

Existential and Positive Theories of Equations in Graph Products

Volker Diekert, Markus Lohrey

Universität Stuttgart, Institut für Informatik
Breitwiesenstr. 20–22, 70565 Stuttgart, Germany
{diekert,lohrey}@informatik.uni-stuttgart.de

Abstract. We prove that the existential theory of equations with normalized rational constraints in a fixed graph product of finite monoids, free monoids, and free groups is PSPACE-complete. Under certain restrictions this result also holds if the graph product is part of the input. As the second main result we prove that the positive theory of equations with recognizable constraints in graph products of finite and free groups is decidable.

1 Introduction

Since the seminal work of Makanin [19] on equations in free monoids, the decidability of various theories of equations in different monoids and groups has been studied, and several new decidability and complexity results have been shown. Let us mention here the results of [25, 27] for free monoids, [6, 15, 20, 21] for free groups, [9] for free partially commutative monoids (trace monoids), [10] for free partially commutative groups (graph groups), and [7] for plain groups (free products of finite and free groups).

In this paper we continue this stream of research. We will present two main results. The first one concerns existential theories of equations. We start with the definition of a class of monoids, which are constructed from finite monoids, free monoids, and free groups using the graph product construction, which is a well-known construction in mathematics. This class of graph products strictly covers all classes mentioned above. Then we prove that for such a graph product the existential theory of equations is PSPACE-complete, where in addition we are allowed to specify constraints for the variables. These constraints are taken from a class of sets, called normalized rational sets, which (in general) lies strictly between the class of recognizable and rational sets. Furthermore under certain restrictions our PSPACE upper-bound holds also in the case that (a suitable description) of the graph product is part of the input.

Our second main result concerns positive theories of equations. We prove that if we restrict our class of graph products to groups, then for each group from the resulting class the positive theory of equations with recognizable constraints for the variables is decidable. Under certain restrictions we obtain an elementary complexity. Up to now only for the class of free groups a decidability result for

the positive theory was known, in particular it was open whether the positive theory of equations for a free partially commutative group is decidable. The full paper of this extended abstract can be found in [8].

2 Preliminaries

An *involution* on a set is a mapping $\bar{}$ such that $\overline{\overline{x}} = x$ for all elements x . For an involution on a monoid we demand in addition that both $\overline{xy} = \overline{y} \overline{x}$ and $\overline{1} = 1$, where 1 is the neutral element of the monoid. Taking the inverse in a group is for instance an involution. In our setting we let Γ be a finite alphabet of constants and $\Delta \subseteq \Gamma$ such that an involution $\bar{}$ is defined on Δ . This involution is extended to Δ^* by $\overline{x_1 \cdots x_n} = \overline{x_n} \cdots \overline{x_1}$. For a monoid M we denote by $\mathcal{I}(M)$ a submonoid of M such that an involution $\bar{}$ is defined on $\mathcal{I}(M)$. In many cases we choose $\mathcal{I}(M)$ to be the submonoid of elements having left- and right-inverses, i.e., $\mathcal{I}(M)$ is the group of units of M , but this is not necessarily the case, for instance for $M = \Gamma^*$ we take $\mathcal{I}(M) = \Delta^*$. We consider only finitely generated monoids. More precisely, we consider monoids M together with a fixed surjective homomorphism $\psi : \Gamma^* \rightarrow M$ such that $\psi^{-1}(\mathcal{I}(M)) = \Delta^*$ and $\psi(\overline{x}) = \overline{\psi(x)}$ for all $x \in \Delta^*$. Moreover, we assume that there is a *normal form mapping* $\nu : M \rightarrow \Gamma^*$, i.e., $\psi(\nu(x)) = x$ for all $x \in M$, such that $\nu(M)$ is a regular subset of Γ^* . Note that it is allowed that $\nu(\overline{x}) \neq \overline{\nu(x)}$ for some $x \in M$. A language $L \subseteq M$ is called

- *recognizable* if $\psi^{-1}(L) \subseteq \Gamma^*$ is regular,
- *normalized rational* if $\nu(L) \subseteq \Gamma^*$ is regular,
- *rational* if $L = \psi(L')$ for some regular language $L' \subseteq \Gamma^*$.

The corresponding classes are denoted by $\text{REC}(M)$, $\text{NRAT}(M)$, and $\text{RAT}(M)$, respectively. We have $\text{REC}(M) \subseteq \text{NRAT}(M) \subseteq \text{RAT}(M)$. The classes $\text{REC}(M)$ and $\text{RAT}(M)$ are classical, see e.g. [4], their definitions do neither depend on ν nor on ψ as can be seen easily. The definition of $\text{NRAT}(M)$ is less robust, it depends on the normal form mapping ν . The classes $\text{REC}(M)$ and $\text{NRAT}(M)$ are Boolean algebras, whereas $\text{RAT}(M)$ is not a Boolean algebra in general. For free monoids we have $\text{REC}(M) = \text{NRAT}(M) = \text{RAT}(M)$. For the canonical normal form mappings which we will use we have: $\text{REC}(M) \neq \text{NRAT}(M) = \text{RAT}(M)$ for free groups [3], $\text{REC}(M) = \text{NRAT}(M) \neq \text{RAT}(M)$ for free partially commutative monoids (trace monoids) [24], and $\text{REC}(M) \neq \text{NRAT}(M) \neq \text{RAT}(M)$ for free partially commutative groups (graph groups). The later holds for instance in $M = \mathbb{Z} \times \mathbb{Z}$.

3 The theory of equations with constraints

Let M be a monoid as above and let \mathcal{C} be a family of subsets of M such that $\mathcal{I}(M) \in \mathcal{C}$. Let Ω be a set of variables and $\overline{\Omega} = \{\overline{X} \mid X \in \Omega\}$ a disjoint copy of Ω . An *equation* is a pair (U, V) with $U, V \in (\Gamma \cup \Omega \cup \overline{\Omega})^*$, it is written as $U = V$. Equations and *constraints* of the form $X \in L$ with $X \in \Omega \cup \overline{\Omega}$ and $L \in \mathcal{C}$

are called *atomic formulae*. From these we construct first order formulae using conjunctions, disjunctions, negations, and universal and existential quantification over variables from Ω . We impose the syntactical restriction that whenever we use a variable $\overline{X} \in \overline{\Omega}$, then this goes together with the implicit constraint $X \in \mathcal{I}(M)$. Given $\psi : \Gamma^* \rightarrow M$, $\mathcal{I}(M)$, the involution $\bar{\cdot} : \mathcal{I}(M) \rightarrow \mathcal{I}(M)$, and a sentence ϕ , i.e., a formula in the sense above without free variables, we can evaluate ϕ over M in the obvious way with the restriction that if a variable X evaluates to $x \in M$, then \overline{X} must evaluate to \overline{x} . The *theory of equations with constraints in \mathcal{C}* , briefly $\text{Th}(M, \mathcal{C})$, denotes the set of all sentences that are true in M . A well-known example of a decidable theory of equations is the Presburger Arithmetic [26]. Translated into our framework this gives the following proposition.

Proposition 1. $\text{Th}(\mathbb{N}^k, \text{RAT}(\mathbb{N}^k))$ and $\text{Th}(\mathbb{Z}^k, \text{RAT}(\mathbb{Z}^k))$ are decidable.

Note that $\text{RAT}(\mathbb{N}^k)$ and $\text{RAT}(\mathbb{Z}^k)$ are the classes of semilinear sets in \mathbb{N}^k and \mathbb{Z}^k , respectively. The following result can be easily deduced from Proposition 1 since the free product $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$ of two copies of $\mathbb{Z}/2\mathbb{Z}$ is isomorphic to the semi-direct product of \mathbb{Z} by $\mathbb{Z}/2\mathbb{Z}$.

Corollary 1. $\text{Th}(\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}, \text{RAT}(\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}))$ is decidable.

The *positive theory of equations with constraints in \mathcal{C}* is the set of all sentences in $\text{Th}(M, \mathcal{C})$ that do not use negations. The *existential theory of equations with constraints in \mathcal{C}* is the set of all sentences in $\text{Th}(M, \mathcal{C})$ that are in prenex normal form without universal quantifiers. We will need the following result, which is a decomposition lemma in the style of the Feferman Vaught Theorem. Its proof in [8] is due to Yuri Matiyasevich (personal communication).

Proposition 2. Let M_1 and M_2 be monoids with classes $\mathcal{C}_1 \subseteq 2^{M_1}$ and $\mathcal{C}_2 \subseteq 2^{M_2}$. Let \mathcal{C} be a class of subsets of $M_1 \times M_2$ such that each $L \in \mathcal{C}$ is a finite union of sets of the form $L_1 \times L_2$ with $L_1 \in \mathcal{C}_1$ and $L_2 \in \mathcal{C}_2$. If both $\text{Th}(M_1, \mathcal{C}_1)$ and $\text{Th}(M_2, \mathcal{C}_2)$ are decidable, then $\text{Th}(M_1 \times M_2, \mathcal{C})$ is decidable, too. The same implication also holds for positive theories.

4 Graph products

Let (V, E) be a finite undirected graph with vertex set V and edge set $E \subseteq \binom{V}{2}$. Every node $n \in V$ is labeled with a monoid M_n which is either a free monoid, a free group, or a finite monoid. In fact, it is enough (and convenient) to assume that M_n is either isomorphic to \mathbb{N} or to \mathbb{Z} , or M_n is finite. If $M_n = \mathbb{N}$, then we let $\Gamma_n = \{a_n\}$ and $\Delta_n = \emptyset$. If $M_n = \mathbb{Z}$, then we let $\Gamma_n = \Delta_n = \{a_n, \overline{a}_n\}$. Finally if M_n is finite, then we let $\Gamma_n = M_n \setminus \{1\}$ and $\Delta_n = \mathcal{I}(M_n) \setminus \{1\}$, where $\mathcal{I}(M_n)$ is the subgroup of units of M_n , i.e., $\mathcal{I}(M_n) = \{a \in M_n \mid \exists b : ab = ba = 1\}$. Thus, for each $n \in V$ we have a canonical homomorphism $\psi_n : \Gamma_n^* \rightarrow M_n$ with $\psi_n^{-1}(\mathcal{I}(M_n)) = \Delta_n^*$. To see this note that if $uv \in \mathcal{I}(M_n)$ and if M_n is finite, then $u, v \in \mathcal{I}(M_n)$, too. The *graph product* defined by (V, E) is the free product of the

monoids M_n , $n \in V$, modulo commutation relations $xy = yx$ for all $x \in M_m$, $y \in M_n$ with $(m, n) \notin E$. Graph products of arbitrary groups and monoids were investigated in [5, 14]. Note that we have defined a commutation, if there is no edge, so an edge corresponds to a rigid ordering. The choice for this convention is due to the representation of elements which is best based on dependence graphs, see e.g. [11]. Before we make our definition more formal let us mention some examples.

If all M_n are equal to \mathbb{N} , then we obtain *free partially commutative monoids*, which are also known as *trace monoids*, see [11] for more details. Extreme cases are free monoids (if $E = \binom{V}{2}$) and free commutative monoids (if $E = \emptyset$). If all M_n are equal to \mathbb{Z} , we obtain *free partially commutative groups*, which are also known as *graph groups* [12]. Again free groups and free commutative groups arise as the extreme cases. If $E = \binom{V}{2}$ and all M_n are groups, then we obtain *plain groups* in the sense of Haring-Smith [16].

Let us proceed with an explicit definition of the graph product using generators and relations. First we may assume that all the alphabets Γ_n are pairwise disjoint. Let $\Gamma = \bigcup_{n \in V} \Gamma_n$ and $\Delta = \bigcup_{n \in V} \Delta_n$. There is a natural involution $\bar{}$ on Δ and this involution has fixed points as soon as some M_n contains an element of order two. We define an *independence relation* $I \subseteq \Gamma \times \Gamma$ by $I = \{(a, b) \in \Gamma \times \Gamma \mid a \in \Gamma_m, b \in \Gamma_n, m \neq n, (m, n) \notin E\}$, which is irreflexive and symmetric. The basic reference monoid for the following consideration is the trace monoid $\mathbb{M} = \Gamma^* / \{ab = ba \mid (a, b) \in I\}$, it is equipped with a partially defined involution. More precisely, since I is compatible with the involution in the sense that $(a, \bar{b}) \in I$ if $(a, b) \in I$ and $b \in \Delta$, we can lift $\bar{} : \Delta \rightarrow \Delta$ to an involution on the recognizable subset $\Delta^* = \mathcal{I}(\mathbb{M})$ of \mathbb{M} . We now define a *trace rewriting system* S , i.e., a subset of $\mathbb{M} \times \mathbb{M}$, by

$$S = \{(a\bar{a}, 1) \mid a \in \Delta\} \cup \{(ab, c) \mid \exists n \in V : a, b, c \in \Gamma_n, ab = c \text{ in } M_n\}.$$

The graph product \mathbb{GP} of the monoids M_n , $n \in V$, over the graph (V, E) is defined as the quotient monoid $\mathbb{GP} = \mathbb{M} / \{\ell = r \mid (\ell, r) \in S\}$. Clearly $\mathbb{GP} = \Gamma^* / (\{ab = ba \mid (a, b) \in I\} \cup \{\ell = r \mid (\ell, r) \in S\})$. Elements of \mathbb{GP} can be represented as words from Γ^* or as traces from \mathbb{M} . It will be always clear from the context, which representation is chosen. Furthermore the canonical homomorphism $\psi : \Gamma^* \rightarrow \mathbb{GP}$ factorizes as $\psi = \psi_1 \circ \psi_2$, where $\psi_1 : \Gamma^* \rightarrow \mathbb{M}$ and $\psi_2 : \mathbb{M} \rightarrow \mathbb{GP}$. Note that the trace monoid \mathbb{M} itself is a graph product, where the vertex set is Γ and the edges are given by the complement of I . The example of a trace monoid shows that rational constraints are too strong in order to obtain decidability results. Since it is undecidable whether $L_1 \cap L_2 = \emptyset$ for $L_1, L_2 \in \text{RAT}(\mathbb{N} \times \{a, b\}^*)$, see [1], the following result holds:

Proposition 3. *Let $\mathbb{M} = \mathbb{N} \times \{a, b\}^*$. Then for M the existential positive theory of equations with constraints in $\text{RAT}(M)$ is undecidable.*

Thus, we have to restrict the class of constraints. We shall consider normalized rational constraints. In order to define a suitable normal form mapping $\nu : \mathbb{GP} \rightarrow \Gamma^*$ we define analogously to string rewriting systems the one-step rewrite

relation $\rightarrow_S \subseteq \mathbb{M} \times \mathbb{M}$ of the trace rewriting system S by $s \rightarrow_S t$ if $s = u \ell v$ and $t = u r v$ for some $(\ell, r) \in S$ and $u, v \in \mathbb{M}$. Its transitive reflexive closure is $\overset{*}{\rightarrow}_S$. The following lemma is fundamental for the following.

Lemma 1. *S is a confluent trace rewriting system, i.e., for all $s, t, u \in \mathbb{M}$ with $s \overset{*}{\rightarrow}_S t$ and $s \overset{*}{\rightarrow}_S u$ there exists $v \in \mathbb{M}$ with $t \overset{*}{\rightarrow}_S v$ and $u \overset{*}{\rightarrow}_S v$.*

Let $\text{RED}(S) = \{u \ell v \mid u, v \in \mathbb{M}, \exists r : (\ell, r) \in S\}$ and $\text{IRR}(S) = \mathbb{M} \setminus \text{RED}(S)$. Thus, $\text{IRR}(S)$ is the set of traces that are irreducible with respect to S . Since $\text{REC}(\mathbb{M})$ is closed under Boolean operations and concatenation, see e.g. [11, Chap. 6], $\text{IRR}(S)$ is recognizable. Since \rightarrow_S is a Noetherian relation, Lemma 1 implies that for each $x \in \mathbb{GP}$ there exists a unique $\mu(x) \in \mathbb{M} \cap \text{IRR}(S)$ with $x = \psi_2(\mu(x))$. The trace $\mu(x)$ is the shortest trace representing x . Now let us fix a linear order on Γ and let $\text{lnf}(t) \in \Gamma^*$ for $t \in \mathbb{M}$ be the lexicographical first word from Γ^* that represents the trace t , see also [2]. Then for $x \in \mathbb{GP}$ we define $\nu(x) = \text{lnf}(\mu(x))$. Since $L \in \text{REC}(\mathbb{M})$ if and only if $\text{lnf}(L) \subseteq \Gamma^*$ is regular [24], we obtain:

Lemma 2. *We have $L \in \text{NRAT}(\mathbb{GP})$ if and only if $\mu(L) \in \text{REC}(\mathbb{M})$ if and only if $\psi_1^{-1}(\mu(L)) \in \text{REC}(\Gamma^*)$.*

In particular we see that $\text{NRAT}(\mathbb{GP})$ does not depend on the chosen lexicographical ordering. It is really a canonical class depending only on the natural trace rewriting system S .

5 Existential theories of equations in graph products

In this section we prove that for the graph product \mathbb{GP} the existential theory of equations with constraints in $\text{NRAT}(\mathbb{GP})$ is decidable. Since we will also deal with complexity issues, we have to define the input length of a formula. We assume some standard binary coding of formulae, where a constraint $X \in L$ is represented by some finite non-deterministic automaton that accepts $\psi_1^{-1}(\mu(L))$. The input length of a formula is the length of this description. In order to obtain existing results for free monoids as special cases, we will put a description of the graph product \mathbb{GP} into the input, too. This description contains the adjacency matrix of (V, E) , and for each node either the multiplication table of M_n if M_n is finite or a flag indicating whether $M_n = \mathbb{N}$ or $M_n = \mathbb{Z}$. In order to obtain convenient complexity bounds we will restrict to graphs (V, E) with a bounded number of *complete thin clans*, see [10] for the definition. It is easy to see that the number of complete thin clans of (V, E) is at most $|V|$, furthermore it is 0 for a complete graph.

Theorem 1. *The following problem is PSPACE-complete for every $k \geq 0$.*

INPUT: A graph product \mathbb{GP} whose underlying graph (V, E) has at most k complete thin clans and an existential formula ϕ with constraints in $\text{NRAT}(\mathbb{GP})$.

QUESTION: Does ϕ belong to $\text{Th}(\mathbb{GP}, \text{NRAT}(\mathbb{GP}))$?

If the number of complete thin clans of (V, E) is not bounded, then the problem above is in EXPSpace.

Remark 1. Formally, Theorem 1 generalizes results of [6, 7, 9, 10, 15, 19, 20, 25]. For this it is enough to give a reduction to the main result of [10].

The next lemma is the main technical tool for proving the theorem above. First we need some further definitions concerning traces. The set $\text{IC} \subseteq \mathbb{M} \cap \text{IRR}(S)$ consists of all traces $a_1 \cdots a_n$, $a_i \in \Gamma$, such that $(a_i, a_j) \in I$ if $i \neq j$. Thus, traces in IC correspond to independence cliques of (Γ, I) . Note that if $u \in \text{IC}$, then the length of u is at most $|\Gamma|$. We identify $u \in \text{IC}$ with the set of symbols that occur in u . For instance for $s \in \mathbb{M}$ the set of maximal symbols $\text{max}(s) = \{a \in \Gamma \mid s = ta\}$ of s and the set of minimal symbols $\text{min}(s) = \{a \in \Gamma \mid s = at\}$ of s belong to IC .

Lemma 3. *Let $x, y, z \in \mathbb{M} \cap \text{IRR}(S)$. Then $xy \xrightarrow{*}_S z$ if and only if there exist $p, s, t, w \in \text{IRR}(S)$ and $u, v \in \text{IC}$ such that*

$$uv \xrightarrow{*}_S w, \quad x = sup, \quad y = \bar{p}vt, \quad z = swt. \quad (1)$$

Note that since $u, v \in \text{IC}$, there exist only finitely many possibilities for w in (1).

Proof (Theorem 1). PSPACE-hardness follows from the fact that for $\{a, b\}^*$ the existential theory of equations with constraints in $\text{REC}(\{a, b\}^*)$ is PSPACE-hard, see [17, Lem. 3.2.3] and [25, Thm. 1]. Membership in PSPACE will be shown by a reduction to the following problem, which was shown to be in PSPACE for every $k \geq 0$ in [10]:

INPUT: A trace monoid \mathbb{M} , specified by an independence relation $I \subseteq \Gamma \times \Gamma$ such that the graph $(\Gamma, (\Gamma \times \Gamma) \setminus I)$ has at most k complete thin clans, a completely defined involution $\bar{\cdot} : \Gamma \rightarrow \Gamma$ that is compatible with I (i.e. $(a, \bar{b}) \in I$ if $(a, b) \in I$), and an existential formula ϕ with constraints in $\text{REC}(\mathbb{M})$.

QUESTION: Is ϕ true in \mathbb{M} with the lifting $\bar{\cdot} : \mathbb{M} \rightarrow \mathbb{M}$ of $\bar{\cdot} : \Gamma \rightarrow \Gamma$? In this problem a set $L \in \text{REC}(\mathbb{M})$ is specified via an automaton for $\psi_1^{-1}(L)$.

Now let k be a fixed bound for the number of complete thin clans, and let \mathbb{GP} be a graph product, specified by a graph (V, E) with at most k complete thin clans. Furthermore let ϕ be an existential formula with constraints in $\text{NRAT}(\mathbb{GP})$. Using standard methods, see e.g. [6], we may assume that ϕ is an existentially quantified conjunction of equations of the form $xy = z$, where $x, y, z \in \Gamma \cup \Omega \cup \bar{\Omega}$, and of constraints $X \in L$ or $X \notin L$, where $X \in \Omega \cup \bar{\Omega}$ and $L \in \text{NRAT}(\mathbb{GP})$. Next we will move from the graph product \mathbb{GP} to its underlying trace monoid \mathbb{M} (it is easy to see that the number of complete thin clans of $(\Gamma, (\Gamma \times \Gamma) \setminus I)$ is also at most k). We replace syntactically every subformula $xy = z$ (resp. $X \in L$) by $\psi_2(xy) = \psi_2(z)$ (resp. $X \in \mu(L)$) and add the negated constraint $X \notin \text{RED}(S)$ for every variable X .¹ We obtain an existential formula which evaluates to *true* in \mathbb{M} if and only if the original formula evaluates to *true* in \mathbb{GP} . Note also that the automaton used to specify $\mu(L)$ is the same as the

¹ Of course this constraint is equivalent to $X \in \text{IRR}(S)$, but we prefer the negated constraint $X \notin \text{RED}(S)$ since an automaton for $\psi_1^{-1}(\text{RED}(S))$ can be easily constructed in polynomial time, whereas the construction of an automaton for $\psi_1^{-1}(\text{IRR}(S))$ would involve an additional complementation with a possible exponential blow-up.

one for L . It remains to eliminate all occurrences of ψ_2 from equations. Since $\Gamma \subseteq \text{IRR}(S)$ and S is confluent, we can replace an equation $\psi_2(xy) = \psi_2(z)$ by $xy \xrightarrow{*}_S z$, which by Lemma 3 is equivalent to an existentially quantified conjunction of equations.

Now we can almost apply the result of [10] cited above. The only remaining problem is that due to the presence of non-invertible generators in \mathbb{GP} , the involution $\bar{}$ may only be partially defined on Γ . But this can be resolved by introducing a new dummy symbol \bar{a} for every $a \in \Gamma \setminus \Delta$ and by adding the constraint $X \in \Gamma^*$ for every variable X . This shows the first statement from Theorem 1.

For the case that the number of complete thin clans is not bounded, an EXPSPACE-algorithm can be deduced from the proof in [10]. \square

6 Positive theories of equations in graph products

In this section we prove our second main result. In the following we throughout assume that all generators in Γ have inverses, i.e. $\Gamma = \Delta$. In particular \mathbb{GP} is a graph product of finite and free groups, and hence itself a group.

Theorem 2. *The following problem is decidable.*

INPUT: A graph product \mathbb{GP} which is a group and a closed positive formula ϕ with constraints in $\text{REC}(\mathbb{GP})$.

QUESTION: Does ϕ belong to $\text{Th}(\mathbb{GP}, \text{REC}(\mathbb{GP}))$?

Complexity issues will be postponed to the end of this section. Note that Theorem 2 cannot be extended to the full class of graph products considered in the previous section. Already for a free monoid $\{a, b\}^*$ the $\forall\exists^3$ -theory of equations is undecidable [13, 22]. Similarly Theorem 2 cannot be extended to the case of normalized rational constraint, since for a free group F of rank 2 a free submonoid $\{a, b\}^*$ belongs to $\text{NRAT}(F)$.

We will prove Theorem 2 by reducing the positive theory of equations with constraints in $\text{REC}(\mathbb{GP})$ to the existential theory of equations with normalized rational constraints in a free extension of \mathbb{GP} , which allows us to apply Theorem 1. Our proof strategy will follow a technique developed in [21, 23] but the presence of partial commutation and recognizable constraints makes the construction more involved.

In a first step we may assume that none of the finite groups M_n , $n \in V$, is a direct product of two finite non-trivial groups since otherwise we could replace n by two non-connected nodes. In particular, if M_n is not $\mathbb{Z}/2\mathbb{Z}$, then there must exist an $a \in \Gamma_n$ such that $a \neq \bar{a}$ in \mathbb{GP} . Next assume that the graph (V, E) consists of two non-empty disjoint components (V_1, E_1) and (V_2, E_2) , which define graph products \mathbb{GP}_1 and \mathbb{GP}_2 , respectively. Then $\mathbb{GP} = \mathbb{GP}_1 \times \mathbb{GP}_2$. Furthermore by Mezei's Theorem, see e.g. [4], every $L \in \text{REC}(\mathbb{GP})$ is a finite union of sets of the form $L_1 \times L_2$ with $L_i \in \text{REC}(\mathbb{GP}_i)$. Thus, we may apply Proposition 2 and proceed with the two graphs (V_1, E_1) and (V_2, E_2) . Hence, for the rest of the proof we may assume that the graph (V, E) is connected. Furthermore since

by Proposition 1 the (positive) theory of equations with rational constraints in \mathbb{Z} is decidable and the same holds for finite monoids for trivial reasons, we may assume that $|V| > 1$. By Corollary 1 we can also exclude the case that V contains exactly two adjacent nodes which are both labeled by $\mathbb{Z}/2\mathbb{Z}$. Thus, we may assume that either the graph (V, E) contains a path consisting of three different nodes or one of the groups labeling the nodes has a generator $x \in \Gamma$ with $\bar{x} \neq x$. Hence, there exist three generators $a, b, c, \in \Gamma$ such that a and b belong to E -adjacent (and hence different) nodes from V , b and c also belong to E -adjacent nodes from V , and finally either a and c belong to different nodes from V or $a \neq \bar{a} = c$. In particular $(a, b), (b, c) \notin I$, i.e., the dependency between a, b , and c being used is $a - b - c$. For the rest of the proof we will fix these three symbols a, b , and c .

Since $L \in \text{REC}(\mathbb{GP})$ if and only if there exists a homomorphism $\rho : \mathbb{GP} \rightarrow H$ onto a finite group H such that $L = \rho^{-1}(\rho(L))$, see e.g. [4], we may fix for the further consideration such a homomorphism ρ and assume that all recognizable constraints are given in the form $\rho(X) = g$ for $X \in \Omega \cup \bar{\Omega}$ and $g \in H$.

We proceed with the definition of a trace rewriting system $R_N^{(h)}$, where $N \subseteq \mathbb{N}$ and $h \in H$. This trace rewriting system will be defined over some free extension of \mathbb{M} . First we need some preliminaries. A *chain* is a trace $a_1 \cdots a_m \in \mathbb{M}$, where $a_1, \dots, a_m \in \Gamma$, and a_i and a_{i+1} belong to E -adjacent (and hence different) nodes from V , $1 \leq i \leq m - 1$. Note that a chain belongs to $\text{IRR}(S)$.

Lemma 4. *For all $h \in H$ there exists a trace $C_h \in \mathbb{M} \cap \text{IRR}(S)$ such that $\min(C_h) = \max(C_h) = c$ and $\rho(C_h) = h$.*

Let C be a chain with $\min(C) = \max(C) = c$ and $|C| > |C_h|$ for all $h \in H$ such that for every node $n \in V$ at least one symbol from Γ_n occurs in C . Since (V, E) is connected, such a C exists. Choose an η with $|b(ab)^\eta| > |C| + 2$, and let $p = b(ab)^\eta C (ba)^\eta b$ and $\ell_i(h) = (ab)^{i \cdot |H|} C_h (ba)^{2 \cdot i \cdot |H|}$ for $i \geq 1$ and $h \in H$. Note that $p \ell_i(h) p \in \text{IRR}(S)$ and $\rho(\ell_i(h)) = h$. For every $i \in \mathbb{N}$ let us take two new constants $k_i, \bar{k}_i \notin \Gamma$ and set $\bar{\bar{k}}_i = k_i$. For every $N \subseteq \mathbb{N}$ and every $h \in H$ we define over the trace monoid $\mathbb{M} * \{k_i, \bar{k}_i \mid i \in N\}^*$, i.e., the free product of our trace monoid \mathbb{M} and the free monoid $\{k_i, \bar{k}_i \mid i \in N\}^*$, the trace rewriting system $R_N^{(h)}$ by $R_N^{(h)} = \{(p \ell_i(h) p, p k_i p), (\bar{p} \bar{\ell}_i(h) \bar{p}, \bar{p} \bar{k}_i \bar{p}) \mid i \in N\}$. Note that $R_N^{(h)}$ is length-reducing and thus, $\rightarrow_{R_N^{(h)}}$ is Noetherian. Let us fix $h \in H$ for the rest of this section. We write R_N and ℓ_i instead of $R_N^{(h)}$ and $\ell_i(h)$, respectively. We write $s \rightarrow_i t$ if the trace t can be obtained from the trace s by an application of one of the rules $(p \ell_i p, p k_i p)$ or $(\bar{p} \bar{\ell}_i \bar{p}, \bar{p} \bar{k}_i \bar{p})$. The next two lemmas are the fundamental statements about the trace rewriting system R_N and the reason for the complicated definition of the traces p and $\ell_i(h)$.

Lemma 5. *Let $i, j \in N \subseteq \mathbb{N}$ and $s, t, u \in \mathbb{M} * \{k_i, \bar{k}_i \mid i \in N\}^*$ such that $s \rightarrow_i t$ and $s \rightarrow_j u$. Then either $t = u$ or there exists a trace $v \in \mathbb{M} * \{k_i, \bar{k}_i \mid i \in N\}^*$ such that $t \rightarrow_j v$ and $u \rightarrow_i v$.*

In particular, R_N is confluent. Since R_N is also Noetherian, for every $s \in \mathbb{M}$ there exists a unique trace $\kappa_N(s) \in \mathbb{M} * \{k_i, \bar{k}_i \mid i \in N\} \cap \text{IRR}(R_N)$ with $s \xrightarrow{*}_{R_N} \kappa_N(s)$.

Lemma 6. For all $s, t \in \mathbb{M}$ there exists an $A \subseteq N$ with $|A| \leq 2$ such that for every $N' \subseteq N \setminus A$ it holds $\kappa_{N'}(st) = \kappa_{N'}(s)\kappa_{N'}(t)$.

6.1 Reduction to the existential theory

In the following, symbols with a tilde like \tilde{x} will denote sequences of arbitrary length over some set, which will be always clear from the context. If say $\tilde{x} = x_1 \cdots x_i$, then $\tilde{x} \in A$ means $x_1 \in A, \dots, x_i \in A$ and $f(\tilde{x})$ for some function f denotes the sequence $f(x_1) \cdots f(x_i)$.

For the rest of the paper let us take some subset $K = \{k_1, \dots, k_n\}$ of our new constants and let $\overline{K} = \{\overline{k}_1, \dots, \overline{k}_n\}$. Let $k, \overline{k} \notin \Gamma \cup K \cup \overline{K}$ be two additional constants, as usual let $\overline{\overline{k}} = k$. The following lemma will be the key for reducing the positive theory to the existential theory, it allows the elimination of one universal quantifier. In this lemma we have to deal with formulae ϕ that are interpreted over the free product $\mathbb{GP} * F(K)$ of the graph product \mathbb{GP} and the free group $F(K)$ generated (as a group) by K . Furthermore different recognizable constraints in ϕ are given by different extensions $\varrho : \mathbb{GP} * F(K) \rightarrow H$ of our fixed morphism $\rho : \mathbb{GP} \rightarrow H$. For $h \in H$ we denote by ϕ_h the formula that results from ϕ by replacing every constraint $\varrho(X) = g$ by $\varrho_h(X) = g$, where ϱ_h is the canonical extension of $\varrho : \mathbb{GP} * F(K) \rightarrow H$ to $\mathbb{GP} * F(K \cup \{k\})$ which is defined by $\varrho_h(k) = h$. Note that $\psi_2 : \mathbb{M} \rightarrow \mathbb{GP}$ can be extended to a canonical morphism from $\mathbb{M} * (K \cup \overline{K})^*$ to $\mathbb{GP} * F(K)$, which will be also denoted by ψ_2 .

Lemma 7. Let $\phi(X, Y_1, \dots, Y_m, \tilde{z})$ be a positive Boolean formula with constraints of the form $\varrho(Y) = g$ for (possibly different) extensions $\varrho : \mathbb{GP} * F(K) \rightarrow H$ of $\rho : \mathbb{GP} \rightarrow H$. Let $K_i \subseteq K$ for $1 \leq i \leq m$. Then for all $\tilde{z} \in \mathbb{GP}$ we have

$$\forall X \in \mathbb{GP} \exists Y_1, \dots, Y_m \left\{ \begin{array}{l} \phi(X, Y_1, \dots, Y_m, \tilde{z}) \wedge \\ \bigwedge_{i=1}^m Y_i \in \mathbb{GP} * F(K_i) \end{array} \right\} \quad \text{in } \mathbb{GP} * F(K) \quad (2)$$

if and only if

$$\bigwedge_{h \in H} \exists Y_1, \dots, Y_m \left\{ \begin{array}{l} \phi_h(k, Y_1, \dots, Y_m, \tilde{z}) \wedge \\ \bigwedge_{i=1}^m Y_i \in \mathbb{GP} * F(K_i \cup \{k\}) \end{array} \right\} \quad \text{in } \mathbb{GP} * F(K \cup \{k\}). \quad (3)$$

Proof. First assume that (3) holds for $\tilde{z} \in \mathbb{GP}$. In order to prove (2), let us choose an arbitrary $s \in \mathbb{GP}$ and let $h = \rho(s)$. Then there exist $t_i \in \mathbb{GP} * F(K_i \cup \{k\})$, $1 \leq i \leq m$, such that $\phi_h(k, t_1, \dots, t_m, \tilde{z})$ holds in $\mathbb{GP} * F(K \cup \{k\})$. Let us define a homomorphism $\sigma : \mathbb{GP} * F(K \cup \{k\}) \rightarrow \mathbb{GP} * F(K)$ by $\sigma(k) = s$ and $\sigma(x) = x$ for $x \in \mathbb{GP} * F(K)$. Since $\rho(s) = h$ and ϕ_h is positive, the sentence $\phi(s, \sigma(t_1), \dots, \sigma(t_m), \tilde{z})$ holds in $\mathbb{GP} * F(K)$ (note that $\sigma(\tilde{z}) = \tilde{z}$). Thus, (2) holds.

For the other direction assume that (2) holds for $\tilde{z} \in \mathbb{GP}$. Define a trace rewriting system T over $\mathbb{M} * (K \cup \overline{K})^*$ by $T = S \cup \{x\overline{x} \rightarrow 1, \overline{x}x \rightarrow 1 \mid x \in K\}$.

Completely analogously to the proof of Theorem 1 we can now change into the trace monoid $\mathbb{M} * (K \cup \overline{K})^*$. We obtain a sentence of the form

$$\forall X \in \text{IRR}(S) \exists Y_1, \dots, Y_m, \tilde{Y} \in \text{IRR}(T) \left\{ \begin{array}{l} \varphi(X, Y_1, \dots, Y_m, \tilde{Y}, \tilde{u}) \wedge \\ \bigwedge_{i=1}^m Y_i \in \mathbb{M} * (K_i \cup \overline{K_i})^* \end{array} \right\} \quad (4)$$

which evaluates to true in $\mathbb{M} * (K \cup \overline{K})^*$. Here $\tilde{u} = \mu(\tilde{z}) \in \text{IRR}(S)$, and the positive Boolean formula φ results from the original positive Boolean formula ϕ by applications of Lemma 3 to equations $xy = z$. These transformations only introduce new existentially quantified variables, which correspond to \tilde{Y} in (4). The constraints in (4) are the same as in (2) (formally we identify a homomorphism $\varrho : \mathbb{G}\mathbb{P} * F(K) \rightarrow H$ with $\psi_2 \circ \varrho : \mathbb{M} * (K \cup \overline{K})^* \rightarrow H$). Let $\mathcal{M} \subseteq \mathbb{M}$ consist all traces in \tilde{u} plus Γ . W.l.o.g we assume that all equations in (4) have the form $xy = z$ for $x, y, z \in \Omega \cup \overline{\Omega} \cup \mathcal{M} \cup \overline{\mathcal{M}}$. Let λ be the maximum of n (the largest index of the constants in K) and the maximal length of the traces in \tilde{u} . Let d be the number of equations in (4). Fix an $h \in H$ in (3) and let $s \in \mathbb{M}$ be the trace

$$s = C_g p \ell_{\lambda+1}(h) p c p \ell_{\lambda+2}(h) p c \dots p \ell_{\lambda+2d+1}(h) p \in \text{IRR}(S), \quad (5)$$

where $g \in H$ is chosen such that $\rho(s) = h$. Then by (4) there exist traces $t_1, \dots, t_m, \tilde{t} \in \text{IRR}(T)$ with $t_i \in \mathbb{M} * (K_i \cup \overline{K_i})^*$ and

$$\varphi(s, t_1, \dots, t_m, \tilde{t}, \tilde{u}) \quad \text{in } \mathbb{M} * (K \cup \overline{K})^*. \quad (6)$$

Let $N = \{\lambda + 1, \dots, \lambda + 2d + 1\}$ and add to \mathcal{M} all traces from $\{s, t_1, \dots, t_m\}$. Then $\varphi(s, t_1, \dots, t_m, \tilde{t}, \tilde{u})$ is a true statement, which contains d atomic statements of the form $xy = z$ with $x, y, z \in \mathcal{M} \cup \overline{\mathcal{M}}$ plus recognizable constraints. Of course some of these atomic statements may be false. But since there are only d equations in (6), we have to remove from N by Lemma 6 at most $2d$ numbers such that for the resulting set N' we have $\kappa_{N'}(x)\kappa_{N'}(y) = \kappa_{N'}(z)$ ($x, y, z \in \mathcal{M} \cup \overline{\mathcal{M}}$) whenever $xy = z$ is a true atomic statement in (6). Since $|N| = 2d + 1$, we have $N' \neq \emptyset$, let $i \in N'$. Note that $k_i \notin K$ since $\lambda \geq n$. We rename the constant k_i into k and abbreviate $\kappa_{\{i\}}(x)$ by $\kappa(x)$. Again by Lemma 6 we have $\kappa(x)\kappa(y) = \kappa(z)$ for every true statement $xy = z$ ($x, y, z \in \mathcal{M} \cup \overline{\mathcal{M}}$) in (6). Furthermore if one of the constraints $\varrho(x) = g$ in (6) is true, where ϱ is an extension of ρ , then also $\varrho_h(\kappa(x)) = g$ holds (note that $\varrho(\ell_i(h)) = \rho(\ell_i(h)) = h = \varrho_h(k)$). Finally $\kappa(\tilde{u}) = \tilde{u}$ since λ was chosen big enough in (5). Altogether it follows that the statement $\varphi_h(\kappa(s), \kappa(t_1), \dots, \kappa(t_m), \kappa(\tilde{t}), \tilde{u})$ is true in $\mathbb{M} * (K \cup \overline{K} \cup \{k, \overline{k}\})^*$. Next we can write $\kappa(s) = s_1 k s_2$ for $s_1, s_2 \in \mathbb{M}$. Let us define a homomorphism $\sigma : \mathbb{M} * (K \cup \overline{K} \cup \{k, \overline{k}\})^* \rightarrow \mathbb{M} * (K \cup \overline{K} \cup \{k, \overline{k}\})^*$ by $\sigma(k) = \overline{s}_1 k \overline{s}_2$, $\sigma(\overline{k}) = s_2 \overline{k} s_1$, and $\sigma(x) = x$ otherwise. Note that $\rho(s_1)h\rho(s_2) = \rho(s) = h$ and hence $\varrho_h(\overline{s}_1 k \overline{s}_2) = \rho(s_1)^{-1}h\rho(s_2)^{-1} = h$ for every extension ϱ of ρ . Thus, the statement $\varphi_h(\sigma(\kappa(s)), \sigma(\kappa(t_1)), \dots, \sigma(\kappa(t_m)), \sigma(\kappa(\tilde{t})), \tilde{u})$ is true in $\mathbb{M} * (K \cup \overline{K} \cup \{k, \overline{k}\})^*$, hence it is also true in $\mathbb{G}\mathbb{P} * F(K \cup \{k\})$. But in this group $\sigma(\kappa(s)) = \sigma(s_1 k s_2) =$

$s_1 \bar{s}_1 k \bar{s}_2 s_2 = k$. Since furthermore $\sigma(\kappa(t_i)) \in \mathbb{M} * (K_i \cup \overline{K_i} \cup \{k, \bar{k}\})^*$, the sentence $\exists Y_1, \dots, Y_m, \tilde{Y} : \varphi_h(k, Y_1, \dots, Y_m, \tilde{Y}, \tilde{z}) \wedge \bigwedge_{i=1}^m Y_i \in \mathbb{GP} * F(K_i \cup \{k\})$ is true in $\mathbb{GP} * F(K \cup \{k\})$ for every $h \in H$. But then also (3) holds, since if (1) from Lemma 3 holds in $\mathbb{GP} * F(K \cup \{k\})$, then also $xy = z$ in $\mathbb{GP} * F(K \cup \{k\})$. \square

Let us fix a formula $\theta(\tilde{Z}) \equiv \forall X_1 \exists Y_1 \dots \forall X_n \exists Y_n \phi(X_1, \dots, X_n, Y_1, \dots, Y_n, \tilde{Z})$, where ϕ is a positive Boolean formula with constraints of the form $\rho(X) = g$. For $h_1, \dots, h_n \in H$ we denote by $\rho_{h_1, \dots, h_n} : \mathbb{GP} * F(K) \rightarrow H$ the canonical extension of ρ with $\rho_{h_1, \dots, h_n}(k_i) = h_i$ for $1 \leq i \leq n$. With ϕ_{h_1, \dots, h_n} we denote the formula, where every constraint $\rho(X) = g$ in ϕ is replaced by $\rho_{h_1, \dots, h_n}(X) = g$. The following theorem is the main result of this section, it can be easily deduced from Lemma 7 by an induction on n .

Theorem 3. *For all $\tilde{z} \in \mathbb{GP}$ we have $\theta(\tilde{z})$ in \mathbb{GP} if and only if*

$$\bigwedge_{h_1 \in H} \exists Y_1 \dots \bigwedge_{h_n \in H} \exists Y_n \left\{ \begin{array}{l} \phi_{h_1, \dots, h_n}(k_1, \dots, k_n, Y_1, \dots, Y_n, \tilde{z}) \\ \wedge \bigwedge_{i=1}^n Y_i \in \mathbb{GP} * F(\{k_1, \dots, k_i\}) \end{array} \right\} \text{ in } \mathbb{GP} * F(K).$$

Since $\mathbb{GP} * F(\{k_1, \dots, k_i\}) \in \text{NRAT}(\mathbb{GP} * F(K))$, Theorem 2 is a consequence of Theorem 1 and Theorem 3. Concerning the complexity, it can be shown that in general our proof of Theorem 2 gives us a non-elementary algorithm due to the construction in our proof of Proposition 2, see [8]. We obtain an elementary algorithm if we restrict to connected graphs (V, E) . For this we have to use the fact that Presburger arithmetic (without negations), which occurs for $\mathbb{GP} = \mathbb{Z}/2\mathbb{Z}$ or $\mathbb{GP} = \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/2\mathbb{Z}$ as a special case, is elementary. More precise complexity bounds will be given in the full version of this paper.

References

1. IJ. J. Aalbersberg and H. J. Hoogeboom. Characterizations of the decidability of some problems for regular trace languages. *Mathematical Systems Theory*, 22:1–19, 1989.
2. A. V. Anisimov and D. E. Knuth. Inhomogeneous sorting. *International Journal of Computer and Information Sciences*, 8:255–260, 1979.
3. M. Benois. Parties rationnelles du groupe libre. *C. R. Acad. Sci. Paris, Sér. A*, 269:1188–1190, 1969.
4. J. Berstel. *Transductions and context-free languages*. Teubner Studienbücher, Stuttgart, 1979.
5. A. V. da Costa. Graph products of monoids. *Semigroup Forum*, 63(2):247–277, 2001.
6. V. Diekert, C. Gutiérrez, and C. Hagenah. The existential theory of equations with rational constraints in free groups is PSPACE-complete. In *Proceedings of the 18th Annual Symposium on Theoretical Aspects of Computer Science (STACS 01)*, number 2010 in Lecture Notes in Computer Science, pages 170–182. Springer, 2001.

7. V. Diekert and M. Lohrey. A note on the existential theory of equations in plain groups. *International Journal of Algebra and Computation*, 2001. to appear.
8. V. Diekert and M. Lohrey. Existential and positive theories of equations in graph products. Technical Report 2001/10, University of Stuttgart, Germany, 2001.
9. V. Diekert, Y. Matiyasevich, and A. Muscholl. Solving word equations modulo partial commutations. *Theoretical Computer Science*, 224(1–2):215–235, 1999.
10. V. Diekert and A. Muscholl. Solvability of equations in free partially commutative groups is decidable. In *Proceedings of the 28th International Colloquium on Automata, Languages and Programming (ICALP 01)*, number 2076 in Lecture Notes in Computer Science, pages 543–554. Springer, 2001.
11. V. Diekert and G. Rozenberg, editors. *The Book of Traces*. World Scientific, 1995.
12. C. Droms. Graph groups, coherence and three-manifolds. *Journal of Algebra*, 106(2):484–489, 1985.
13. V. G. Durnev. Undecidability of the positive $\forall\exists^3$ -theory of a free semi-group. *Sibirsky Matematicheskie Jurnal*, 36(5):1067–1080, 1995.
14. E. R. Green. *Graph Products of Groups*. PhD thesis, The University of Leeds, 1990.
15. C. Gutiérrez. Satisfiability of equations in free groups is in PSPACE. In *32nd Annual ACM Symposium on Theory of Computing (STOC'2000)*, pages 21–27. ACM Press, 2000.
16. R. H. Haring-Smith. Groups and simple languages. *Transactions of the American Mathematical Society*, 279:337–356, 1983.
17. D. Kozen. Lower bounds for natural proof systems. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science, (FOCS 77)*, pages 254–266. IEEE Computer Society Press, 1977.
18. M. Lohrey. Confluence problems for trace rewriting systems. *Information and Computation*, 170:1–25, 2001.
19. G. S. Makanin. The problem of solvability of equations in a free semigroup. *Math. Sbornik*, 103:147–236, 1977. (Russian); English translation in *Math. USSR Sbornik* 32 (1977).
20. G. S. Makanin. Equations in a free group. *Izv. Akad. Nauk SSR, Ser. Math.* 46:1199–1273, 1983. (Russian); English translation in *Math. USSR Izv.* 21 (1983).
21. G. S. Makanin. Decidability of the universal and positive theories of a free group. *Izv. Akad. Nauk SSSR, Ser. Mat.* 48:735–749, 1984. (Russian); English translation in: *Math. USSR Izvestija*, 25, 75–88, 1985.
22. S. S. Marchenkov. Unsolvability of positive $\forall\exists$ -theory of a free semi-group. *Sibirsky Matematicheskie Jurnal*, 23(1):196–198, 1982.
23. Y. I. Merzlyakov. Positive formulas on free groups. *Algebra i Logika Sem.*, 5(4):25–42, 1966. (Russian).
24. E. Ochmański. Regular behaviour of concurrent systems. *Bulletin of the European Association for Theoretical Computer Science (EATCS)*, 27:56–67, 1985.
25. W. Plandowski. Satisfiability of word equations with constants is in PSPACE. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science (FOCS 99)*, pages 495–500. IEEE Computer Society Press, 1999.
26. M. Presburger. Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. In *Comptes Rendus du Premier Congrès des Mathématiciennes des Pays Slaves*, pages 92–101, 395, Warsaw, 1927.
27. K. U. Schulz. Makanin's algorithm for word equations — Two improvements and a generalization. In *Word Equations and Related Topics*, number 572 in Lecture Notes in Computer Science, pages 85–150. Springer, 1991.