf}(x)$*. Then there exists some c with $0 \le c \le d$ satisfying*

$$\mathrm{nf}(x_1) = y_1 \cdots y_c \quad \textit{and} \quad \mathrm{nf}(x_2) = y_{c+1} \cdots y_d\,.$$

Proof. We begin with the following observation. Let $f, f' \in F$ with $f = x[m; i, j]$, $f' = x[m'; i', j']$ and $m \in \{0,1\}$, $m' \in \{2,3\}$. Then $\omega_\ell(f) \le p + \frac{1}{2} \le \omega_\ell(f')$, hence $f \sqsubseteq_\ell f'$. Therefore in our $\sqsubseteq_\ell$-sorted sequence there is some c with $0 \le c \le d$ such that $y_i \subseteq \mathrm{nf}(x_1)$ if and only if $i \le c$ for all $1 \le i \le d$. It remains to show that $y_1 \cdots y_c = \mathrm{nf}(x_1)$ and $y_{c+1} \cdots y_d = \mathrm{nf}(x_2)$. This part is identical to the proof of Corollary 15 with the local view instead of the global one. □

Another immediate consequence of Proposition 14 is the following statement.

Corollary 17. *Let $f, f' \in F$, $f \neq f'$, be segments such that $f \sqsubseteq_\ell f'$. Then $f' \sqsubseteq_g f$ implies* $\mathrm{alph}(y_f) \times \mathrm{alph}(y_{f'}) \subseteq I$.

For the final step we consider the normal forms of x, x_1, x_2 and express them as products of factor traces y_f associated with segments $f \in F$. We sort F such that $f_1 \sqsubseteq_\ell \cdots \sqsubseteq_\ell f_d$. Let $\pi \in \mathrm{Perm}(d)$ be the permutation such that $f_{\pi(1)} \sqsubseteq_g \cdots \sqsubseteq_g f_{\pi(d)}$ is sorted with respect to $\sqsubseteq_g$. By Corollary 17 we have $\mathrm{alph}(y_i) \times \mathrm{alph}(y_j) \subseteq I$ whenever $(i-j)(\pi(i) - \pi(j)) < 0$. (As above, y_i is the factor trace associated with $f_i \in F$.) This completes the proof of Theorems 8.

Example 18. Let $\Gamma = \{a, \overline{a}, b, \overline{b}, c, \overline{c}\}$ with $(a,b) \in I$ and $(a,c), (b,c) \in D$. Consider the thin clan $\{a, \overline{a}\}$ and the situation given by Figure 2, i.e., $x = x_1 x_2$ with $x_1 = a\,\overline{a}\,a\,a\,\overline{a}\,b\,\overline{b}\,c\,\overline{a}\,\overline{a}\,\overline{b}$, $x_2 = a\,a\,\overline{a}\,b\,b\,c\,\overline{a}\,a\,a\,\overline{a}\,b\,b$ and $p = 7, q = 14$.

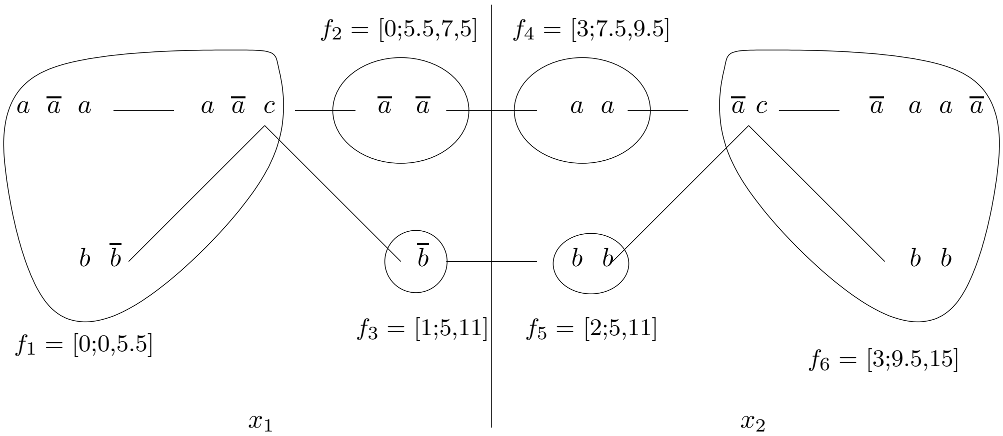

Fig. 2.

We list first the global and local positions of all $b, \overline{b}$:

b_i	1	2	3	4	5	6	7
$\ell(b_i)$	3.5	3.5	5.5	9.5	9.5	12.5	12.5
$g(b_i)$	3.5	3.5	7.5	7.5	7.5	12.5	12.5

For instance, let v be the first b (or $\overline{b}$, resp.) in this example. Then, $s(v) = 0$, $t(v) = 6$, and $g(v) = m(s(v), t(v)) = 3.5$.

We have as cutting points $C = \{0, 5.5, 7.5, 9.5, 15\}$. The 0-segments are $f_1 = [0; 0, 5.5]$, and $f_2 = [0; 5.5, 7.5]$. The 3-segments are $f_4 = [3; 7.5, 9.5]$ and $f_6 = [3; 9.5, 15]$. We have one 1-segment $f_3 = [1; 5, 11]$ and one 2-segment $f_5 = [2; 5, 11]$. The local and global weights are as follow:

f_i	1	2	3	4	5	6
$\omega_\ell(f_i)$	$\frac{11}{4}$	$\frac{13}{2}$	$\frac{11}{2}$	$\frac{17}{2}$	$\frac{19}{2}$	$\frac{49}{4}$
$\omega_g(f_i)$	$\frac{11}{4}$	$\frac{13}{2}$	$\frac{15}{2}$	$\frac{17}{2}$	$\frac{15}{2}$	$\frac{49}{4}$

In the example the segments f_3 and f_5 have the same global weight, but f_3 will appear before f_5 in the sorted sequences of F. The normal form $\mathrm{nf}(x)$ will correspond to $f_1 f_2 f_3 f_4 f_5 f_6$, whereas $\mathrm{nf}(x_1)$ is $f_1 f_3 f_2$ and $\mathrm{nf}(x_2)$ is $f_4 f_5 f_6$. We also have $\mathrm{alph}(f_2 f_4) \times \mathrm{alph}(f_3 f_5) \subseteq I$.

4. The existential theory of trace monoids with involution

For a moment let $(M, \overline{\ })$ be any finitely generated monoid with involution generated by Γ, and let $\psi : \Gamma^* \to M$ be a surjective homomorphism (which respects the involution). Let Ω be a set of variables (or unknowns) together with an involution without fixed points $\overline{\ } : \Omega \to \Omega$. Let $\mathcal{C}$ be a family of subsets of M which we call *constraints*.

The *existential theory of equations with* $\mathcal{C}$*-constraints* in the monoid M is defined as follows. Atomic formulae are either of the form $\alpha = \beta$, where $\alpha, \beta \in (\Gamma \cup \Omega)^*$ or of the form $X \in C$, where X is in Ω and $C \subseteq M$ belongs to the family $\mathcal{C}$. An *existentially quantified* formula is a block of existentially quantified variables followed by a Boolean combination of atomic formulae. It is *closed*, if there are no free variables. A closed formula is also called a *sentence*. If the interpretation of a variable X is $m \in M$, then the interpretation of $\overline{X}$ must be $\overline{m} \in M$. The existential theory of equations with $\mathcal{C}$-constraints in M is the set of all existentially quantified sentences which are *true* in M.

A subset $L \subseteq M$ is called *recognizable*, if $\psi^{-1}(L)$ is a regular word language in the usual sense of automata theory. Languages of the form $\psi(L)$ are called *rational*, if $L \subseteq \Gamma^*$ is regular. They can also be defined by regular (or rational) expressions. Kleene's Theorem states that in Γ^* the classes of rational and recognizable languages coincide, so we call them regular for simplicity. However, in general, the class of rational languages is strictly larger than the one of recognizable languages. In particular, if $\mathbb{M}$ is a trace monoid with $(a, b) \in I$, then $(ab)^* \subseteq \mathbb{M}$ is rational, but not recognizable since $\psi^{-1}((ab)^*)$ is the set of words with an equal number of a and b. If G is an infinite group, then the singleton $\{1\}$ is rational, but not recognizable because a subgroup is recognizable if and only if it is of finite index.

In this section we use recognizable trace languages as constraints. As before, ψ means the canonical homomorphism from words to traces, thus $\psi : \Gamma^* \to \mathbb{M}$ and $\mathbb{M} = \mathbb{M}(\Gamma, I)$ is a trace monoid with involution. If not stated otherwise we assume that a recognizable trace language is specified by some I*-diamond* NFA $\mathcal{A}$, where NFA means non-deterministic finite automaton and the I-diamond property means that for all states p, q of $\mathcal{A}$ and all pairs $(a, b) \in I$ there is a path from p to q labeled by ab if and only if there is a path from p to q labeled by ba. Such an NFA recognizes a a regular language $K \subseteq \Gamma^*$ with $\psi^{-1}(\psi(K)) = K$, hence it defines a recognizable trace language $\psi(K) \subseteq \mathbb{M}$. In Section 6 we show that we can start with constraints that are presented in lexicographical normal form, which is sometimes a more compact specification of recognizable constraints.

Definition 19. *Let ETMI denote the following decision problem:*

INPUT: A graph $(\Gamma, I, \overline{})$ *and an existentially quantified sentence with recognizable constraints in a trace monoid with involution* $\mathbb{M} = M(\Gamma, I)$.

QUESTION: Is the sentence true *in* $\mathbb{M}$*?*

The proof of the following statement is the main contribution of the paper.

Theorem 20. *The following assertions hold.*

(i) The problem ETMI is PSPACE*–hard.*

(ii) There exist a constant k and a polynomial $q(x)$ such that the problem ETMI can be solved in space $\tau^{k\tau} \cdot q(n)$, where n denotes the length of the input and τ is the number of thin clans, $\tau = \tau(\Gamma, D)$. In particular, it can be solved in EXPSPACE.

Corollary 21. *If the input to the problem ETMI is restricted such that the parameter τ is bounded by some constant, then the problem is* PSPACE*–complete.*

The PSPACE–hardness follows directly from a result of Kozen [16], since the *empty intersection* problem of regular sets is a special instance of the problem ETMI restricted to inputs with $\tau = 0$. So we do not discuss PSPACE–hardness in the proofs of Theorem 20 and Corollary 21 anymore.

Moreover, for inputs with $\tau = 0$ the assertions of Theorem 20 and Corollary 21 follow by [6], since $\tau = 0$ is exactly the case of free monoids with involution. The road map is therefore as follows. We assume that the input satisfies $\tau > 1$. (The case $\tau = 1$ does not exist.) Then we reduce it non-deterministically to the case at most $\tau - 1$ in such a way that the input size n is increased at most by factor $p(\tau)$, where $p(x)$ is some fixed polynomial. The degree of p is actually quadratic, only. So for some (small) constant k' we reach the situation $\tau = 0$ with an input of size at most $\tau^{k'\tau} \cdot n$. Then we apply the polynomially space bounded algorithm of [6] in order to get a non-deterministic algorithm using space at most $\tau^{k\tau} \cdot q(n)$.

The reduction is done stepwise. Let Φ be the input to the problem ETMI.

Step 1: Using De Morgan's laws we may assume that there are no negations at all and that the atomic formulae of Φ are of either form: $\alpha = \beta$, $\alpha \neq \beta$, $X \in L$, or $X \notin L$, where L is a recognizable trace language.

A formula $\min(X) \neq \min(Y)$ stands for

$$\bigvee_{a \in \Gamma} \big((X \in a\mathbb{M} \ \wedge \ Y \notin a\mathbb{M}) \vee (X \notin a\mathbb{M} \ \wedge \ Y \in a\mathbb{M})\big).$$

With the help of this macro we can replace an inequality $\alpha \neq \beta$ by the equivalent formula $\exists Z \exists X \exists Y : \alpha = ZX \ \wedge \ \beta = ZY \ \wedge \ \min(X) \neq \min(Y)$.

However, there is a general, more space efficient strategy we are following here: we can avoid disjunctions by making non-deterministic guesses. In particular, instead of writing the macro $\min(X) \neq \min(Y)$ we guess the corresponding letter a and and one side in the disjunction $(X \in a\mathbb{M} \ \wedge \ Y \notin a\mathbb{M}) \vee (X \notin a\mathbb{M} \ \wedge \ Y \in a\mathbb{M})$.

After the first phase we may therefore assume that Φ is an existentially quantified sentence over a conjunction of atomic formulae of either form: $\alpha = \beta$, $X \in L$, or $X \notin L$.

Step 2: We reduce to the case that the involution is without fixed points. Assume that the set of fixed points $\Delta = \{a \in \Gamma \mid a = \overline{a}\}$ is not empty. (Otherwise we skip this step.)

Let Δ' be a disjoint copy of Δ, and let $\Gamma' = \Gamma \cup \Delta'$. For each $a \in \Delta$ we define $\overline{a} = a'$ and $\overline{a'} = a$. It is clear how to extend I to I' such that I' is compatible with the involution, and $\mathbb{M}'$ becomes a trace monoid where the involution on Γ' has no fixed points. We define a homomorphism $\iota : \mathbb{M} \to \mathbb{M}'$ by $\iota(a) = aa'$ for $a \in \Delta$ and $\iota(a) = a$ for $a \in \Gamma \setminus \Delta$. There is also a projection $\pi : \mathbb{M}' \to \mathbb{M}$, defined by erasing all $a' \in \Delta'$. Obviously, $\pi\iota(x) = x$, hence $\iota(x) = \iota(y)$ in $\mathbb{M}'$ implies $x = y$ in $\mathbb{M}$, hence ι is injective. Since $aa' = \overline{(aa')}$ we have $\iota(\overline{x}) = \overline{\iota(x)}$ for all $x \in \mathbb{M}$ and

$\pi(\overline{x}) = \overline{\pi(x)}$ for all $x \in \iota(\mathbb{M})$. Moreover, if $L \subseteq \mathbb{M}$ is recognizable, then $\iota(L) \subseteq \mathbb{M}'$ is also recognizable. Indeed, $\iota(L) = \pi^{-1}(L) \cap \iota(\mathbb{M})$. For the specification of $\pi^{-1}(L)$ we may use the same automaton as for L by adding self-loops labeled by letters from Δ'. This automaton is again I-diamond.

It remains to see that $\iota(\mathbb{M})$ is recognizable. In order to avoid a state explosion we do not use a single automaton, but we write $\iota(\mathbb{M})$ as an intersection of at most $|\Delta|$ recognizable sets: For each $a \in \Delta$ we have an automaton M_a consisting of two states, p_a, q_a, with p_a being initial and final. There is a transition from p_a to q_a labeled by a, and one from q_a to p_a labeled by a'. Moreover, p_a has self-loops labeled by c for every $c \in \Gamma' \setminus \{a, a'\}$, and q_a has self-loops labeled by d, for every d with $(a, d) \in I$. Then $\iota(\mathbb{M}) = \bigcap_{a\in\Delta} L(M_a)$.

We transform the sentence Φ as follows: Each subformula of the form $\exists X \varphi$ is replaced by $\exists X (\bigwedge_{a\in\Delta} X \in L(M_a) \ \wedge \ \varphi)$, each constant $a \in \Delta$ appearing in an equation is replaced by aa' and finally each constraint $X \in L$ is replaced by $\bigwedge_{a\in\Delta} X \in L(M_a) \wedge X \in \pi^{-1}(L)$.

We obtain a sentence Φ' over $\mathbb{M}'$ which has the same truth value as the sentence Φ after Step 1. The length of Φ' is at most polynomial in the length of Φ, but the switch from Γ to Γ' did not increase the number of thin clans. So, since this step is not repeated, we may in fact we assume that $\Gamma = \Gamma'$. The sentence Φ' is denoted by Φ again.

After Step 2 we are in the following situation: Φ is an existentially quantified sentence of a conjunction over equations and recognizable constraints, and the involution $\overline{} : \Gamma \to \Gamma$ has no fixed points, i.e., $a \neq \overline{a}$ for all $a \in \Gamma$. Moreover, we may assume that all equations of Φ are in triangulated form, $X = X_1X_2$ with $X, X_1, X_2 \in \Gamma \cup \Omega$.

Step 3: This step is repeated until $\tau = 0$.

We choose some thin clan in (Γ, D). As in the previous section we write the clan in the form $A \cup \overline{A}$ with $A \cap \overline{A} = \emptyset$. We define $\widehat{D} = D \cup \Gamma \times (A \cup \overline{A}) \cup (A \cup \overline{A}) \times \Gamma$ and $\widehat{I} = \Gamma \times \Gamma \setminus \widehat{D}$. Recall that $\widehat{\mathbb{M}} = M(\Gamma, \widehat{I})$ is a trace monoid with involution, where the number of thin clans in $(\Gamma, \widehat{D})$ is at most $\tau - 1$.

Let $\hat{\psi}$ mean the canonical homomorphism $\hat{\psi} : \widehat{\mathbb{M}} \to \mathbb{M}$. We transform non-deterministically Φ into some sentence $\widehat{\Phi}$ over $\widehat{\mathbb{M}}$. The handling of recognizable constraints is trivial: A constraint $X \in L$ ($X \notin L$ resp.) is replaced by $X \in \hat{\psi}^{-1}(L)$ ($X \notin \hat{\psi}^{-1}(L)$ resp.). The same I-diamond automaton as for L serves also for $\hat{\psi}^{-1}(L)$ since it is $\widehat{I}$-diamond.

For a replacement of an equation $X = X_1X_2$ we make several non-deterministic guesses. We guess some $d \leq 3\tau^2+8\tau+7$ and we choose d new, existentially quantified variables $Y_1, \ldots, Y_d$. We guess a permutation $\pi \in \mathrm{Perm}(d)$ and some c with $0 \leq c \leq d$. Now, $X = X_1X_2$ is replaced by the following system of equations and constraints.

$$\begin{aligned} X &= Y_{\pi(1)} \ldots Y_{\pi(d)}, \\ X_1 &= Y_1 \ldots Y_c, \end{aligned}$$

$$X_2 = Y_{c+1} \dots Y_d,$$
$$Y_i \times Y_j \subseteq I \text{ for all } i, j, \text{ where } (i-j)(\pi(i) - \pi(j)) < 0.$$

Here, a formula $X \times Y \subseteq I$ stands for

$$\bigwedge_{a \in \Gamma} (X \in (\Gamma \setminus \{a\})^* \ \vee \ Y \in I(a)^*).$$

A formula as above is called a *commutation constraint.* If $I = \emptyset$, then the macro means nothing but $X \in \{1\} \ \vee \ Y \in \{1\}$.

In fact, we make again guesses, so instead of writing a commutation constraint we actually write down a conjunction of purely alphabetic constraints of type $Y \in B^*$ with $B \subseteq \Gamma$. The system of equations can be triangulated, so that we end up with a form where we can repeat Step 3. The correctness of Step 3 is a direct consequence of Theorem 8. It remains to analyze the space requirement after all loops of Step 3. At the end all constraints are purely alphabetic or of the form $X \in \psi^{-1}(L)$ ($X \notin \psi^{-1}(L)$ resp.) where L is a constraint in the form after Step 2. The number of new alphabetic constraints is bounded by the number of new variables since a conjunction of alphabetic constraints can be written as a single alphabetic constraint.

The system of equations is triangulated, so it is enough to calculate the number of equations. After triangulation each equation is replaced by $2d$ new equations, so the upper bound is the polynomial $p(\tau) = 6\tau^2 + 16\tau + 14$. Clearly for some constant k we have $\prod_{i=2}^{\tau} p(i) \le \tau^{k\tau}$. This shows Theorem 20 and Corollary 21.

The result of Theorem 20 cannot be extended to rational constraints due to the following fact which we state for sake of completeness.

Theorem 22. *[22,7] The existential theory of equations with rational constraints in $\mathbb{M}$ is decidable if and only if $\mathbb{M}$ is a free product of free commutative monoids.*

5. Equations over graph groups

5.1. *Elementary properties of graph groups*

Graph groups (or *free partially commutative groups*) arise at many places in mathematics and they are well-studied objects under various names, see e.g. [2,4]. The most standard approach is to define a graph group as the quotient group of a free group by a partial commutation relation between generators. This is the usual setting of graph groups as investigated by Droms in [12]. Our definition of a graph group is slightly more general, since we allow that the involution has fixed points and this leads to torsion elements of order 2.

We start with an alphabet with involution $(\Gamma, \overline{})$ and, as always, we assume that the independence relation $I \subseteq \Gamma \times \Gamma$ is compatible with the involution. We define the *graph group* $G(\Gamma, I)$ by

$$M(\Gamma, I)/\{a\overline{a} = 1 \mid a \in \Gamma\}.$$

This is a group, since $\overline{\overline{a}} = a$. The canonical homomorphism is denoted by $\varphi : M(\Gamma, I) \to G(\Gamma, I)$. If the reference to (Γ, I) is clear, we write $\mathbb{G}$ instead of $G(\Gamma, I)$. A trace $x \in \mathbb{M}$ is called *reduced*, if it does not contain any factor of the form $a\overline{a}$ with $a \in \Gamma$. It is well-known and easy to verify that every group element $x \in \mathbb{G}$ has a unique reduced representation. This means there is a unique reduced trace $\widetilde{x} \in \mathbb{M}$ such that $\varphi(\widetilde{x}) = x$.

We have the following basic lemma.

Lemma 23. *Let $\widetilde{x}, \widetilde{y}, \widetilde{z} \in \mathbb{M}$ be reduced traces representing the group elements $x, y, z \in \mathbb{G}$. Then we have $xy = z$ in $\mathbb{G}$ if and only if there are traces $p, q, r \in \mathbb{M}$ such that $\widetilde{x} = pq, \widetilde{y} = \overline{q}r$, and $\widetilde{z} = pr$.*

Proof. Let $xy = z$ in $\mathbb{G}$ and let $q \in \mathbb{M}$ be of maximal length such that we can write $\widetilde{x} = pq$ and $\widetilde{y} = \overline{q}r$. Then we have $z = pr$ in $\mathbb{G}$. However, p and r are reduced. Thus, if a factor $a\overline{a}$ occurs in pr, then a is a maximal letter in p and $\overline{a}$ is a minimal letter in r. This is true, because $(a, b) \in I$ implies $(\overline{a}, b) \in I$. Since q is of maximal length, the trace pr is reduced. Hence $\widetilde{z} = pr$. The other direction is trivial. $\square$

5.2. *Normalized regular subsets*

The set of all reduced traces is in one-to-one correspondence with $\mathbb{G}$, and we have a normal form mapping $\rho : \mathbb{G} \to \mathbb{M}$, defined by $\rho(x) = \widetilde{x}$.

Since $\rho(\mathbb{G})$ is defined by some finite set of forbidden factors $a\overline{a}$ with $a \in \Gamma$, it is a recognizable subset of $\mathbb{M}$. A group language $L \subseteq \mathbb{G}$ is called *normalized regular* (or *normalized rational* in [8]), if the set of normal forms $\rho(L) \subseteq \mathbb{M}$ is a recognizable trace language. In particular, if $A \subseteq \Gamma$ is a subset, then $A^* = \{x \in \mathbb{G} \mid \mathrm{alph}(\widetilde{x}) \subseteq A\} \subseteq \mathbb{G}$ is normalized regular, because $A^* \cap \rho(\mathbb{G}) \subseteq \mathbb{M}$ is recognizable. Since $\rho(\mathbb{G} \setminus L) = \rho(\mathbb{G}) \setminus \rho(L)$, the class of normalized regular languages is an effective Boolean algebra.

If not stated otherwise, a normalized regular language L is given by an I-diamond NFA accepting $\rho(L)$.

If $\mathbb{G}$ is a free group, then every rational language is normalized regular. This follows from [3]. In general, the class of normalized regular languages is strictly contained in the class of rational subsets, since $(ab)^* \subseteq \mathbb{G}$ is not normalized regular, if $(a, b) \in I$, $a \neq \overline{a}$, and $b \neq \overline{b}$. On the other hand, all finite subsets are normalized regular, hence if $\mathbb{G}$ is infinite, then the class of normalized regular languages is strictly larger than the class of recognizable subsets. Thus, normalized regular subsets form a class between recognizable and rational languages, and they are in one-to-one correspondence with recognizable trace languages containing only reduced traces. Finally note that if $L \subseteq \mathbb{M}$ is recognizable then $\varphi(L)$ need not be normalized regular, in general. Take for example three letters a, b, c with $(a, b) \in I$, $(a, c) \in D$ and $(b, c) \in D$. The trace language $L = (ac\overline{c}b)^*$ is recognizable. However, $\varphi(L) = (ab)^* \subseteq \mathbb{G}$ is not normalized regular, if $a \neq \overline{a}$, and $b \neq \overline{b}$.

5.3. *The existential theory of graph groups*

Definition 24. *Let ETGG denote the following decision problem:*

INPUT: A graph $(\Gamma, I, \overline{})$ *and an existentially quantified sentence with normalized regular constraints in a graph group* $\mathbb{G} = G(\Gamma, I)$.

QUESTION: Is the sentence true *in* $\mathbb{G}$*?*

Theorem 25. *The following assertions hold.*

(i) *The problem ETGG is* PSPACE*–hard.*
(ii) *There exist a constant* k *and a polynomial* $q(x)$ *such that the problem ETGG can be solved in space* $\tau^{k\tau} \cdot q(n)$*, where* n *denotes the length of the input and* τ *is the number of thin clans,* $\tau = \tau(\Gamma, I)$*. In particular, it can be solved in* EXPSPACE*.*

Corollary 26. *If the input to the problem ETGG is restricted such that the parameter* τ *is bounded by some constant, then the problem becomes* PSPACE*–complete.*

The lower bounds (PSPACE–hardness) in Theorem 25 and Corollary 26 follow exactly the same way as in Theorem 20. For the upper bounds it is enough to show the following proposition.

Proposition 27. *There is a polynomial time reduction of Problem ETGG to Problem ETMI, which does not change the underlying graph* (Γ, I).

Proof. The input to Problem ETGG is an existentially quantified sentence and we ask whether it is true in the graph group $\mathbb{G}$. We may assume that negations appear only with atomic formulae. A formula of type $\alpha \neq \beta$ is equivalent with $\exists X : \alpha X = \beta \wedge X \notin \{1\}$ (note that $\{1\}$ is a normalized regular constraint). Hence we may assume that there are negations only with normalized regular constraints.

As above, we may assume that all equations are of the form $xy = z$ with $x, y, z \in \Gamma \cup \Omega$.

We transform the formula as follows. Each constraint $X \in L$ (resp. $X \notin L$) is replaced by $X \in \rho(L)$ (resp. $X \in \rho(\mathbb{G}) \ \wedge X \notin \rho(L)$). Note that the constraint $X \in \rho(\mathbb{G})$ can be expressed as the conjunction of $|\Gamma|$ many constraints which forbid the factors $a\overline{a}$, $a \in \Gamma$. By Lemma 23 we can replace each equation $xy = z$ by

$$\exists P, Q, R : x = PR \ \wedge \ y = \overline{R}Q \ \wedge \ z = PR.$$

We do not need any constraints for P, Q or R. $\square$

6. Constraints in lexicographical normal form

Two words representing the same trace have the same length. Fixing some linear order on Γ we can choose for a given trace $x \in \mathbb{M}$ the lexicographical first word representing x. This word is called the *lexicographical normal form* of x and it is denoted by $\mathrm{lex}(x) \in \Gamma^*$. Clearly $\psi(\mathrm{lex}(x)) = x$. It is well-known [1] that $\mathrm{lex}(\mathbb{M}) \subseteq \Gamma^*$

is a regular word language. Hence, if $L \subseteq \mathbb{M}$ is recognizable, then $\psi^{-1}(L) \cap \mathrm{lex}(\mathbb{M})$ is regular, too; and L is a homomorphic image of some regular subset of $\mathrm{lex}(\mathbb{M})$. Ochmański's Theorem says that this correspondence is one-to-one between recognizable subsets of $\mathbb{M}$ and regular subsets of $\mathrm{lex}(\mathbb{M})$, see [23], [24, Thm. 6.3.12]. Hence, for the specification of a constraint (i.e., a recognizable trace language L) we may use some NFA, where the accepted language is a subset of $\mathrm{lex}(\mathbb{M})$. Vice versa, if the accepted language of an NFA is a subset of $\mathrm{lex}(\mathbb{M})$, then it specifies a recognizable trace language. Note that the structure of an NFA accepting lexicographical normal forms is quite different from an I-diamond automaton. If we start with an I-diamond automaton we can use a simple product automaton construction for $\psi^{-1}(L) \cap \mathrm{lex}(\mathbb{M})$. However, the size of a minimal NFA recognizing $\mathrm{lex}(\mathbb{M})$ can be exponential in the size of the alphabet.

Example 28. Let Γ be an alphabet which contains $2n+1$ letters with the following ordering:

$$a_1 < \cdots < a_n < b_1 \cdots < b_n < c.$$

Let the dependency among these letters be induced by $(c, a_i), (a_i, b_i)$ for all $1 \leq i \leq n$. Then every NFA recognizing $\mathrm{lex}(\mathbb{M})$ has at least 2^n states. Indeed, for a subset $J \subseteq \{1, \ldots, n\}$ let A_J (resp. B_J) be the lexicographical normal form of the product in over all a_j (resp. b_j) with $j \in J$. Consider an NFA recognizing $\mathrm{lex}(\mathbb{M})$ and let p_J be the state on an accepting path for the word cA_JB_J after reading the prefix cA_J. If the automaton has less than 2^n states, then $p_J = p_K$ for some $J \neq K$. We may assume $K \setminus J \neq \emptyset$, but then the automaton accepts the word cA_JB_K which is not in lexicographical normal form, hence a contradiction.

The reverse operation, starting with an NFA and constructing an I-diamond automaton is more complicated, as we will see below.

In the following we assume that the total order chosen for Γ is compatible with the clans of (Γ, D). More precisely, we choose $<$ by ordering the clans and then choosing some total order on each clan. With such an ordering it follows from Proposition 29 that we can start with an existentially quantified sentence with constraints in lexicographical normal form, and, still, the sentence can be evaluated in ExpSpace.

Proposition 29. *Let (Γ, I) be an independence alphabet with τ thin clans. Let $\mathcal{A}$ be an NFA with s states such that $L(\mathcal{A}) \subseteq lex(\mathbb{M})$. Then we can build an I-diamond NFA $\mathcal{A}'$ with $L(\mathcal{A}') = \psi^{-1}(\psi(L(\mathcal{A})))$ and where the number of states is at most $\tau^{2\tau} s^{2\tau^2+1}$.*

The proof of the previous proposition is based on a lemma that describes how some linearization of a trace prefix y of $x \in \mathbb{M}$ fits into the lexicographical normal form $\mathrm{lex}(x)$ of x.

Lemma 30. *Let $x \in \mathbb{M}$ and let y be a prefix of x. Decompose $lex(x) = v_0u_0 \cdots v_ku_k$ such that the positions of y in $lex(x)$ correspond to the subword $v_0 \cdots v_k$ and k is* *minimal with this property. In particular,* $v_i \neq 1$ *for all* $0 < i \leq k$, $u_i \neq 1$ *for all* $0 \leq i < k$, *and* $(u_i, v_j) \in I$ *for all* $i < j$. *Then we have* $k \leq \tau^2 - \tau$.

Proof. We may assume that $\tau > 0$ and $k > 0$. Let $A_i = I(v_i \cdots v_k)$ for $0 < i \leq k$. Obviously, $(A_i)_i$ is a monotonically increasing sequence, $A_i \subseteq A_{i+1}$ for all i. Moreover, $A_1 \neq \emptyset$ (since $u_0 \neq 1$) and $A_k \subsetneq \Gamma$. If Γ contains a thick clan, then it is necessarily contained in $\Gamma \setminus A_k$ We split the sequence $(A_i)_i$ into:

$$A_1 = \cdots = A_{i_1} \subsetneq A_{i_1+1} = \cdots = A_{i_2} \subsetneq \cdots \subsetneq A_{i_{m-1}+1} = \cdots = A_{i_m} = A_k.$$

We note that $m \leq \tau$. This is clear, because for $1 \leq i \leq i_m$ each A_i is a non-empty union of thin clans. Next, assume that we have $A_i = A_{i+1}$ for some $1 \leq i < k$. Let b be the first letter of v_i and d be the first letter of v_{i+1}. Note that neither b nor d belong to any thick clan because $A_1 \neq \emptyset$. The only way the first position in u_i can be filled with some letter c is that we have $b < c < d$ and $c \in A_i$. If b and d belong to the same thin clan then there is no such c due to the choice of our ordering. Hence, if $A_i = A_{i+1} = \cdots = A_j$, then $j - i + 1 < \tau$, because the first letters of $v_i, \ldots, v_j$ yield an ordered sequence of thin clans, all of them disjoint with A_1. Hence the claim of the lemma. □

The proof of Proposition 29 follows now by defining the states of $\mathcal{A}'$ as tuples $(p_0, A_1, q_1, \ldots, p_{k-1}, A_k, q_k, p_k)$, where $A_i \subseteq \Gamma$ $(1 \leq i \leq k)$, p_i $(0 \leq i \leq k)$ and q_i $(1 \leq i \leq k)$ are states of $\mathcal{A}$ and $k \leq \tau^2 - \tau$ as in Lemma 30. The idea is that $\mathcal{A}'$ guesses an accepting run on $\mathrm{lex}(x)$, with A_i the alphabets defined in the proof of that lemma, and p_i, q_i the states reached in the run after reading $v_0 u_0 \cdots v_i$ and $v_0 u_0 \cdots v_{i-1} u_{i-1}$, respectively. Another way to see this is to replace in the sequence $v_0 u_0 \cdots v_k u_k$ each v_i by q_i, p_i and each u_i by A_i. We do not need to remember the initial state q_0. Moreover, it is enough to remember for each thin clan A the least i such that $A \subseteq A_i$. Hence the number of states can be bounded by $\tau^{2\tau} s^{2\tau^2+1}$. The automaton $\mathcal{A}'$ is I-diamond.

Corollary 31. *The problems ETMI and ETGG are in* ExpSpace *(resp.* PSpace*, if* τ *is bounded by some constant), even if the constraints for the input are given in lexicographical normal form.*

Remark 32. The reader may ask why we did not use lexicographical normal forms already in Section 3. The reason is simple: if Γ contains letters a, b with $(a, b) \in I$, then the lexicographical normal form is not compatible with the involution. Indeed, let $(a, b) \in I$ and $a < b$, then we must have $\overline{b} < \overline{a}$ or we are not compatible with the involution. If $b < \overline{a}$, then $b\overline{a}a$ is in lexicographical normal form since $(a, \overline{a}) \in D$, but $\overline{a}a\overline{b}$ is not in lexicographical normal form. If $\overline{a} < b$, then $\overline{a}b\overline{b}$ is in lexicographical normal form, but $b\overline{b}a$ is not in lexicographical normal form.

Acknowledgments

We are grateful to Yuri Matiyasevich for various contributions which were at the origin of this work. We also thank Marcus Schaefer who suggested better complexity bounds with respect to the conference version of the present paper. Finally we thank an anonymous referee for the detailed report which helped to improve the presentation.

References

[1] A. V. Anisimov and D. E. Knuth. Inhomogeneous sorting. *International Journal of Computer and Information Sciences*, 8:255–260, 1979.

[2] A. Baudisch. Subgroups of semifree groups. *Acta Mathematica Academiae Scientiarum Hungaricae*, 38:19–28, 1981.

[3] M. Benois. Parties rationelles du groupe libre. *C. R. Acad. Sci. Paris, Sér. A*, 269:1188–1190, 1969.

[4] N. Brady and J. Meier. Connectivity at infinity for right angled Artin groups. *Transactions of the American Mathematical Society*, 353:117–132, 2001.

[5] P. Cartier and D. Foata. *Problèmes combinatoires de commutation et réarrangements.* Number 85 in Lecture Notes in Mathematics. Springer-Verlag, Berlin Heidelberg, 1969.

[6] V. Diekert, C. Gutiérrez, and C. Hagenah. The existential theory of equations with rational constraints in free groups is PSPACE-complete. *Information and Computation*, 202:105–140, 2005. The conference version appeared at STACS'01, LNCS 2010, pages 170 – 182, 2001.

[7] V. Diekert and M. Lohrey. Word equations over graph products. In P. K. Pandya and J. Radhakrishnan, editors, *Proceedings of the 23rd Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2003), Mumbai (India)*, number 2914 in Lecture Notes in Computer Science, pages 156–167. Springer-Verlag, 2003.

[8] V. Diekert and M. Lohrey. Existential and positive theories of equations in graph products. *Theory of Computing Systems*, 37:133–156, 2004.

[9] V. Diekert, Yu. Matiyasevich, and A. Muscholl. Solving word equations modulo partial commutations. *Theoretical Computer Science*, 224:215–235, 1999. Special issue of LFCS'97.

[10] V. Diekert and A. Muscholl. Solvability of equations in free partially commutative groups is decidable. In F. Orejas, P. G. Spirakis, and J. van Leeuwen, editors, *Proc. 28th International Colloquium on Automata, Languages and Programming (ICALP'01)*, number 2076 in Lecture Notes in Computer Science, pages 543–554, Berlin Heidelberg, 2001. Springer-Verlag.

[11] V. Diekert and G. Rozenberg, editors. *The Book of Traces.* World Scientific, Singapore, 1995.

[12] C. Droms. Isomorphisms of graph groups. *Proc. American Mathematical Society*, 100:407–408, 1987.

[13] C. Gutiérrez. Satisfiability of equations in free groups is in PSPACE. In *Proceedings 32nd Annual ACM Symposium on Theory of Computing, STOC'2000*, pages 21–27. ACM Press, 2000.

[14] R. M. Keller. Parallel program schemata and maximal parallelism I. Fundamental results. *Journal of the Association for Computing Machinery*, 20(3):514–537, 1973.

[15] A. Kościelski and L. Pacholski. Makanin's algorithm is not primitive recursive. *The-*

oretical Computer Science, 191:145–156, 1998.

[16] D. Kozen. Lower bounds for natural proof systems. In *Proc. of the 18th Ann. Symp. on Foundations of Computer Science, FOCS'77*, pages 254–266, Providence, Rhode Island, 1977. IEEE Computer Society Press.

[17] G. S. Makanin. The problem of solvability of equations in a free semigroup. *Math. Sbornik*, 103:147–236, 1977. English transl. in Math. USSR Sbornik 32 (1977).

[18] G. S. Makanin. Equations in a free group. *Izv. Akad. Nauk SSR*, Ser. Math. 46:1199–1273, 1982. English transl. in Math. USSR Izv. 21 (1983).

[19] Yu. Matiyasevich. Some decision problems for traces. In S. Adian and A. Nerode, editors, *Proceedings of the 4th International Symposium on Logical Foundations of Computer Science (LFCS'97), Yaroslavl, Russia, July 6–12, 1997*, number 1234 in Lecture Notes in Computer Science, pages 248–257, Berlin Heidelberg, 1997. Springer-Verlag. Invited lecture.

[20] A. Mazurkiewicz. Concurrent program schemes and their interpretations. DAIMI Rep. PB 78, Aarhus University, Aarhus, 1977.

[21] A. Mazurkiewicz. Trace theory. In W. Brauer et al., editors, *Petri Nets, Applications and Relationship to other Models of Concurrency*, number 255 in Lecture Notes in Computer Science, pages 279–324, Berlin Heidelberg, 1987. Springer-Verlag.

[22] A. Muscholl. *Decision and complexity issues on concurrent systems.* Habilitationsschrift (postdoctoral thesis), Universität Stuttgart, 1999.

[23] E. Ochmański. Regular behaviour of concurrent systems. *Bulletin of the European Association for Theoretical Computer Science (EATCS)*, 27:56–67, Oct. 1985.

[24] E. Ochmański. Recognizable trace languages. In V. Diekert and G. Rozenberg, editors, *The Book of Traces*, chapter 6, pages 167–204. World Scientific, Singapore, 1995.

[25] W. Plandowski. Satisfiability of word equations with constants is in PSPACE. *Journal of the Association for Computing Machinery*, 51:483–496, 2004. Conference version appeared at FOCS'99, IEEE Computer Society Press, pages 495-500, 1999.

[26] M. Schaefer, E. Sedgwick, and D. Štefankovič. Recognizing string graphs in NP. *Journal of Computer and System Sciences*, 2003:365–380, 2003. Conference version appeared at STOC'02, ACM Press, pages 1–6, 2002.

[27] K. U. Schulz. Makanin's algorithm for word equations — Two improvements and a generalization. In K. U. Schulz, editor, *Word Equations and Related Topics*, number 572 in Lecture Notes in Computer Science, pages 85–150, Berlin Heidelberg, 1991. Springer-Verlag.