Theories of HNN-extensions and amalgamated products

Markus Lohrey[1] and Géraud Sénizergues[2]

[1] Universität Stuttgart, FMI, Germany
[2] Université Bordeaux I, LaBRI, France
lohrey@informatik.uni-stuttgart.de, ges@labri.u-bordeaux.fr

Abstract. It is shown that the existential theory of $\mathbb{G}$ with rational constraints, over an HNN-extension $\mathbb{G} = \langle \mathbb{H}, t; t^{-1}at = \varphi(a)(a \in A) \rangle$ is decidable, provided that the same problem is decidable in the base group $\mathbb{H}$ and that A is a finite group. The positive theory of $\mathbb{G}$ is decidable, provided that the existential positive theory of $\mathbb{G}$ is decidable and that A and $\varphi(A)$ are proper subgroups of the base group $\mathbb{H}$ with $A \cap \varphi(A)$ finite. Analogous results are also shown for amalgamated products. As a corollary, the positive theory and the existential theory with rational constraints of any finitely generated virtually-free group is decidable.

1 Introduction

Theories of equations over groups are a classical research topic at the borderline between algebra, mathematical logic, and theoretical computer science. This line of research was initiated by the work of Lyndon, Tarski, and others in the first half of the 20th century. A major driving force for the development of this field was a question that was posed by Tarski around 1945: Is the first-order theory of a free group F of rank two, i.e, the set of all statements of first-order logic with equations as atomic propositions that are true in F, decidable. Decidability results for fragments of this theory were obtained by Makanin (for the existential theory of a free group) [15] and Merzlyakov and Makanin (for the positive theory of a free group) [16, 17]. A complete (positive) solution of Tarski's problem was finally announced in [9]; the complete solution is spread over a series of papers. The complexity of Makanin's algorithm for deciding the existential theory of a free group was shown to be not primitive recursive in [10]. Based on [19], a new PSPACE algorithm for the existential theory of a free group, which also allows to include rational constraints for variables, was presented in [2].

Beside these results for free groups, also extensions to larger classes of groups were obtained in the past: [4, 5, 8, 20]. In [3], a general transfer theorem for existential and positive theories was shown: the decidability of the existential theory is preserved by graph products over groups — a construction that generalizes both free and direct products, see e.g. [7]. Moreover, it is shown in [3] that for a large class of graph products, the positive theory can be reduced to the existential theory. The aim of this paper is to prove similar transfer theorems for HNN-extensions and amalgamated free products. These two operations are of fundamental importance in combinatorial group theory [14]; they are recalled in Section 2 by equations (1) and (3).

One of the first important applications of HNN-extensions was a more transparent proof of the celebrated result of Novikov and Boone on the existence of a finitely presented group with an undecidable word problem, see e.g. [14]. Such a group can be

constructed by a series of HNN-extensions starting from a free group. This shows that there is no hope to prove a transfer theorem for HNN-extensions, similar to the one for graph products from [3]. Therefore we mainly consider HNN-extensions and amalgamated free products, where the subgroup A in (1) and (3), respectively, is finite. Those groups which can be built up from finite groups using the operations of amalgamated free products and HNN-extensions, both subject to the finiteness restrictions above, are precisely the virtually-free groups [1] (i.e., those groups with a free subgroup of finite index). Virtually-free groups have strong connections to formal language theory and infinite graph theory [18].

In Section 3, we consider existential theories. For an HNN-extension $\mathbb{G}$ of the form (1) where the subgroup A is finite, we prove that the existential theory of $\mathbb{G}$ with rational constraints is decidable if the existential theory of $\mathbb{H}$ with rational constraints is decidable (Thm. 1). In Section 4, we consider positive theories. For an HNN-extension $\mathbb{G}$ where the two isomorphic subgroups A and $\varphi(A)$ have finite intersection, we prove that the positive theory of $\mathbb{G}$ is decidable if the positive existential theory of $\mathbb{G}$ is decidable (Thm. 2). From Thm. 1 and 2 and their analogues for amalgamated free products we deduce that every finitely generated virtually-free group has a decidable existential theory with rational constraints as well as a decidable positive theory (Thm. 4). Our exposition will put emphasis on the case of HNN-extensions and just mention the adaptations to amalgamated free products. Full proofs can be found in the three manuscripts [11–13].

2 Preliminaries

The powerset of a set A is denoted by $\mathcal{P}(A)$. With $\mathrm{RAT}(\mathbb{M})$ (resp. $\mathcal{B}(\mathrm{RAT}(\mathbb{M}))$) we denote the class of all rational (resp. boolean combinations of rational) subsets of a monoid $\mathbb{M}$. The free product of two monoids $\mathbb{M}_1$ and $\mathbb{M}_2$ is denoted by $\mathbb{M}_1 * \mathbb{M}_2$. For a monoid $\mathbb{M}$, a bijection $h : \mathbb{M} \to \mathbb{M}$ is an *anti-automorphism* if $h(1_\mathbb{M}) = 1_\mathbb{M}$ and $h(a \cdot b) = h(b) \cdot h(a)$ for all $a, b \in \mathbb{M}$. It is called *involutive*, if $h^2(a) = a$ for all $a \in \mathbb{M}$. For two groups A and B, $\mathrm{PGI}(A, B)$ denotes the set of all partial isomorphisms from A to B, i.e., isomorphisms from some subgroup $C \leq A$ to some subgroup $D \leq B$. Let $\mathrm{PGI}\{A, B\} = \mathrm{PGI}(A, B) \cup \mathrm{PGI}(B, A) \cup \mathrm{PGI}(A, A) \cup \mathrm{PGI}(B, B)$.

HNN-extensions and amalgamated free products See [14] for background in combinatorial group theory. Let Γ be an alphabet and let $\Gamma^{-1} = \{a^{-1} \mid a \in \Gamma\}$ be a disjoint copy of Γ. A pair (Γ, R) with $R \subseteq (\Gamma \cup \Gamma^{-1})^*$ is called a *group presentation*. Elements in R are also called *relations*. The group presented by (Γ, R) is usually denoted by $\langle \Gamma; R \rangle$, and is defined as the quotient monoid $(\Gamma \cup \Gamma^{-1})^*/\rho$, where ρ is the smallest congruence relation on the free monoid $(\Gamma \cup \Gamma^{-1})^*$, which contains all pairs in $\{(aa^{-1}, \varepsilon), (a^{-1}a, \varepsilon) \mid a \in \Gamma\} \cup \{(r, \varepsilon) \mid r \in R\}$; note that this quotient is indeed a group. Instead of $\langle \Gamma; \{r_i \mid i \in I\} \rangle$, we also write $\langle \Gamma; r_i (i \in I) \rangle$. Clearly, every group is isomorphic to a group of the form $\langle \Gamma; R \rangle$ (we do not assume Γ to be finite). For a group $G \simeq \langle \Gamma; R \rangle$, an alphabet Σ with $\Sigma \cap \Gamma = \emptyset$ and a new set of relations $P \subseteq (\Gamma \cup \Sigma \cup \Gamma^{-1} \cup \Sigma^{-1})^*$ we denote with $\langle G, \Sigma; P \rangle$ the group $\langle \Sigma \cup \Gamma;\ P \cup R \rangle$.

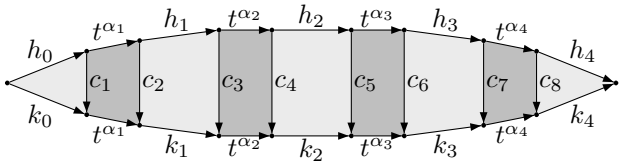
Let $\mathbb{H}$ be a group (the base group), together with two proper subgroups $A \leq \mathbb{H}, B \leq \mathbb{H}$ and an isomorphism $\varphi : A \to B$. Let $t \notin \mathbb{H}$ be a new generator. Then, the group

$$\mathbb{G} = \langle \mathbb{H}, t; t^{-1}at = \varphi(a)(a \in A) \rangle \tag{1}$$

is called an *HNN-extension of* $\mathbb{H}$ *by the stable letter* t, where A and B are *associated*. It is well known that $\mathbb{H}$ is a subgroup of $\mathbb{G}$. Clearly, there is a natural projection $\pi_{\mathbb{G}} : \mathbb{H} * \{t, t^{-1}\}^* \to \mathbb{G}$. An element s from the free product $\mathbb{H} * \{t, t^{-1}\}^*$ can be written as

$$s = h_0 t^{\alpha_1} h_1 \cdots t^{\alpha_n} h_n, \tag{2}$$

where $n \in \mathbb{N}, \alpha_i \in \{1, -1\}$, and $h_i \in \mathbb{H}$. It is called a *reduced sequence* iff it has neither a factor of the form $t^{-1}at$ with $a \in A$ nor tbt^{-1} with $b \in B$. We denote by $\mathrm{Red}(\mathbb{H}, t)$ the set of all *reduced* t*-sequences*; one has $\mathbb{G} = \pi_{\mathbb{G}}(\mathrm{Red}(\mathbb{H}, t))$. Reduced t-sequences turned out to be the right representations for elements from $\mathbb{G}$ for the purpose of deciding $\mathrm{Th}_{\exists}(\mathbb{G}, \mathrm{RAT}(\mathbb{G}))$. Let $\sim$ be the smallest congruence over $\mathbb{H} * \{t, t^{-1}\}^*$ generated by the rules $at \sim t\varphi(a)$ for all $a \in A$ and $bt^{-1} \sim t^{-1}\varphi^{-1}(b)$ for all $b \in B$. The congruence $\approx$ is the kernel of $\pi_{\mathbb{G}} : \mathbb{H} * \{t, t^{-1}\}^* \to \mathbb{G}$. Note that $u \sim v$ implies $u \approx v$. Moreover, if $u, v \in \mathrm{Red}(\mathbb{H}, t)$, then $u \approx v$ iff $u \sim v$. The fact $u \sim v$ for $u, v \in \mathrm{Red}(\mathbb{H}, t)$ can be visualized by a Van Kampen diagram (see [14]) in the group $\mathbb{G}$ of the following form, where $u = h_0 t^{\alpha_1} h_1 t^{\alpha_2} h_2 t^{\alpha_3} h_3 t^{\alpha_4} h_4$, $v = k_0 t^{\alpha_1} k_1 t^{\alpha_2} k_2 t^{\alpha_3} k_3 t^{\alpha_4} k_4$ with $h_0, k_0, \ldots, h_4, k_4 \in \mathbb{H}$ and $c_1, \ldots, c_8 \in A \cup B$. Light-shaded (resp. dark-shaded) areas represent relation in $\mathbb{H}$ (resp. group identities of the form $at = t\varphi(a)$ ($a \in A$) and $bt^{-1} = t^{-1}\varphi^{-1}(b)$ ($b \in B$)).



(†)

Now assume that $\mathbb{H}$ and $\mathbb{J}$ are groups with proper subgroups $A < \mathbb{H}$ and $B < \mathbb{J}$ and let $\varphi : A \to B$ be an isomorphism. Then

$$\mathbb{G} = \langle \mathbb{H} * \mathbb{J}, a = \varphi(a)(a \in A) \rangle \tag{3}$$

is called an *amalgamated free product* of $\mathbb{H}$ and $\mathbb{J}$, where A and B are associated.

Logical theories Let us fi x a countable group $\mathbb{G}$, let $\mathcal{C} \subseteq \mathcal{P}(\mathbb{G})$ be a set of *constraints*, and let Ω be an infi nite set of variables ranging over $\mathbb{G}$. Formulas of fi rst-order logic over $\mathbb{G}$ with constraints from $\mathcal{C}$ are built up from atomic formulas of the form $x \in L$ ($L \in \mathcal{C}$, $x \in \Omega$) and equations $u = v$ ($u, v \in (\Omega \cup \{x^{-1} \mid x \in \Omega\} \cup \mathbb{G})^*$) using boolean connectives and quantifi cations over variables. A formula θ is called *positive* if there are no negations in φ, i.e., conjunction and disjunction are the only boolean operators in θ. A formula is called *existential* (resp. existential positive) if it is of the form $\exists x_1 \cdots \exists x_n : \psi(x_1, \ldots, x_n)$, where ψ is a boolean (resp. a positive boolean) combination of atomic formulas. We denote with $\mathrm{Th}_+(\mathbb{G}, \mathcal{C})$ (resp. $\mathrm{Th}_{\exists}(\mathbb{G}, \mathcal{C}), \mathrm{Th}_{\exists+}(\mathbb{G}, \mathcal{C})$) the set of all positive (resp. existential, existential positive) sentences that are true in $\mathbb{G}$. We briefly write $\mathrm{Th}_X(\mathbb{G})$ for $\mathrm{Th}_X(\mathbb{G}, \emptyset)$ ($X \in \{\exists, +, \exists+\}$).

3 Existential theories

The following theorem is our main result concerning existential theories:

Theorem 1. $\mathrm{Th}_\exists(\mathbb{G}, \mathrm{RAT}(\mathbb{G}))$ *is decidable in the following two cases:*

(1) $\mathbb{G} = \langle \mathbb{H}, t; t^{-1}at = \varphi(a)(a \in A)\rangle$ *is an HNN-extension, where* A *and* $\varphi(A)$ *are proper subgroups of* $\mathbb{H}$ *with* A *finite, and* $\mathrm{Th}_\exists(\mathbb{H}, \mathrm{RAT}(\mathbb{H}))$ *is decidable.*

(2) $\mathbb{G} = \langle \mathbb{H} * \mathbb{J}, a = \varphi(a)(a \in A)\rangle$ *is an amalgamated free product, where* A *is finite, and* $\mathrm{Th}_\exists(\mathbb{H}, \mathrm{RAT}(\mathbb{H}))$ *and* $\mathrm{Th}_\exists(\mathbb{J}, \mathrm{RAT}(\mathbb{J}))$ *are decidable.*

The statements (1) and (2) in Thm. 1 are orthogonal to the corresponding result for graph products from [3]: none of the three operations (HNN-extensions, amalgamated free products, and graph products) is a special case of another one. At the end of Section 3, we will mention several variants of Thm. 1, which can be obtained by similar techniques. In the following we will sketch a proof of (1) from Thm. 1. Before we go into the details, we will first present some material concerning rational subsets of HNN-extensions, which is of independent interest.

3.1 Rational subsets of HNN-extensions

Let us fix throughout this section an HNN-extension $\mathbb{G}$ of a base group $\mathbb{H}$ as described by (1), where A is *finite*. We now define a notion of finite automata which will be well-suited for deciding $\mathrm{Th}_\exists(\mathbb{G}, \mathrm{RAT}(\mathbb{G}))$.

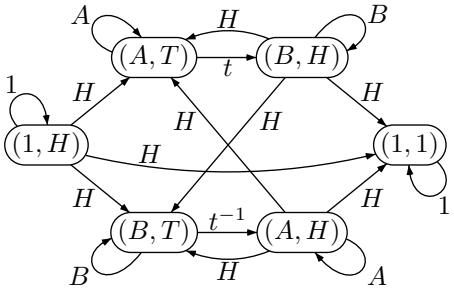
A *finite t-automaton* over $\mathbb{H} * \{t, t^{-1}\}^*$ with labeling set $\mathcal{F} \subseteq \mathcal{P}(\mathbb{H})$ is a 5-tuple

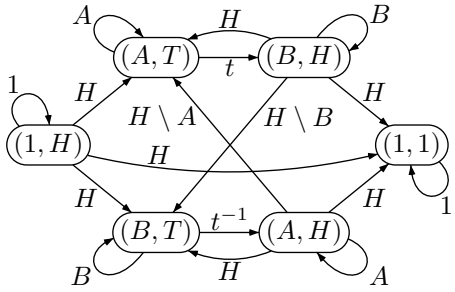
$$\mathcal{A} = \langle \mathcal{L}, \mathsf{Q}, \Delta, \mathsf{I}, \mathsf{T}\rangle, \tag{4}$$

where: (i) $\mathcal{L}$ is a finite subset of $\mathcal{F} \cup \mathcal{P}(A) \cup \mathcal{P}(B) \cup \{\{t\}, \{t^{-1}\}\}$, (ii) Q is a finite set of states, (iii) $\mathsf{I} \subseteq \mathsf{Q}$ is the set of initial states, (iv) $\mathsf{T} \subseteq \mathsf{Q}$ is the set of terminal states, and (v) $\Delta \subseteq \mathsf{Q} \times \mathcal{L} \times \mathsf{Q}$ is the set of transitions. Such an automaton induces a representation map $\mu_\mathcal{A} : \mathbb{H} * \{t, t^{-1}\}^* \to \mathcal{P}(\mathsf{Q} \times \mathsf{Q})$ defined as follows, where $x \in \mathbb{H} \cup \{t, t^{-1}\} \setminus \{1\}$ and $s \in \mathbb{H} * \{t, t^{-1}\}^*$ is of the form (2):

$$\begin{aligned}
\mu_{\mathcal{A},0}(1) &= \{(q,q) \mid q \in Q\} \cup \{(q,r) \in Q \times Q \mid \exists (q, L, r) \in \Delta : 1 \in L\} \\
\mu_{\mathcal{A},0}(x) &= \{(q,r) \in Q \times Q \mid \exists (q, L, r) \in \Delta : x \in L\} \\
\mu_\mathcal{A}(s) &= \mu_{\mathcal{A},0}(h_0) \circ \mu_{\mathcal{A},0}(t^{\alpha_1}) \circ \mu_{\mathcal{A},0}(h_1) \cdots \mu_{\mathcal{A},0}(t^{\alpha_n}) \circ \mu_{\mathcal{A},0}(h_n).
\end{aligned}$$

$\mathcal{A}$ recognizes the set $\mathrm{L}(\mathcal{A}) = \{s \in \mathbb{H} * \{t, t^{-1}\}^* \mid (\mathsf{I} \times \mathsf{T}) \cap \mu_\mathcal{A}(s) \neq \emptyset\}$. Let $\mathcal{G}_6 = (\mathcal{T}_6, \mathcal{E}_6)$ and $\mathcal{R}_6 = (\mathcal{T}_6, \mathcal{E}'_6)$ be the following two graphs:





Let $\widehat{\mathcal{E}}_6 = \{(p,\ell,q) \mid \exists (p,L,q) \in \mathcal{E}_6, \ell \in L\}$. One can check that $\mathcal{G}_6$ (resp. $\mathcal{R}_6$) endowed with the unique initial state $(1,H)$ and the unique final state $(1,1)$ is a finite t-automaton recognizing $\mathbb{H} * \{t,t^{-1}\}^*$ (resp. $\mathrm{Red}(\mathbb{H},t)$). Nodes of $\mathcal{G}_6$, i.e., elements of $\mathcal{T}_6$, are called *vertex-types*. We define a finite partial semigroup $\langle \mathcal{T}, \cdot \rangle$, where $\mathcal{T} = \mathcal{T}_6 \times \mathbb{B} \times \mathcal{T}_6$ and $\mathbb{B} = \langle \{0,1\}, \vee \rangle$ is the monoid of booleans. The partial product on $\mathcal{T}$ is defined by:

$$\forall (p,b,q),(p',b',q') \in \mathcal{T}_6 \times \mathbb{B} \times \mathcal{T}_6 : (p,b,q)\cdot(p',b',q') = \begin{cases} (p, b \vee b', q') & \text{if } q = p' \\ \text{undefined} & \text{otherwise} \end{cases}$$

The structure $\langle \mathcal{P}(\mathcal{T}), \cdot \rangle$ is thus a (total) monoid. Elements of $\mathcal{T}$ are also called *path-types*. We define an involution $\mathbb{I}_{\mathcal{R}} : \mathcal{T}_6 \to \mathcal{T}_6$ by $(A,T) \leftrightarrow (A,H)$, $(B,T) \leftrightarrow (B,H)$, and $(1,H) \leftrightarrow (1,1)$. It induces an involution $\mathbb{I}_{\mathcal{T}} : \mathcal{T} \to \mathcal{T}$ defined by: $\mathbb{I}_{\mathcal{T}}(p,b,q) = (\mathbb{I}_{\mathcal{R}}(q), b, \mathbb{I}_{\mathcal{R}}(p))$. This map $\mathbb{I}_{\mathcal{T}}$ is an anti-automorphism of $\mathcal{T}$ and also induces an involutive anti-automorphism of $\langle \mathcal{P}(\mathcal{T}), \cdot \rangle$ that will be denoted by $\mathbb{I}_{\mathcal{T}}$ too. We associate with every element $(p,b,q) \in \mathcal{T}$ its *initial group* $\mathrm{Gi}(p,b,q) = p_1(p) \in \{1,A,B\}$ and its *end group* $\mathrm{Ge}(p,b,q) = p_1(q) \in \{1,A,B\}$. Here p_1 is the projection onto the first component. For $s \in \mathbb{H} * \{t,t^{-1}\}^*$ let $b(s) = 1$ if s contains at least one occurrence of t or t^{-1}, otherwise $b(s) = 0$. Define

$$\gamma_t(s) = \{(p,b(s),q) \in \mathcal{T}_6 \times \mathbb{B} \times \mathcal{T}_6 \mid (p,q) \in \mu_{\mathcal{R}_6}(s)\} \in \mathcal{P}(\mathcal{T}). \tag{5}$$

Let $\mathcal{TA} = \{(\theta, b(x), \theta') \in \mathcal{T} \mid (\theta, x, \theta') \in \widehat{\mathcal{E}}_6\}$ be the set of *atomic path types*. A *normal finite t-automaton* over $\mathcal{F}$ is a 6-tuple $\mathcal{A} = \langle \mathcal{L}, \mathsf{Q}, \tau, \Delta, \mathsf{I}, \mathsf{T} \rangle$, where $\langle \mathcal{L}, \mathsf{Q}, \Delta, \mathsf{I}, \mathsf{T} \rangle$ is as in (4) and $\tau : Q \to \mathcal{T}_6$ maps each state to a vertex-type such that

$$\tau(\mathsf{I}) = \{(1,H)\},\ \tau(\mathsf{T}) = \{(1,1)\},\ \forall (q,L,r) \in \Delta : \{\tau(q)\} \times L \times \{\tau(r)\} \subseteq \widehat{\mathcal{E}}_6,$$

$$[\mathrm{L}(\mathcal{A})]_{\approx} = [\mathrm{L}(\mathcal{A}) \cap \mathrm{Red}(\mathbb{H},t)]_{\approx}, \tag{6}$$

$$\forall s,s' \in \mathbb{H} * \{t,t^{-1}\}^* : s \sim s' \Rightarrow \mu_{\mathcal{A}}(s) = \mu_{\mathcal{A}}(s'), \tag{7}$$

$$\forall \tilde{\theta} \in \gamma_t(s), \tilde{\theta}' \in \gamma_t(s') : \tilde{\theta}\cdot\tilde{\theta}' \text{ defined in } \mathcal{T} \Rightarrow$$

$$\mu_{\mathcal{A},1}(\tilde{\theta}\cdot\tilde{\theta}', s\cdot s') = \mu_{\mathcal{A},1}(\tilde{\theta}, s) \cdot \mu_{\mathcal{A},1}(\tilde{\theta}', s'),$$

$$\forall \theta \in \mathcal{T}_6 : \mu_{\mathcal{A},1}((\theta,0,\theta),1) = \mathrm{id}_{\tau^{-1}(\theta)}.$$

Here, $\mu_{\mathcal{A},1}((\theta,b,\theta'),s) = \mu_{\mathcal{A}}(s) \cap \tau^{-1}(\theta) \times \tau^{-1}(\theta')$; it does not depend on $b \in \{0,1\}$. $\mathcal{A}$ is said to be *strict* if, instead of (6), it fulfills the condition $\mathrm{L}(\mathcal{A}) \subseteq \mathrm{Red}(\mathbb{H},t)$.

Lemma 1. *We have:*

- $R \in \mathrm{RAT}(\mathbb{G})$ *iff* $R = \pi_{\mathbb{G}}(\mathrm{L}(\mathcal{A}))$ *for some normal finite t-automaton $\mathcal{A}$ with labeling set* $\mathrm{RAT}(\mathbb{H})$.
- *If* $R \in \mathcal{B}(\mathrm{RAT}(\mathbb{G}))$ *then* $R = \pi_{\mathbb{G}}(\mathrm{L}(\mathcal{A}))$ *for some strict normal finite t-automaton $\mathcal{A}$ with labeling set* $\mathcal{B}(\mathrm{RAT}(\mathbb{H}))$.

3.2 Deciding $\mathbf{Th}_{\exists}(\mathbb{G}, \mathbf{RAT}(\mathbb{G}))$

AB-algebras and AB-homomorphisms In this section, we introduce an algebraic structure which is devised for handling equations with rational constraints in an HNN-extension. Let A, B be two groups (later, these will be the two subgroups A and $B =$

$\varphi(A)$ from (1)) and Q be some finite set (it will be the state set of a t-automaton). Let $\mathsf{B}(\mathsf{Q}) = (\mathcal{P}(\mathsf{Q} \times \mathsf{Q}), \cdot)$ be the monoid of binary relations over Q and let $\mathsf{B}^2(\mathsf{Q})$ be the direct product $\mathsf{B}(\mathsf{Q}) \times \mathsf{B}(\mathsf{Q})$. For $m \in \mathsf{B}(\mathsf{Q})$ let $m^{-1} = \{(p,q) \in \mathsf{Q} \times \mathsf{Q} \mid (q,p) \in m\} \in \mathsf{B}(\mathsf{Q})$. Let $\mathbb{I}_\mathsf{Q} : \mathsf{B}^2(\mathsf{Q}) \to \mathsf{B}^2(\mathsf{Q})$ be the involutive anti-automorphism defined by $\mathbb{I}_\mathsf{Q}(m, m') = (m'^{-1}, m^{-1})$. An *$AB$-algebra* is a structure $\langle \mathbb{M}, \cdot, 1_\mathbb{M}, \mathbb{I}, \iota_A, \iota_B, \gamma, \mu, \delta \rangle$, where $\langle \mathbb{M}, \cdot, 1_\mathbb{M} \rangle$ is a monoid, $\iota_A : A \to \mathbb{M}, \iota_B : B \to \mathbb{M}$ are injective monoid homomorphisms, $\mathbb{I} : \mathbb{M} \to \mathbb{M}$ is an involutive anti-automorphism, and $\gamma : \mathbb{M} \to \mathcal{P}(\mathcal{T})$, $\mu : \mathcal{T} \times \mathbb{M} \to \mathsf{B}^2(\mathsf{Q})$, and $\delta : \mathcal{T} \times \mathbb{M} \to \mathrm{PGI}\{A, B\}$ are total mappings fulfilling the axioms (8)–(13) below.

For all $m, m' \in \mathbb{M}$ and all $\tilde{\theta} \in \gamma(m), \tilde{\theta}' \in \gamma(m')$:

$$\gamma(m) \cdot \gamma(m') \subseteq \gamma(m \cdot m') \tag{8}$$

$$\tilde{\theta} \cdot \tilde{\theta}' \text{ defined } \Rightarrow \mu(\tilde{\theta} \cdot \tilde{\theta}', m \cdot m') = \mu(\tilde{\theta}, m) \cdot \mu(\tilde{\theta}', m') \tag{9}$$

$$\mathrm{dom}(\delta(\tilde{\theta}, m)) \subseteq \mathrm{Gi}(\tilde{\theta}), \quad \mathrm{im}(\delta(\tilde{\theta}, m)) \subseteq \mathrm{Ge}(\tilde{\theta}) \tag{10}$$

$$\tilde{\theta} \cdot \tilde{\theta}' \text{ defined } \Rightarrow \delta(\tilde{\theta} \cdot \tilde{\theta}', m \cdot m') = \delta(\tilde{\theta}, m) \circ \delta(\tilde{\theta}', m') \tag{11}$$

For all $a \in A$, $b \in B$, $m \in \mathbb{M}$, and $\tilde{\theta} \in \gamma(m)$:

$$\mathbb{I}(\iota_A(a)) = \iota_A(a^{-1}), \quad \mathbb{I}(\iota_B(b)) = \iota_B(b^{-1}), \tag{12}$$

$$\gamma(\mathbb{I}(m)) = \mathbb{I}_\mathcal{T}(\gamma(m)),\ \mu(\mathbb{I}_\mathcal{T}(\tilde{\theta}), \mathbb{I}(m)) = \mathbb{I}_\mathsf{Q}(\mu(\tilde{\theta}, m)),\ \delta(\mathbb{I}_\mathcal{T}(\tilde{\theta}), \mathbb{I}(m)) = \delta(\tilde{\theta}, m)^{-1} \tag{13}$$

Let $\mathcal{M}_i = \langle \mathbb{M}_i, \cdot, 1_{\mathbb{M}_i}, \iota_{A,i}, \iota_{B,i}, \mathbb{I}_i, \gamma_i, \mu_i, \delta_i \rangle$ ($i \in \{1, 2\}$) be two AB-algebras with the same underlying groups A, B and set Q. An *AB-homomorphism* from $\mathcal{M}_1$ to $\mathcal{M}_2$ is a monoid homomorphism $\psi : \mathbb{M}_1 \to \mathbb{M}_2$ fulfilling the five properties (14)–(18) below:

$$\forall a \in A\, \forall b \in B : \psi(\iota_{A,1}(a)) = \iota_{A,2}(a) \ \wedge\ \psi(\iota_{B,1}(b)) = \iota_{B,2}(b) \tag{14}$$

$$\forall m \in \mathbb{M}_1 : \mathbb{I}_2(\psi(m)) = \psi(\mathbb{I}_1(m)) \tag{15}$$

$$\forall m \in \mathbb{M}_1 : \gamma_2(\psi(m)) \supseteq \gamma_1(m) \tag{16}$$

$$\forall m \in \mathbb{M}_1\, \forall \tilde{\theta} \in \gamma_1(m) : \mu_2(\tilde{\theta}, \psi(m)) = \mu_1(\tilde{\theta}, m) \tag{17}$$

$$\forall m \in \mathbb{M}_1\, \forall \tilde{\theta} \in \gamma_1(m) : \delta_2(\tilde{\theta}, \psi(m)) = \delta_1(\tilde{\theta}, m) \tag{18}$$

In the following we will introduce two particular AB-algebras.

The AB-algebra $\mathbb{H}_t$ From now on, we fix an HNN-extension (1) with A and $B = \varphi(A)$ finite and a strict normal finite t-automaton $\mathcal{A} = \langle \mathcal{L}, \mathsf{Q}, \tau, \Delta, \mathsf{I}, \mathsf{T} \rangle$ with labeling set $\mathcal{B}(\mathrm{RAT}(\mathbb{H}))$. We define an AB-algebra

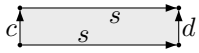
$$\langle \mathbb{H} * \{t, t^{-1}\}^*, \cdot, 1_\mathbb{H}, \iota_A, \iota_B, \mathbb{I}_t, \gamma_t, \mu_t, \delta_t \rangle$$

with underlying monoid $\mathbb{H} * \{t, t^{-1}\}^*$ and set of states Q as follows: ι_A (resp. ι_B) is the natural injection from A (resp. B) into $\mathbb{H} * \{t, t^{-1}\}^*$, and $\mathbb{I}_t$ is the unique involutive anti-automorphism $\mathbb{H} * \{t, t^{-1}\}^* \to \mathbb{H} * \{t, t^{-1}\}^*$ such that $\mathbb{I}_t(h) = h^{-1}$ for $h \in \mathbb{H}$, $\mathbb{I}_t(t) = t^{-1}$, and $\mathbb{I}_t(t^{-1}) = t$. The map γ_t was already defined in (5). The maps μ_t :

$\mathcal{T} \times \mathbb{H} * \{t, t^{-1}\}^* \to \mathsf{B}^2(\mathsf{Q})$ and $\delta_t : \mathcal{T} \times \mathbb{H} * \{t, t^{-1}\}^* \to \mathrm{PGI}\{A, B\}$ are defined as follows, where $s \in \mathbb{H} * \{t, t^{-1}\}^*$ and $\tilde{\theta} \in \mathcal{T}$:

$$\mu_t(\tilde{\theta}, s) = (\mu_{\mathcal{A},1}(\tilde{\theta}, s), (\mu_{\mathcal{A},1}(\mathbb{I}_{\mathcal{T}}(\tilde{\theta}), \mathbb{I}_t(s)))^{-1})$$
$$\delta_t(\tilde{\theta}, s) = \{(c, d) \in \mathrm{Gi}(\tilde{\theta}) \times \mathrm{Ge}(\tilde{\theta}) \mid cs \sim sd\}.$$

Note that $(c, d) \in \delta(\tilde{\theta}, s)$ implies that in the group $\mathbb{G}$ there is a Van Kampen diagram as shown on the right, which (for $s \in \mathrm{Red}(\mathbb{H}, t)$) is a diagram of the form (†); note that $c, d \in A \cup B$. E.g., if $\alpha_2 = \alpha_3 = 1$ and $h_2 = k_2$ in (†), then $(c_3, c_6) \in \delta_t(((A, T), 1, (B, H)), th_2t)$. (‡)



One can check that the monoid congruence $\sim$ is compatible with $\mathbb{I}_t$, ι_A, ι_B, γ_t, μ_t, and δ_t (here, (7) is important) so that the quotient $\mathbb{H}_t = \mathbb{H} * \{t, t^{-1}\}^* / \sim$ is naturally endowed with the structure of an AB-algebra (which we denote again with $\mathbb{H}_t$)

$$\mathbb{H}_t = \langle \mathbb{H}_t, \cdot, 1_{\mathbb{H}}, \iota_A, \iota_B, \mathbb{I}_\sim, \gamma_\sim, \mu_\sim, \delta_\sim \rangle. \tag{19}$$

Intuitively, the values $\gamma_\sim(s)$, $\mu_\sim(\tilde{\theta}, s)$, and $\delta_\sim(\tilde{\theta}, s)$ (for $\tilde{\theta} \in \gamma_\sim(s)$) store all information about a sequence s that is relevant when s appears in a solution of a system of equations. Since A, B, and Q are finite, this is only a finite amount of information.

Normal systems of equations A normal system of (dis)equations with constraints from $\mathcal{B}(\mathrm{RAT}(\mathbb{G}))$ is a tuple

$$\mathcal{S}_{\mathbb{G}} = ((u_i = u_i')_{1 \le i \le n}, (u_i \neq u_i')_{n < i \le 2n}, \mu_{\mathcal{A}}, \mu_{\mathcal{U}}), \tag{20}$$

where u_i, u_i' are words over an alphabet of unknowns $\mathcal{U}$, $|u_i| = 1, |u_i'| = 2$ for $1 \le i \le n$, $|u_i| = 1 = |u_i'|$ for $n < i \le 2n$, $\mu_{\mathcal{A}}$ is the representation map associated with the strict normal t-automaton $\mathcal{A}$ from the previous paragraph, and $\mu_{\mathcal{U}} : \mathcal{U} \to \mathsf{B}(\mathsf{Q})$. A solution of the system (20) is any monoid homomorphism $\sigma_{\mathbb{G}} : \mathcal{U}^* \to \mathbb{G}$ such that for all $1 \le i \le n$, $n < j \le 2n$, and $U \in \mathcal{U}$:

$$\sigma_{\mathbb{G}}(u_i) = \sigma_{\mathbb{G}}(u_i'), \;\; \sigma_{\mathbb{G}}(u_j) \neq \sigma_{\mathbb{G}}(u_j'), \;\; \mu_{\mathcal{A},1}(((1, H), b, (1, 1)), \sigma_{\mathbb{G}}(U)) = \mu_{\mathcal{U}}(U),$$

where $b \in \{0, 1\}$ ($\mu_{\mathcal{A},1}$ does not depend on the concrete value of b). Since $\mathcal{A}$ is strict normal, $\mu_{\mathcal{A},1}(\tilde{\theta}, g)$ for $g \in \mathbb{G}$ can be defined as $\mu_{\mathcal{A},1}(\tilde{\theta}, s)$ for any $s \in \mathrm{Red}(\mathbb{H}, t)$ with $\pi_{\mathbb{G}}(s) = g$. Using Lemma 1, one can reduce $\mathrm{Th}_\exists(\mathbb{G}, \mathrm{RAT}(\mathbb{G}))$ to the question whether a system of the form (20) has a solution. Thus, we may assume to have a system of the form (20) and we aim to decide whether it has a solution.

The AB-algebra $\mathbb{W}_t$ Whereas our first AB-algebra $\mathbb{H}_t$ from (19) depends on the "concrete" base group $\mathbb{H}$, we now introduce a second "generic" AB-algebra $\mathbb{W}_t$, which depends on our input system (20), but it depends only superficially on $\mathbb{H}$. The idea is to factorize the $\mathbb{G}$-values of a concrete solution of our given system (20) into "generic" symbols, which generate our new AB-algebra $\mathbb{W}_t$. Every generic symbol can be instantiated in $\mathbb{G}$ so that the original solution in $\mathbb{G}$ is recovered.

In order to carry out the above factorization, we introduce for every atomic type $\tilde{\theta} \in \mathcal{TA}$, every $\alpha \in \mathsf{B}^2(\mathsf{Q})$, and every $\beta \in \mathrm{PGI}(\mathrm{Gi}(\tilde{\theta}), \mathrm{Ge}(\tilde{\theta}))$, $54 \cdot n$ (n is from

(20)) many different new "generic" symbols $W_1, \ldots, W_{54n}$ and define: $\gamma(W_i) = \{\tilde{\theta}\}$, $\mu(\tilde{\theta}, W_i) = \alpha$, and $\delta(\tilde{\theta}, W_i) = \beta$. Let $\mathcal{W}$ be the new alphabet obtained in this way. By adding for every $W \in \mathcal{W}$ a new copy to $\mathcal{W}$, we can define on $\mathcal{W}$ an involution $\mathbb{I}$ without fixpoints (i.e., $\mathbb{I}(W) \neq W$) such that (13) holds for every $m = W \in \mathcal{W}$. Let us now consider the free product $\mathcal{W}^* * A * B$. We denote by $\iota_A : A \to \mathcal{W}^* * A * B$ (resp. $\iota_B : B \to \mathcal{W}^* * A * B$) the natural embedding of A (resp. B) into $\mathcal{W}^* * A * B$. We define the AB-algebra

$$\langle \mathcal{W}^* * A * B, \cdot, 1, \iota_A, \iota_B, \mathbb{I}, \mu, \gamma, \delta \rangle$$

with underlying monoid $\mathcal{W}^* * A * B$ and set of states Q as follows: $\mathbb{I}$ is extended as the unique involutive anti-automorphism $\mathcal{W}^* * A * B \to \mathcal{W}^* * A * B$ such that $\mathbb{I}(\iota_A(a)) = \iota_A(a^{-1})$ for $a \in A$ and $\mathbb{I}(\iota_B(b)) = \iota_B(b^{-1})$ for $b \in B$. The mapping $\gamma : \mathcal{W} \to \mathcal{P}(\mathcal{TA})$ is extended to $\iota_A(A) \cup \iota_B(B)$ by

$$\begin{aligned}
\forall a \in A \setminus \{1\} : \gamma(\iota_A(a)) &= \{((A,T),0,(A,T)),((A,H),0,(A,H))\},\\
\forall b \in B \setminus \{1\} : \gamma(\iota_B(b)) &= \{((B,T),0,(B,T)),((B,H),0,(B,H))\},\\
\gamma(1) &= \{(\theta,0,\theta) \mid \theta \in \mathcal{T}_6\},
\end{aligned}$$

and finally to the full free product $\mathcal{W}^* * A * B$ by

$$\forall g_1, \ldots, g_k \in \mathcal{W} \cup \iota_A(A) \cup \iota_B(B) : \gamma(g_1 \cdots g_k) = \gamma(g_1) \cdots \gamma(g_k).$$

The mappings $\mu : \mathcal{T} \times \mathcal{W} \to \mathsf{B}^2(\mathsf{Q})$ and $\delta : \mathcal{T} \times \mathcal{W} \to \mathrm{PGI}\{A, B\}$ are extended as follows:

$$\begin{aligned}
\forall a \in A\ \forall \tilde{\theta} \in \gamma(\iota_A(a)) : \delta(\tilde{\theta}, \iota_A(a)) &= \delta_t(\tilde{\theta}, a), \mu(\tilde{\theta}, \iota_A(a)) = \mu_t(\tilde{\theta}, a)\\
\forall b \in B\ \forall \tilde{\theta} \in \gamma(\iota_B(b)) : \delta(\tilde{\theta}, \iota_B(b)) &= \delta_t(\tilde{\theta}, b),\ \mu(\tilde{\theta}, \iota_B(b)) = \mu_t(\tilde{\theta}, b)
\end{aligned}$$

Finally, the maps μ and δ are extended to $\mathcal{W}^* * A * B$ in the only way such that for all $m \in \iota_A(A) \cup \iota_B(B) \cup \mathcal{W}, \tilde{\theta} \in \mathcal{T} \setminus \gamma(m)$: $\mu(\tilde{\theta}, m) = \emptyset$, $\delta(\tilde{\theta}, m) = \{(1,1)\}$ (the trivial partial isomorphism), and axioms (9) and (11) are respected. Let $\equiv$ be the smallest monoid congruence on $\mathcal{W}^* * A * B$ which contains all pairs (cW, Wd) with $W \in \mathcal{W}$ and $(c,d) \in \delta(\tilde{\theta}, W)$ for the unique $\tilde{\theta} \in \gamma(W)$. Let $\mathbb{W} := \mathcal{W}^* * A * B / \equiv$ be the quotient monoid, i.e., we enforce for every $W \in \mathcal{W}$ diagrams of the form (‡) (with $s = W$). One can check that $\equiv$ is compatible with $\mathbb{I}$, ι_A, ι_B, γ, μ, and δ, so that $\mathbb{W}$ inherits from $\mathcal{W}^* * A * B$ the structure of an AB-algebra. Let $\mathcal{W}_t$ be the set of all $W \in \mathcal{W}$ such that for some $s \in \mathbb{H} * \{t, t^{-1}\}^*$: (i) $\gamma(W) \subseteq \gamma_t(s)$ and (ii) the unique $\tilde{\theta} \in \gamma(W)$ fulfills $\mu(\tilde{\theta}, W) = \mu_t(\tilde{\theta}, s)$ and $\delta(\tilde{\theta}, W) = \delta_t(\tilde{\theta}, s)$. Thus, $\mathcal{W}_t$ is the set of all generic symbols that can be realized by a concrete sequence $s \in \mathbb{H} * \{t, t^{-1}\}^*$. With $\mathcal{W}_{\mathbb{H}} \subseteq \mathcal{W}_t$ we denote the set of those $W \in \mathcal{W}_t$ such that moreover $\gamma(W) = \{(\theta, 0, \theta')\}$, where $(\theta, H, \theta') \in \mathcal{E}_6$. Let $\mathbb{W}_t$ (resp. $\mathbb{W}_{\mathbb{H}}$) be the substructure of $\mathbb{W}$ generated by the subset of monoid generators $\iota_A(A) \cup \iota_B(B) \cup \mathcal{W}_t$ (resp. $\iota_A(A) \cup \iota_B(B) \cup \mathcal{W}_{\mathbb{H}}$). It is easy to see that $\psi(\mathbb{W}_{\mathbb{H}}) \subseteq \mathbb{H}$ for every AB-homomorphism $\psi : \mathbb{W}_t \to \mathbb{H}_t$.

The algorithm Recall that we have to check, whether the normal system of (dis)equations (20) has a solution.

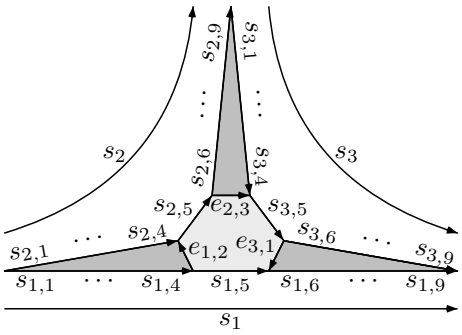
Step 1 Consider an equation $u_i = u_i'$ from (20), where w.l.o.g. $u_i = U_1$ and $u_i' = U_2U_3$ for $U_1, U_2, U_3 \in \mathcal{U}$; disequations can be treated similarly. Let $\sigma_{\mathbb{G}}$ be a solution for (20). We can choose reduced t-sequences s_1, s_2, and s_3 such that $\sigma_{\mathbb{G}}(U_j) = \pi_{\mathbb{G}}(s_j)$. Then there exists factorizations $s_j = s_{j,1} \cdots s_{j,9}$ and elements $e_{1,2}, e_{2,3}, e_{3,1} \in A \cup B$ such that the Van-Kampen diagram describing the group relation $\pi_{\mathbb{G}}(s_1) = \pi_{\mathbb{G}}(s_2 s_3)$ (i.e., $s_1 \approx s_2 s_3$) decomposes into four pieces, represented by the four relations

$$s_{1,1}\, s_{1,2}\, s_{1,3}\, s_{1,4}\, e_{1,2} \sim s_{2,1}\, s_{2,2}\, s_{2,3}\, s_{2,4} \tag{21}$$

$$s_{2,6}\, s_{2,7}\, s_{2,8}\, s_{2,9} \sim e_{2,3}\, \mathbb{I}_t(s_{3,4})\, \mathbb{I}_t(s_{3,3})\, \mathbb{I}_t(s_{3,2})\, \mathbb{I}_t(s_{3,1}) \tag{22}$$

$$e_{3,1}\, s_{1,6}\, s_{1,7}\, s_{1,8}\, s_{1,9} \sim s_{3,6}\, s_{3,7}\, s_{3,8}\, s_{3,9} \tag{23}$$

$$s_{1,5} = e_{1,2}\, s_{2,5}\, e_{2,3}\, s_{3,5}\, e_{3,1} \text{ in the base group } \mathbb{H}, \tag{24}$$



see the diagram on the right, where the light-shaded area represents a relation in the group $\mathbb{H}$. Dark-shaded areas are diagrams of the form (†) from Section 2. The $s_{j,k}$ ($k \neq 5$) belong to $\mathbb{H} * \{t, t^{-1}\}^*$, while the $s_{j,5} \in \mathbb{H}$. Decomposing e.g. the sequence $s_{1,1}s_{1,2}s_{1,3}s_{1,4}$ into 4 parts allows us to choose all the $s_{1,k}$ ($1 \leq k \leq 4$) either trivial or of some (guessed) atomic type in $\mathcal{TA}$. We now replace every $s_{j,k}$ by a new generic symbol $W_{i,j,k} \in \mathcal{W}_t$ (or possibly 1); the additional index i refers to the equation $u_i = u_i'$, where $W_{i,j,k}$ comes from. Note that for every i we need 27 symbols $W_{i,j,k}$, this explains the factor $54 = 2 \cdot 27$ in the definition of the alphabet $\mathcal{W}$. The values of the mappings $\mathbb{I}$, γ, μ, and δ on $W_{i,j,k}$ have to be chosen such that the generic symbol $W_{i,j,k}$ captures all the relevant data about the concrete sequence $s_{j,k}$. For instance, $\gamma(W_{i,j,k})$ only contains the guessed type for $s_{j,k}$. In this way, we can translate system (20) into a new system of equations over $\mathbb{W}_t$ (corresponding essentially to (21)–(23)) and another system over $\mathbb{H}$ (corresponding to (24)). Thus, we reduce the problem, whether (20) has a solution, to a finite disjunction of problems of the following form:

INPUT: Finitely many pairs $(v_j, v_j') \in \mathbb{W}_t \times \mathbb{W}_t$ ($j \in J$), with $\gamma(v_j) = \gamma(v_j') \neq \emptyset$, and a (ordinary) system $\mathcal{S}_{\mathbb{H}}$ of equations and disequations in the base group $\mathbb{H}$ and with constraints from $\mathrm{RAT}(\mathbb{H})$; the set of unknowns of $\mathcal{S}_{\mathbb{H}}$ is included in $\mathcal{W}_{\mathbb{H}}$.

QUESTION: Does there exist an AB-homomorphism $\sigma_t : \mathbb{W}_t \to \mathbb{H}_t$ such that

$$\forall j \in J : \sigma_t(v_j) = \sigma_t(v_j') \text{ and simultaneously} \tag{25}$$

$$\sigma_t \text{ solves the system } \mathcal{S}_{\mathbb{H}}? \tag{26}$$

Step 2 We reduce the question, whether (25) and (26) holds for some σ_t to the problem, where the input is the same as above, but the question is:

QUESTION: Do there exist AB-homomorphisms $\sigma_{\mathbb{W}} : \mathbb{W}_t \to \mathbb{W}_t$, $\psi_t : \mathbb{W}_t \to \mathbb{H}_t$ with

$$\forall j \in J : \sigma_{\mathbb{W}}(v_j) = \sigma_{\mathbb{W}}(v_j') \text{ and simultaneously} \tag{27}$$

$$\sigma_{\mathbb{W}} \circ \psi_t \text{ solves the system } \mathcal{S}_{\mathbb{H}}? \tag{28}$$

This reduction is a direct corollary of a *factorization* property for the solutions σ_t of (25): σ_t is a solution iff it can be factorized as $\sigma_{\mathbb{W}} \circ \psi_t$ for AB-homomorphisms $\sigma_{\mathbb{W}} : \mathbb{W}_t \to \mathbb{W}_t$ and $\psi_t : \mathbb{W}_t \to \mathbb{H}_t$. The proof consists in decomposing σ_t into a sort of elementary AB-homomorphisms of the form $W \mapsto cW_1dW_2eW_3f$ ($c, d, e, f \in A \cup B, W_i \in \mathcal{W}_t$), followed by some $\psi_t : \mathbb{W}_t \to \mathbb{H}_t$: we start with the trivial decomposition $\sigma_{\mathbb{W}} = \mathrm{id}_{\mathbb{W}}, \psi_t = \sigma_t$ and then reason by induction over the multiset $\{\!\{d(\sigma_{\mathbb{W}}(v_j), \sigma_{\mathbb{W}}(v'_j)) \mid j \in J\}\!\}$, where d is a kind of distance on $\mathbb{W}$.

Step 3 We introduce the group $\mathbb{U} = \langle \mathbb{W};\ W \cdot \mathbb{I}(W) = 1\ (W \in \mathcal{W}) \rangle$. This group turns out to be obtained from a base group $\mathbb{K}$, which is a semi-direct product of the finite group A by a free group of finite rank, by a finite number of HNN-extensions with associated subgroups *strictly smaller* than A. Using the main result of [2], one can show that $\mathrm{Th}_\exists(\mathbb{K}, \mathrm{RAT}(\mathbb{K}))$ is decidable. Moreover, by induction on the cardinality of A, also $\mathrm{Th}_\exists(\mathbb{U}, \mathrm{RAT}(\mathbb{U}))$ is decidable. One can show that for every AB-homomorphism $\sigma_{\mathbb{W}} : \mathbb{W}_t \to \mathbb{W}_t$ and every generator $W \in \mathcal{W}_{\mathbb{H}}$ one has $\sigma_{\mathbb{W}}(W) \in (A \cup B)\mathcal{W}_{\mathbb{H}}(A \cup B)$. This implies that for the restriction $\sigma_{\mathbb{H}} = \sigma_{\mathbb{W}}{\restriction}_{\mathbb{W}_{\mathbb{H}}}$ of $\sigma_{\mathbb{W}}$ in (27) there are only finitely many possibilities. By enumerating all these mappings $\sigma_{\mathbb{H}}$ and substituting them into (27) and (28), we reduce the simultaneous satisfiability of (27) and (28) to: (i) on one hand solving finitely many specialized instances of (27), which reduce to the theory $\mathrm{Th}_\exists(\mathbb{U}, \mathrm{RAT}(\mathbb{U}))$, and (ii) on the other hand, for every specialized instance from the previous point, solving a corresponding system of the form $\sigma_{\mathbb{H}}(\mathcal{S}_{\mathbb{H}})$, which reduces to $\mathrm{Th}_\exists(\mathbb{H}, \mathrm{RAT}(\mathbb{H}))$. Following this strategy we prove Thm. 1.

Using the embedding of an amalgamated free product (3) into the HNN-extension $\langle \mathbb{H} * \mathbb{J}', t; t^{-1}at = \varphi(a)(a \in A) \rangle$ by the map defined by $h \in \mathbb{H} \mapsto t^{-1}ht$, $j \in \mathbb{J} \mapsto j'$ (where $\mathbb{J}' = \{j' \mid j \in \mathbb{J}\}$ is a copy of $\mathbb{J}$, disjoint from $\mathbb{H}$, and φ maps every element of A to its copy in $\mathbb{J}'$, see [14, Thm. 2.6. p. 187]), we obtain statement (2) of Thm. 1. Let us finally discuss some variations of Thm. 1.

Remark 1. Thm. 1(1) remains true when $\mathrm{Th}_\exists(X, \mathrm{RAT}(X))$ ($X \in \{\mathbb{H}, \mathbb{G}\}$) is replaced by: $\mathrm{Th}_{\exists+}(X)$ (variant 1), $\mathrm{Th}_\exists(X)$ (variant 2), or $\mathrm{Th}_{\exists+}(X, \mathrm{RAT}(X))$ (variant 3). If $\mathrm{Th}_{\exists+}(\mathbb{H}, \{A_1, \ldots, A_n\})$ is decidable, where every A_i is a finitely generated subgroup of $\mathbb{H}$ containing A, then also $\mathrm{Th}_{\exists+}(\mathbb{G}, \{A_i, \langle A_i, t \rangle \mid 1 \leq i \leq n\})$ is decidable (variant 4), where $\langle A_i, t \rangle$ is the subgroup of $\mathbb{G}$ generated by $A_i \cup \{t\}$. These variants can be also shown if $\mathbb{H}$ is a cancellative monoid instead of a group (only A and B have to be groups). Finally, variant 2 still holds for amalgamated products of cancellative monoids.

4 Positive theories

The following two theorems are our main results concerning positive theories:

Theorem 2. $\mathrm{Th}_+(\mathbb{G})$ *is decidable in the following two cases:*

(1) $\mathbb{G} = \langle \mathbb{H}, t; t^{-1}at = \varphi(a)(a \in A) \rangle$ *is an HNN-extension, where* A *and* $\varphi(A)$ *are proper subgroups of* $\mathbb{H}$ *with* $A \cap \varphi(A)$ *finite, and* $\mathrm{Th}_{\exists+}(\mathbb{G})$ *is decidable.*

(2) $\mathbb{G} = \langle \mathbb{H} * \mathbb{J}, a = \varphi(a)(a \in A) \rangle$ *is an amalgamated free product with* A *finite and* $\mathrm{Th}_{\exists+}(\mathbb{G})$ *is decidable.*

In Thm. 2 we cannot allow a cancellative monoid for $\mathbb{H}$, because the positive theory of $\{a,b\}^* \simeq \mathbb{N} * \mathbb{N}$ is undecidable [6]. For the same reason, we cannot include rational constraints: $\{a,b\}^*$ is a rational subset of the free group of rank 2.

Let us sketch a proof of (1) from Thm. 2. Our strategy for reducing $\mathrm{Th}_+(\mathbb{G})$ to $\mathrm{Th}_{\exists+}(\mathbb{G})$ is similar to [16, 17]: From a positive sentence ψ, which is interpreted over $\mathbb{G}$, we construct an existential positive sentence ψ' with subgroup constraints of a very special form, which is interpreted over a multiple HNN-extension $\mathbb{G}'$ of $\mathbb{G}$, where only finite subgroups of $\mathbb{G}$ are associated. Roughly speaking, ψ' results from ψ by replacing the universally quantified variables by the stable letters of the HNN-extension $\mathbb{G}'$. Let $\mathbb{G}$ be an HNN-extension as in Thm. 2. Let $X \leq A \cap \varphi(A)$ be a (necessarily finite) subgroup of $\mathbb{H}$. With $\mathrm{In}(X)$ we denote the group of all automorphisms f of X such that for some $g \in \mathbb{G}$ we have: $f(c) = g^{-1}cg$ for all $c \in X$. For new constants $k_1, \ldots, k_m \notin \mathbb{G}$ and $f_1, \ldots, f_m \in \mathrm{In}(X)$ we define the multiple HNN-extension

$$\mathbb{G}^{f_1,\ldots,f_m}_{k_1,\ldots,k_m} = \langle \mathbb{G}, k_1, \ldots, k_m; k_i^{-1} c k_i = f_i(c)\ (c \in X, 1 \leq i \leq m) \rangle. \tag{29}$$

The following theorem yields the reduction from $\mathrm{Th}_+(\mathbb{G})$ to $\mathrm{Th}_{\exists+}(\mathbb{G})$.

Theorem 3. *There is a subgroup $X \leq A \cap B \leq \mathbb{H} \leq \mathbb{G}$ such that for every formula $\psi(z_1, \ldots, z_m) \equiv \forall x_1 \exists y_1 \cdots \forall x_n \exists y_n\, \phi(x_1, \ldots, x_n, y_1, \ldots, y_n, z_1, \ldots, z_m)$, where ϕ is a positive boolean combination of equations (with constants) over the group $\mathbb{G}$, and for all $u_1, \ldots, u_m \in \mathbb{G}$ we have: $\psi(u_1, \ldots, u_m)$ in $\mathbb{G}$ iff*

$$\bigwedge_{f_1 \in \mathrm{In}(X)} \exists y_1 \cdots \bigwedge_{f_n \in \mathrm{In}(X)} \exists y_n \left\{ \begin{array}{l} \bigwedge\limits_{1 \leq i \leq n} y_i \in \mathbb{G}^{f_1,\ldots,f_i}_{k_1,\ldots,k_i} \wedge \\ \phi(k_1, \ldots, k_n, y_1, \ldots, y_n, u_1, \ldots, u_m) \text{ in } \mathbb{G}^{f_1,\ldots,f_n}_{k_1,\ldots,k_n} \end{array} \right\} \tag{30}$$

In [3], a result analogous to Theorem 3 for the case that $\mathbb{G}$ is a free product was shown. In this case, the new generators $k_1, \ldots, k_n$ do not interact with the group $\mathbb{G}$, i.e., the HNN-extension $\mathbb{G}^{f_1,\ldots,f_n}_{k_1,\ldots,k_n}$ is replaced by the free product $\mathbb{G} * F_n$, where F_n is the free group generated by $k_1, \ldots, k_n$. For the more general case that $\mathbb{G}$ is an HNN-extension, we cannot avoid some nontrivial interaction between k_i and $\mathbb{G}_i$. This interaction is expressed by the identities $k_i^{-1} c k_i = f_i(c)$ ($c \in X$) in the HNN-extension $\mathbb{G}^{f_1,\ldots,f_n}_{k_1,\ldots,k_n}$. Note that the sentence in (30) is not interpreted in a single HNN-extension of $\mathbb{G}$. But it is not difficult to construct an HNN-extension $\mathbb{G}'$ of $\mathbb{G}$ such that each of the groups $\mathbb{G}^{f_1,\ldots,f_n}_{k_1,\ldots,k_n}$ can be embedded into $\mathbb{G}'$. Moreover, each single HNN-extension that leads from $\mathbb{G}$ to $\mathbb{G}'$ associates X with itself as in (29). In this way, we can construct from (30) an existential positive sentence $\Psi = (\exists y_\sigma \in \mathbb{G}_\sigma)_{\sigma \in J}\, \chi((k_\sigma)_{\sigma \in J}, (y_\sigma)_{\sigma \in J}, u_1, \ldots, u_m)$ (for some index set J larger than n in (30)) such that (30) iff Ψ is true in $\mathbb{G}'$. Moreover, all constraint-groups $\mathbb{G}_\sigma$ in Ψ are generated by $\mathbb{G}$ and some of the stable letters k_σ. To complete the proof of (1) in Thm. 2, notice that an iterated application of variant 4 from Remark 1 (recall that X is finite) enables us to reduce $\mathrm{Th}_{\exists+}(\mathbb{G}', \{\mathbb{G}_\sigma \mid \sigma \in J\})$ to $\mathrm{Th}_{\exists+}(\mathbb{G})$. A proof of (2) in Thm. 2 follows a similar strategy.

We conclude this paper with an application to virtually-free groups. A finitely generated group $\mathbb{G}$ is *virtually-free*, if it has a free subgroup of finite index. Since these groups have finite decompositions over finite groups by means of the operations (1) and (3) with A finite [1], we obtain from Thm. 1 and 2:

Theorem 4. *If* $\mathbb{G}$ *is virtually-free, then* $\mathrm{Th}_{\exists}(\mathbb{G}, \mathrm{RAT}(\mathbb{G}))$ *and* $\mathrm{Th}_{+}(\mathbb{G})$ *are decidable.*

Thm. 4 immediately leads to the question, whether also the full first-order theory of a virtually-free group is decidable. This is certainly a difficult question. The full proof of Kharlampovich and Myasnikov for the decidability of the theory of a free group (see [9] for an overview) takes several hundred pages. Moreover, there seems to be no obvious reduction from the theory of a virtually-free group to the theory of a free group.

References

1. W. Dicks and M. J. Dunwoody. *Groups Acting on Graphs*. Cambridge Univ. Press, 1989.
2. V. Diekert, C. Gutiérrez, and C. Hagenah. The existential theory of equations with rational constraints in free groups is PSPACE-complete. *Inf. Comput.*, 202(2):105–140, 2005.
3. V. Diekert and M. Lohrey. Word equations over graph products. In *Proc. FSTTCS 2003*, LNCS 2914, pages 156–167. Springer, 2003.
4. V. Diekert and M. Lohrey. Existential and positive theories of equations in graph products. *Theory Comput. Syst.*, 37(1):133–156, 2004.
5. V. Diekert and A. Muscholl. Solvability of equations in free partially commutative groups is decidable. *Int. J. Algebra Comput.*, 2006. to appear.
6. V. G. Durnev. Undecidability of the positive $\forall\exists^3$-theory of a free semi-group. *Sibirsky Matematicheskie Jurnal*, 36(5):1067–1080, 1995. English translation.
7. E. R. Green. *Graph Products of Groups*. PhD thesis, The University of Leeds, 1990.
8. B. Khan, A. G. Myasnikov, and D. E. Serbin. On positive theories of groups with regular free length function. Manuscript, 2005.
9. O. G. Kharlampovich and A. Myasnikov. Tarski's problem about the elementary theory of free groups has a positive solution. *Electron. Res. Announc. AMS*, 4(14):101–108, 1998.
10. A. Kościelski and L. Pacholski. Makanin's algorithm is not primitive recursive. *Theor. Comput. Sci.*, 191(1-2):145–156, 1998.
11. M. Lohrey and G. Sénizergues. Equations in HNN-extensions. Manuscript, 2006.
12. M. Lohrey and G. Sénizergues. Positive theories of HNN-extensions and amalgamated free products. Manuscript, 2006.
13. M. Lohrey and G. Sénizergues. Rational subsets of HNN-extensions. Manuscript, 2006.
14. R. C. Lyndon and P. E. Schupp. *Combinatorial Group Theory*. Springer, 1977.
15. G. S. Makanin. Equations in a free group. *Math. USSR, Izv.*, 21:483–546, 1983. English translation.
16. G. S. Makanin. Decidability of the universal and positive theories of a free group. *Math. USSR, Izv.*, 25:75–88, 1985. English translation.
17. Y. I. Merzlyakov. Positive formulas on free groups. *Algebra i Logika Sem.*, 5(4):25–42, 1966. In Russian.
18. D. E. Muller and P. E. Schupp. Groups, the theory of ends, and context-free languages. *J. Comput. Syst. Sci.*, 26:295–310, 1983.
19. W. Plandowski. Satisfiability of word equations with constants is in PSPACE. *J. ACM*, 51(3): 483–496, 2004.
20. E. Rips and Z. Sela. Canonical representatives and equations in hyperbolic groups. *Invent. Math.*, 120:489–512, 1995.