On equations and first-order theory of one-relator monoids

Albert Garreta[*] Robert D. Gray[†]

Abstract

We investigate systems of equations and the first-order theory of one-relator monoids. We describe a family $\mathcal{F}$ of one-relator monoids of the form $\langle A \mid w = 1 \rangle$ where for each monoid M in $\mathcal{F}$, the longstanding open problem of decidability of word equations with length constraints reduces to the Diophantine problem (i.e. decidability of systems of equations) in M. We achieve this result by finding an interpretation in M of a free monoid, using only systems of equations together with length relations. It follows that each monoid in $\mathcal{F}$ has undecidable positive AE-theory, hence in particular it has undecidable first-order theory. The family $\mathcal{F}$ includes many one-relator monoids with torsion $\langle A \mid w^n = 1 \rangle$ $(n > 1)$. In contrast, all one-relator groups with torsion are hyperbolic, and all hyperbolic groups are known to have decidable Diophantine problem. We further describe a different class of one-relator monoids with decidable Diophantine problem.

1 Introduction

Two important longstanding open algorithmic problems in algebra are the decidability of the conjugacy problem for one-relator groups, and of the word problem for one-relator monoids. Each of these problems is a special case of the much more general and open question of whether the Diophantine problem (decidability of systems of equations) is decidable in one-relator groups, or in one-relator monoids. A positive answer to any of these would give a positive resolution to one (or both) of the open questions about the word and conjugacy problems mentioned above. On the other hand, if the Diophantine problem turns out to be undecidable for one-relator monoids or one-relator groups, then this would give a natural undecidable decision problem for these classes, which could lead to further undecidability results for these classes that would then have the potential to shed new light on fundamental questions like the word and conjugacy problems.

[*] Albert Garreta, Department of Mathematics, University of the Basque Country, 48080 Bilbao, Spain, *albert.garreta@ehu.eus*,

[†] Robert D. Gray, School of Mathematics, University of East Anglia, Norwich NR4 7TJ, England, UK, *Robert.D.Gray@uea.ac.uk*

The first named author was supported by the ERC grant PCG-336983, by the Spanish Government grant IT974-16, and by the Ministry of Economy, Industry and Competitiveness of the Spanish Government Grant MTM2017-86802-P.

2010 Mathematics Subject Classification. 20M05, 20F70, 20F05, 20F10, 03B25

Key words and phrases. one-relator monoids, Diophantine problem, first-order theory, word equations with length constraints.

The Diophantine problem for one-relator groups has recently received attention in the literature, with positive results obtained for solvable Baumslag-Solitar groups $BS(1,n) = \langle a, b \mid a^{-1}ba = b^n \rangle$ ($n \in \mathbb{Z}$); see [38] (in the same paper the authors also solve the problem for wreath products of the form $A \wr \mathbb{Z}$, where A is a finitely generated abelian group). The aim of this paper is to initiate the study of Diophantine problems (and related model-theoretic questions) for one-relator monoids. We shall obtain both positive and negative (undecidability) results, and will also establish a close connection between these problems and the problem of solving word equations with length constraints, which is a longstanding open problem in computer science.

Our main result describes a family of one-relator monoids $\mathcal{F}$ such that for any $M \in \mathcal{F}$ it is possible to reduce decidability of word equations with length constraints —a longstanding open problem in computer science— to the Diophantine problem in M. We further prove decidability of the Diophantine problem for a certain class of one-relator monoids. As a corollary we obtain undecidability of the positive AE-theory (hence of the first-order theory) of any one-relator monoid belonging to $\mathcal{F}$. To the best of our knowledge, this provides the first examples of one-relator monoids with undecidable positive AE-theory (with coefficients), excluding the free monoid. Other examples of one-relator monoids with undecidable first-order theory with coefficients can be found in [37, Theorem 1].

Equations in monoids and groups have been widely studied during the past decades, being of interest in several areas, ranging from computer science to group and model theory. For a detailed account of the history, motivation and key results in this area we refer the reader to the survey articles [21, 35, 40, 61]. By the *Diophantine problem* we mean the algorithmic problem of determining if any given system of equations has a solution or not. Two classical results due to Makanin show that the Diophantine problem is decidable in any free monoid [47] and in any free group [48]. Based on Makanin's algorithm, Razborov [58] provided a powerful description of the sets of solutions to systems of equations in free groups via what were later called Makanin-Razaborov diagrams. This played a key part in the solution to the Tarski problems [39, 63] regarding groups elementary equivalent to a free group.

In subsequent years new decidability algorithms and descriptions of solutions have appeared: in [56] Plandowski describes a polynomial space algorithm for deciding word equations based on a compression technique. In [33] Jeż shows that word equations can be solved in non-deterministic linear space, and in [15] it is proved that the solution set of a word equation is an EDT0L language (in particular, it is an indexed language), furthermore this set can be computed in polynomial space [22]. More recently, in [65] Sela presents the first in a sequence of papers devoted to investigating the structure of sets of solutions to systems of equations over a free semigroup via a Makanin—Razborov diagram analogue. Diophantine problems have been extensively considered also in different classes of groups and monoids, see e.g. [14, 19, 20, 22, 23, 28, 43, 43, 62]. For us the most relevant result in this direction is the decidability of the Diophantine problem in hyperbolic groups [16, 60].

A variation relevant to the present paper is the problem of word equations with *length constraints* (in short, WELCs). This consist of a (system of) word equation(s) together with finitely many linear inequalities involving the length of solutions (see Subsection 2.1 for a formal definition). The problem of determining whether WELCs are decidable has been open for decades now and is of major interest in computer science. Some partial cases and variations have been successfully studied in [13, 17, 26, 42]. As hinted at in [17], extending word equations with constraints that involve some type of length relation or letter-counting seems to always lead to undecidability. Indeed, many problems closely related to WELCs are undecidable, as shown in some of the previous references.

WELCs are of interest in industry where they are applied for program verification, code debugging, security analysis, document spanning, etc. A WELC is a particular instance of a so-called Satisfibility Modulo Theory (SMT) problem, which, roughly speaking, is a satisfiability problem for a first order sentence that combines different types of formulas from different languages (such as the language of monoids, which allows to write word equations, and the language of Presburger arithmetic, which allows to write linear integer equations and inequalities). In practice, such problems are usually tackled by so-called SMT solvers, which are programs that rely on different heuristics for solving certain types of SMT problems (different SMT solvers support different possible languages and fragments of a theory). Usually, SMT solvers are desgined to be fast and usable in real life, which in turn means that often they are not complete i.e. it is not guaranteed that the solver will be able to correctly solve a given input. We refer to [8, 18] for further information on general SMT solvers and their applications. There exists a variety of fast SMT solvers which can handle in particular word equations with rational constraints and length constraints [1, 6, 7, 10, 25, 66, 68] (we stress again that these programs are not complete, i.e. they cannot successfully solve any input problem).

A further point of interest is that WELCs are reducible to the problem of solving systems of integer-coefficient polynomial equations in $\mathbb{Z}$ [52]. Thus a proof of undecidability of WELCs would provide a new solution to Hilbert's 10th Problem, which states that equations in the ring $\mathbb{Z}$ are undecidable [51].

We would like to emphasize how the Diophantine problem generalizes and contains many well-known and studied algorithmic problems. Notably, and as already mentioned, both the word problem and the conjugacy problem are particular cases of the Diophantine problem (see [3, 4, 53, 54, 55, 69, 70] for definitions and results regarding the conjugacy problem in monoids). Moreover, the left and right divisibility problems in monoids, as well as decidability of Green's orders $\leqslant_{\mathcal{R}}$ and $\leqslant_{\mathcal{L}}$ are particular instances of the Diophantine problem. Thus proving that the latter is decidable in some specific group or monoid implies that any of the previously mentioned problems are decidable. Conversely, undecidability of any of the mentioned problems implies undecidability of the Diophantine problem. In a similar vein, systems of equations are particular instances of positive AE-formulas, which in turn are first-order formulas. Hence similar considerations hold for the problem of decidability of the positive AE-theory, or of the first-order theory, of a monoid or group.

There are several natural classes of one-relator monoids for which the word problem has been shown to be decidable. Specifically, Adjan [2] showed that all one-relator monoids defined by presentations of the form $\langle A \mid w = 1\rangle$ have decidable word problem. Monoid presentations where all of the relations are of the form $w = 1$ are commonly called *special* presentations. Adjan solved the word problem for special one-relator monoids by showing that the group of units of such a monoid is a one-relator group, and then reducing the word problem of such a monoid to the word problem of its group of units. Then decidability of the word problem for the special one-relator monoid follows from Magnus's theorem. Similarly, in [69] Zhang proves that the conjugacy problem is decidable in the monoid $\langle A \mid w = 1\rangle$ provided it is decidable in the group of units of the monoid. Other results where an algorithmic problem in a monoid is reduced to the group of units can be found in [46, 70]. These results immediately suggest the following question: *Can the Diophantine problem of a special one-relator monoid be reduced to the Diophantine problem in its group of units?* Notice that by Proposition 3.32, a positive answer to this question would imply that the Diophantine problem is decidable in all special one-relator monoids with torsion. Note that it follows from the main result of [41] that a one-relator monoid of the form $\langle A \mid u = 1\rangle$ has torsion (that is, has a non-identity element of finite order) if and only if $u = w^k$ for some $k \geqslant 2$. Moreover, as we will prove in this paper, a positive answer to this question would imply decidability of WELCs (see Corollary C).

A modern approach to finite special monoid presentations using techniques from the theory of string rewriting systems is given by Zhang in [70]. Zhang's methods will play an important role in the results we prove in this paper for special one-relator monoids.

We shall now explain the main results of the paper in more detail. Before doing so, we first need to give some background notions.

Given any one-relator monoid presentation of the form $\langle A \mid r = 1\rangle$, defining a monoid M, there is a unique decomposition of the word $r \equiv r_1 r_2 \dots r_k$ such that each r_i belongs to $A^+ = A^*\backslash\{1\}$, each of the words r_i represents an invertible element of M, and no proper non-empty prefix of r_i is invertible, for all $1 \leqslant i \leqslant k$. The words r_i $(1 \leqslant i \leqslant k)$ in this decomposition are called the *minimal invertible pieces* of r. Adjan [2] gives an algorithm for computing this decomposition for any one-relator special monoid. Minimal invertible pieces are a key concept for relating a special monoid with its group of units. The key idea used in Adjan's algorithm for computing the minimal invertible pieces is the following fact:

(†) If $\alpha, \beta, \gamma \in A^*$ are words such that $\alpha\beta$ and $\beta\gamma$ both represent invertible elements of the monoid M then all of the words α, β and γ also represent invertible elements of M.

This is because $\alpha\beta$ being invertible implies β is left invertible, while $\beta\gamma$ being invertible implies β is right invertible, hence β is invertible, from which it then quickly also follows that α and γ are also invertible. We say that the words $\alpha\beta$ and $\beta\gamma$ *overlap* in the word β. Adjan's algorithm begins with the defining relator word r from the presentation $\langle A \mid r = 1\rangle$ which clearly represents an invertible element of M (since $r = 1$ in M) and first considers overlaps that r has with itself. If there are overlaps then applying (†) this gives rise to new shorter words that we know are invertible, and then the process is repeated with these words and is iterated until no further overlaps are discovered. We refer the reader to [41, Section 1] for a detailed description of the this overlap algorithm. We will not need full details of the algorithm here, but we will use the key fact (†) above about overlaps when giving examples to which our main results apply. Let us illustrate this now with an example.

Example 1.1. Let M be the one-relator monoid $\langle a, b \mid abcdcdabab = 1\rangle$. Since ab is both a prefix and a suffix of the defining relator word $abcdcdabab$, applying the fact (†) above about overlaps with $\alpha\beta \equiv abcdcdabab \equiv \beta\gamma$ where $\beta \equiv ab$ it it follows that the words $\beta \equiv ab$, $\gamma \equiv cdcdabab$, and $\alpha \equiv abcdcdab$ are all invertible. Then overlapping the invertible word ab with the invertible word $abcdcdab$ it follows from (†) that $cdcdab$ and $abcdcd$ are both invertible. Then overlapping the invertible words $cdcdab$ and $abcdcd$ we deduce that cd is invertible. This shows that this monoid presentation can be written as $\langle a, b \mid (ab)(cd)(cd)(ab)(ab) = 1\rangle$ where the parentheses indicate a decomposition of the defining relator into invertible pieces ab and cd. Moreover, since ab and cd do not overlap with themselves, or each other, the Adjan algorithm will not compute any smaller invertible pieces and hence this is the decomposition of the relator into minimal invertible pieces. In particular $\{ab, cd\}$ is the set of minimal invertible pieces of the relator in this example.

For all the concrete examples of one-relator monoids that we give in this paper, the decomposition of the defining relator into minimal invertible pieces can be computed by repeated application of (†) in exactly the same manner as in Example 1.1. In each case, we shall refer to this as the decomposition into minimal invertible pieces computed by the Adjan overlap algorithm.

Given a set S and a tuple of nonnegative integers $\vec{\lambda} = (\lambda_s \mid s \in S)$, by $|\cdot|_{\vec{\lambda}}$ we denote the $\vec{\lambda}$-*weighted word-length* in S^* defined as

$$|w|_{\vec{\lambda}} =_{\text{def}} \sum_{s\in S} \lambda_s |w|_s, \quad (w \in S^*),$$

where $|w|_s$ denotes the number of occurrences of the letter s in w. By $\mathsf{L}_{\vec{\lambda}}$ we denote the $\vec{\lambda}$-*length relation* defined as $\mathsf{L}_{\vec{\lambda}}(w, u)$ if and only if $|w|_{\vec{\lambda}} \leqslant |u|_{\vec{\lambda}}$. Note that if $\lambda_s = 1$ for all $s \in S$ then $|\cdot|_{\vec{\lambda}}$ and $\mathsf{L}_{\vec{\lambda}}$ are just the standard word length and the standard length relation, which we denote simply as $|\cdot|$ and L, respectively. Hence $\mathsf{L}(u, v)$ holds if and only if $|u| \leqslant |v|$, for any two words $u, v \in S^*$. The tuple $(S^*, \cdot, 1, =, \mathsf{L}_{\vec{\lambda}})$ refers to the free monoid S^* equipped with the relation $\mathsf{L}_{\vec{\lambda}}$. This is the natural structure on which to write systems of word equations with ($\vec{\lambda}$-weighted) length constraints. See Subsection 2.1 for further details.

The main tool we use for reducing one problem to another is that of *interpretability by systems of equations* or by *positive existential formulas* (Definition 2.3). This is nothing more than the usual notion of interpretability [31, 50] restricting all formulas to be systems of equations or disjunctions of systems of equations, respectively.

Among other results, in this paper we prove the following.

Theorem A (Theorems 3.23 and 3.26)**.** *Let M be the one-relator monoid $\langle A \mid r = 1\rangle$. Write $r \equiv r_1 r_2 \dots r_k$ such that $r_i \in A^+$ for all $i = 1, \dots, k$, each of the words r_i represents an invertible element of M, and no proper non-empty prefix of r_i is invertible, for all $1 \leqslant i \leqslant k$. Set $\Delta = \{r_i \mid 1 \leqslant i \leqslant k\}$, so Δ is the set of minimal invertible pieces of the relator r. Suppose that:*

(C1) no word from Δ is a proper subword of any other word from Δ, and

(C2) there exist distinct words $\gamma, \delta \in \Delta$ with a common first letter a.

Then there exists a free monoid D of finite rank $n \geqslant 2$ and a tuple of positive integer weights $\vec{\lambda} = (\lambda_1, \dots, \lambda_n)$ such that the free monoid with weighted length relation $(S, \cdot, 1, \mathsf{L}_{\vec{\lambda}})$ is interpretable in M by systems of equations. Consequently, the problem of solving systems of word equations with weighted length constraints is reducible to the problem of solving systems of equations in M.

If additionally to (C1) and (C2) we have:

(C3) no word in Δ starts with a^2,

then the above result holds with $\mathsf{L}_{\vec{\lambda}}$ being the standard length relation L, i.e. $\mathsf{L}(u, v)$ if and only if $|u| \leqslant |v|$, for $u, v \in D$. Consequently, in this case, the problem of solving systems of word equations with length constraints is reducible to the problem of solving systems of equations in M.

Example 1.2. We now give several examples to which Theorem A applies. For each of these examples the decomposition into minimal invertible pieces can be computed using Adjan overlap algorithm in the same way as in Example 1.1.

Some examples of monoids satisfying conditions (C1), (C2) and (C3) are $\langle a, b, c \mid (ab)(ac)(ab) = 1\rangle$ and $\langle a, b, c \mid ((ab)(ac)(ab))^n = 1\rangle$ for $n \geqslant 1$, where we indicate the minimal invertible pieces with parentheses. In all these examples the set of minimal invertible pieces is $\Delta = \{ab, ac\}$. Indeed, considering overlaps of the defining relator word $((ab)(ac)(ab))^n$ with itself implies that $(ab)(ac)(ab)$ is invertible, and then overlapping this word with itself we deduce that ab and ac are both invertible. Since this pair of words do not overlap with themselves, or each other, it follows that these are the minimal invertible pieces. This set of words $\Delta = \{ab, ac\}$ clearly satisfies conditions (C1), (C2) and (C3).

In the two-generated case we have examples satisfying all of (C1), (C2) and (C3) such as $\langle a, b \mid (ababb)(abaabb)(ababb) = 1\rangle$ and $\langle a, b \mid ((aba^nb^{n+1})(aba^{n+1}b^{n+1})(aba^nb^{n+1}))^m = 1\rangle$, for all $n, m \geqslant 1$, where again we identify the decomposition into minimal invertible pieces using parentheses. For this second family of examples, by overlapping the relator word with itself we deduce that $(aba^nb^{n+1})(aba^{n+1}b^{n+1})(aba^nb^{n+1})$ is invertible and hence overlapping this word with itself we deduce that each of aba^nb^{n+1} and $aba^{n+1}b^{n+1}$ is an invertible word. Since this pair of words do not overlap with themselves or with each other, it follows that the set of minimal invertible pieces for this example is $\Delta = \{aba^nb^{n+1}, aba^{n+1}b^{n+1}\}$. It is then straightforward to verify that this set of words satisfies conditions (C1), (C2) and (C3).

Dropping (C3) there are simpler two-generated examples which satisfy both (C1) and (C2) e.g. $\langle a, b \mid ((aab)(abb)(aab))^n = 1\rangle$ $(n \geqslant 1)$ with set of minimal invertible pieces $\{aab, abb\}$. As seen in these examples, the family of one-relator monoids satisfying conditions (C1), (C2), and (C3) includes many one-relator monoids with torsion $\langle A \mid w^n = 1\rangle$, $n > 1$, which by Proposition 3.32 have hyperbolic group of units and hyperbolic undirected Cayley graph. We stress again that one-relator groups with torsion are hyperbolic and thus have decidable Diophantine problem [16, 60].

In another direction we prove the following result, which can be used to obtain many examples of special one-relator monoids with decidable Diophantine problem, as described in Section 4.

Theorem B (Theorem 3.1). *Let $M = \langle A \mid w = 1\rangle$ and suppose that every letter in w is invertible in M. Let $G = \langle B \mid w = 1\rangle$ where $B \subseteq A$ is the set of letters that appear in w. Then G is a one-relator group, and if the Diophantine problem is decidable in G then it is decidable in M.*

Comparing Theorem B with Theorem A, in both results we decompose the defining relator $r \equiv r_1r_2 \dots r_k$ into words r_i that are invertible in M. Theorem B is the case where all the words r_i have size one, i.e. they are single letters, in which case the theorem shows a reduction of the Diophantine problem of M to its group of units.

In Section 4 we provide some examples of monoids satisfying the hypotheses of Theorem B, as well as a list of questions and open problems.

An immediate consequence of Theorem A is the following

Corollary C. *If word equations with length constraints are undecidable, then so is the Diophantine problem in any one-relator monoid of the form $\langle A \mid w = 1\rangle$ satisfying conditions*

(C1), (C2) and (C3). On the other hand, proving that the Diophantine problem is decidable in some of these monoids would imply that word equations with length constraints are decidable.

In particular, if the Diophantine problem is decidable for all one-relator monoids with torsion $\langle A \mid w^n = 1\rangle$, with $n > 1$, then this would imply that word equations with length constraints are decidable.

In addition to the Diophantine problem, we also obtain results about the decidability of the first-order theory, and more precisely of the positive AE-theory, of some one-relator monoids. The first-order theory with coefficients of a free nonabelian semigroup was shown to be undecidable by Quine [57] (all free structures in this paragraph are implicitly assumed to be nonabelian). Quine's result was strengthened in [24, 49] by proving that the positive AE-theory with coefficients of a free semigroup is undecidable. This contrasts with the aforementioned decidability result of Makanin for systems of equations, and also with the fact that the first-order theory of free groups is decidable as part of the solution to Tarski problems [39]. A consequence of Theorem A is the following

Theorem D (Theorem 3.25)**.** *Let M be a monoid with presentation $\langle A \mid w = 1\rangle$ for some set A and some word $w \in A^*$ satisfying the conditions (C1) and (C2) of Theorem A. Then the positive AE-theory with coefficients of M is undecidable. In particular, the first-order theory with coefficients of M is undecidable.*

The paper is organized as follows: in Section 2.1 we provide all the necessary background regarding equations, first-order theory, and tools for obtaining reductions of one algorithmic problem to another. Section 3 contains the main results of the paper. Section 3 finishes with a small subsection where we obtain results regarding the hyperbolicity of the group of units and of the Cayley graph of some one-relator monoids. Finally, in Section 4 we present examples and applications of our results, and we provide a list of open questions.

2 Preliminaries

In this section we provide the necessary background definitions and results from model and semigroup theory that will be needed in this article. In Subsections 2.1 and 2.2 we shall state the model-theoretic definitions for general structures, although throughout the paper these will be used only on monoids, or on monoids with some extra function or relation such as a length relation. Further background on model theory can be found in [31, 50]. See [5] for notions of computational and complexity theory, [32] for semigroup and monoid theory background, and [45] for notions in combinatorial group theory.

2.1 Equations, first-order theory, and other problems

We follow Sections 1.1. and 1.3 from [31]. We fix X and A to denote a finite set of variables and a finite set of constants, respectively.

We describe structures by tuples $S = (U, f_1, f_2, \dots, r_1, r_2, \dots, c_1, c_2, \dots)$, where U is the domain of the structure, the f_i are function symbols, the r_i are relation symbols, and the c_i are constant symbols. The equality relation $=$ is always assumed to be one of the relations of S and is usually omitted from the list $r_1, \dots$ The tuple $(f_1, f_2, \dots, r_1, r_2, \dots, c_1, c_2, \dots)$ is the language (or signature) of S. We make the convention that this tuple is implicitly enlarged with as many elements from U as needed. These extra elements are called *coefficients* (or *parameters*). Sometimes we identify the whole structure with its domain. For example, we denote the free monoid generated by A simply by A^*, omitting any reference to the concatenation operation $\cdot$ or the identity element 1 or the equality relation $=$.

An *equation* in a structure S with language L is an atomic formula in the language L with coefficients. Recall that an *atomic formula* is one that makes no use of quantifiers, conjunctions, disjunctions, or negations. Thus an equation in S is a formula constructed using only variables, constant elements from U (because we allow the use of coefficients by convention), functions f_i, and a single relation r_i. For example if S is a monoid generated by A then an equation in S is a formal expression of the form $w_1(X, A) = w_2(X, A)$, where $w_1(X, A)$ and $w_2(X, A)$ are words in $(A \cup X)^*$. A *solution* to such equation is a map $f : X \to S$ such that $w_1(f(X), A) = w_2(f(X), A)$ is true in S. By $w_i(f(X), A)$ we refer to the word obtained from w_i after replacing each variable $x \in X$ by the word $f(x)$. A *system of equations* in S is a conjunction of equations in S. Alternatively one can define equations as formulas of the form $\exists x_1 \dots \exists x_n \phi(x_1, \dots, x_n)$ where ϕ is an atomic formula as above on variables $x_1, \dots, x_n$, with the relation in ϕ being equality. We use these two formulations interchangeably.

Equations in a free monoid A^* receive the special name of *word equations.* One can consider equations in more complicated structures, such as the structure $(A^*, \cdot, 1, =, \mathsf{L})$ obtained from the free monoid A^* (which we identify with the tuple $(A^*, \cdot, 1, =)$) by adding the length relation L defined by the rule $\mathsf{L}(u, v)$ if and only if $|u| \leqslant |v|$, for all $u, v \in A^*$, where $|\cdot|$ denotes length of words and 1 is the identity element. A system of equations in $(A^*, \cdot, 1, =, \mathsf{L})$ is called a system of *word equations with length constraints.* This is a system of word equations Σ together with a finite conjunction $\mathcal{C}$ of formal expressions of the form $\mathsf{L}(w_1, w_2)$, each called a *length constraint*, where $w_1, w_2 \in (X \cup A)^*$. A map $f : X \to A^*$ is a solution to such system if it is a solution to Σ and $|w_1(f(X), A)| \leqslant |w_2(f(X), A)|$ for each length constraint $\mathsf{L}(w_1, w_2)$ appearing in $\mathcal{C}$.

Alternatively to the length constraint one can consider the more general notion of *weighted length constraint*, which we define now. Let $\vec{k} = (k_a \mid a \in A)$ be a tuple of natural numbers, one for each constant $a \in A$. Then by $|\cdot|_{\vec{k}}$ we denote the map $|\cdot|_{\vec{k}} : A^* \to \mathbb{N}$ defined by

$$|h|_{\vec{k}} = \sum_{a \in A} k_s n_a(a),$$

where $n_a(h)$ is the number of times that the letter s appears in h. We call $|\cdot|_{\vec{k}}$ the *$\vec{k}$-weighted* length function of A^*. We further let $\mathsf{L}_{\vec{k}}$ denote the relation in A^* defined by the rule $\mathsf{L}_{\vec{k}}(h, g)$ if and only if $|h|_{\vec{k}} \leqslant |g|_{\vec{k}}$, and call $\mathsf{L}_{\vec{k}}$ the *$\vec{k}$-weighted length relation* in A^*. Note that if $\vec{k}$ consists solely of 1's then $|\cdot|_{\vec{k}}$ is the usual length of words $|\cdot|$ and $\mathsf{L}_{\vec{k}}$ is the length relation L.

The *Diophantine problem* in a structure S, denoted $\mathcal{D}(S)$, refers to the algorithmic problem of determining if a given system of equations in S (with coefficients belonging to a fixed computable set) has a solution. One says that $\mathcal{D}(S)$ is *decidable* if there exists an algorithm (i.e. a Turing machine [5]) that performs such task.

Given two algorithmic problems P_1 and P_2, we say that P_1 is *reducible* to P_2 if there exists an algorithm that solves P_1 using an oracle for the problem P_2 (i.e. a black-box algorithm that 'magically' solves P_2 —see Definition 3.4 in [5]). Thus in this case if P_1 is unsolvable then so is P_2: indeed, if P_2 was solvable then replacing the oracle in the definition above by an algorithm that solves P_2 would yield an algorithm that solves P_1, a contradiction. As an example, $\mathcal{D}(\mathbb{Z})$ is undecidable for $\mathbb{Z}$ the ring of integers (this is the answer to Hilbert's 10th problem [51]), and hence $\mathcal{D}(M)$ is undecidable for any structure M such that $\mathcal{D}(\mathbb{Z})$ is reducible to $\mathcal{D}(M)$.

Let L be some language. A *positive AE-sentence* in L is a first-order sentence of the form

$$\forall x_1 \dots \forall x_n \exists y_1 \dots \exists y_m \psi(x_1, \dots, x_n, y_1, \dots, y_m)$$

where ψ is a quantifier-free formula without negations on the language L. The positive *AE-theory* of a structure S is the set of all positive AE-sentences in the language of S that are true in S. Analogously to the Diophantine problem, the positive AE-theory of S is said to be *decidable* if there exists and algorithm that, given a positive AE-sentence, decides whether or not it holds in S.

One can generalize the notions in the paragraph above by replacing positive AE-sentences by any family of first-order sentences Φ. In particular, if Φ is the set of all first-order sentences then one speaks of the *first-order theory*, or the *elementary theory*, of a structure. It is important to note that if the first-order theory is decidable then so is the Diophantine problem, the positive AE-theory, the positive universal theory (identity checking), etc.

2.2 Reductions and interpretability

In this subsection we introduce the notion of interpretability with respect to some class of formulas. This is a powerful tool which, in particular, implies reducibility of the decision problem for such class of formulas. It is nothing else than the classical model-theoretical notion of interpretability [31, 50], with the modification that formulas are required to be of some specific form (such as systems of equations). We follow Section 1.3 of [50] (alternatively, see Sections 2.1 and 5.3 of [31]).

Definition 2.1. Let M be a structure, n a natural number, and Φ a set of formulas in the language of M. A subset $S \subseteq M^n$ is called *definable* in M *by formulas in* Φ (in short, Φ-definable) if there exists a formula

$$\Sigma_S(x_1, \dots, x_n, y_1, \dots, y_k) \in \Phi,$$

with free variables $(x_1, \dots, x_n, y_1, \dots, y_k) = (\vec{x}, \vec{y})$, such that for any $\vec{m} \in M^n$, one has that $\vec{m} \in S$ if and only if there exists $\vec{y}_0 \in M^k$ such that $\Sigma_S(\vec{m}, \mathbf{y}_0)$ is true in M. In this case Σ_S is said to *define* S in M.

We will make use of the following two classes of formulas Φ:

1. Systems of equations. In this case we replace the prefix $\Phi-$ by e-, speaking of *e-definability.*

2. Disjunctions of systems of equations. In this case we speak of *PE-definability.* See below for an explanation of this terminology.

Remark 2.2. It is well known that any disjunction of systems of equations is equivalent to a positive existential sentence with coefficients (hence the name PE-definability), i.e. formulas that can be constructed using only existential quantifiers, conjunctions, disjunctions, variables, and coefficents from the structure.

For example, the set of all elements that commute with a given element $m \in M$ is defined by the equation $xm = mx$. Likewise, the set of all elements of M that are squares is defined by the equation $x = y^2$. The set of all elements that commute with m or are a square is defined by the PE-formula $(xm = mx) \vee (x = y^2)$.

Observe that, by definition, e-interpretability and PE-interpretability allow the use of any coefficients in the domain of the structures at hand.

Definition 2.3. Let $\mathcal{A}$ and $\mathcal{M}$ be two structures and let Φ be a family of formulas in the language of $\mathcal{M}$. Let further A and M be the domains of $\mathcal{A}$ and of $\mathcal{M}$, respectively. Then $\mathcal{A}$ is called *interpretable* in $\mathcal{M}$ by formulas Φ (in short, Φ-interpretable) if there exists $n \in \mathbb{N}$, a subset $S \subseteq M^n$ and a bijective[1] map, called *interpreting* map, $\phi : S \to A$, such that:

1. S is Φ-definable in $\mathcal{M}$.

2. For every function $f = f(x_1, \ldots, x_n)$ in the language of $\mathcal{A}$, the preimage by ϕ of the graph of f, i.e. the set $\{(s_1, \ldots, s_k, s_{k+1}) \in S^{k+1} \mid \phi(s_{k+1}) = f(\phi(s_1), \ldots, \phi(s_k))\} \subseteq M^{n(k+1)}$, is Φ-definable in $\mathcal{M}$.

3. Similarly, for every relation r of $\mathcal{A}$ (including the equality relation $=$), the preimage by ϕ of the graph of r is Φ-definable in $\mathcal{M}$.

Similarly as before, if Φ consists of all systems of equations in the language of $\mathcal{M}$ then we speak of e-interpretability, and if Φ consists of all disjunctions of systems of equations we speak of PE-interpretability. Note that a PE-interpretation is, in particular, an e-interpretation.

The next two results are fundamental and they constitute the main reason we use interpretability in this paper. These are standard results whose proofs follow immediately from the Reduction Theorem 5.3.2 in [31] and Remark 3 after it (alternatively, see Lemma 2.7 of [27]).

Proposition 2.4 (Interpretability is transitive)**.** *Interpretability is a transitive relation. That is, given three structures M_1, M_2, and M_3, if M_1 is e- or PE-interpretable in M_2 and M_2 is e or PE-interpretable in M_3, then M_1 is e- or PE-interpretable in M_3, respectively.*

Proposition 2.5 (Reduction of problems)**.** *Let M_1 and M_2 be two structures on languages L_1 and L_2, respectively. Assume M_1 is e-interpretable or PE-interpretable in M_2. Then the Diophantine problem in M_1 is reducible to the Diophantine problem in M_2. As a consequence, if the second problem is decidable, then so is the first.*

Similarly, the problem of deciding if any given first-order formula in the language L_1 holds in M_1 is reducible to the problem of deciding if any given formula in the language L_2 holds in M_2. Consequently, if the first-order theory of M_2 is decidable then so is the first-order theory of M_1. The same statement holds when replacing first-order theory by positive AE-theory.

2.3 Monoid presentations and rewriting systems

Let A be a non-empty alphabet. In this paper we will use $\equiv$ to denote graphical equality, that is, for two words $w_1, w_2 \in A^*$, the expression $w_1 \equiv w_2$ means w_1 and w_2 are equal as word in

[1] The most general formulation of interpretability uses onto maps instead of bijective maps. Since only bijective maps appear in the interpretations of this paper, we have chosen to use this more restricted version of interpretability. This is similar to the approach followed in Section 1.3 of [50]. For the definition of interpretability with onto maps see Section 5.4 of [31] or Section 1.3 of [50].

A^*. A *rewriting system* R over A is a subset of $A^* \times A^*$. We call $\langle A \,|\, R\rangle$ a *monoid presentation.* This monoid presentation is said to be finite if both A and R are finite, and infinite otherwise. The elements of R are called *rewrite rules* of the rewriting system, and they are called the *defining relations* of the presentation. A rewrite rule $(u, v) \in R$ is often written as $u = v$ when writing the presentation $\langle A \mid R\rangle$. For $u, v \in A^*$ we write $u \rightarrow_R v$ if there are words $\alpha, \beta \in A^*$ and a rewrite rule (l, r) in R such that $u \equiv \alpha l \beta$ and $v \equiv \alpha r \beta$. Let $\rightarrow_R^*$ denote the reflexive transitive closure of $\rightarrow_R$, and let $\leftrightarrow_R^*$ denote the reflexive transitive symmetric closure of $\rightarrow_R$. The monoid defined by the presentation $\langle A \mid R\rangle$ is the set $A^*/\leftrightarrow_R^*$ of equivalence classes of the equivalence relation $\leftrightarrow_R^*$ with multiplication defined by $(w_1/\leftrightarrow_R^*) \cdot (w_2/\leftrightarrow_R^*) = w_1 w_2/\leftrightarrow_R^*$ for all $w_1, w_2 \in A^*$. When the set of rewrite rules is clear from context, we shall omit the subscript R and simply write $\rightarrow$, $\rightarrow^*$ and $\leftrightarrow^*$. A word u is called *reduced* if no rewrite rule can be applied to it, that is, there is no word v with $u \rightarrow v$. A rewriting system R is called *Noetherian* if there is no infinite chain of words $u_i \in A^*$ with $u_i \rightarrow u_{i+1}$ for all $i \geqslant 1$. The rewriting system system is called *confluent* if whenever $u \rightarrow^* u_1$ and $u \rightarrow^* u_2$ there is a word $v \in A^*$ such that $u_1 \rightarrow^* v$ and $u_2 \rightarrow^* v$. A *complete rewriting system* is one that is both Noetherian and confluent. If R is a complete rewriting system then each $\leftrightarrow^*$-class contains a unique reduced word. It follows that if R is a complete rewriting system over an alphabet A then the set of reduced words of this system provides a set normal forms (that is, unique representatives) for the elements of the monoid M defined by the presentation $\langle A \mid R\rangle$. In this situation we call $\langle A \mid R\rangle$ a *complete presentation defining the monoid M*. We say that a word is *reduced with respect the complete presentation $\langle A \mid R\rangle$* if it is a reduced word with respect to the complete rewriting system R. If in addition either A or R is infinite then this is called an *infinite complete presentation.*

3 One-relator monoids

Our interest in this section is in the Diophantine problem for one-relator monoids with presentation $\langle A \mid w = 1\rangle$. Throughout this section M will denote the one-relator monoid defined by the one-relator presentation $\langle A \,|\, w = 1\rangle$. We shall see how this problem relates to other known difficult decidability problems. Before exploring those links we first observe one situation where the Diophantine problem is decidable.

Theorem 3.1. *Let $M = \langle A \,|\, w = 1\rangle$ and suppose that every letter in w is invertible in M. Let $G = \langle B \mid w = 1\rangle$ where $B \subseteq A$ is the set of letters that appear in w. Then G is a one-relator group, and if the Diophantine problem is decidable in G then it is decidable in M.*

Proof. The monoid M is isomorphic to the moniod free product $G * C^*$ where $C = A \backslash B$. Both G and C^* satisfy Assumption 17 from [20] (a cancellativity condition which satisfied by any group and any free monoid) and Assumption 18 from [20] (decidability of the Diophantine problem). Hence applying [20, Theorem 19] (taking $\mathcal{C}_\sigma$ to be just $\{U_\sigma, V_\sigma\}$) we obtain that the Diophantine problem of M is decidable. □

Example 3.2. As an easy example of an application of the previous theorem, we see that the Diophantine problem is decidable in the monoid $M = \langle a, b, c, d \mid aba = 1\rangle$. Indeed, the monoid $G = \langle a, b \mid aba = 1\rangle$ is the infinite cyclic group. To see this, since $a(ba) = 1$ it follows that a is right invertible in the monoid with right inverse ba, and since $(ab)a = 1$ it follows that a is left invertible with left inverse ab. Hence a is invertible in the monoid G with inverse $ab = ba$ (because $ab = ab(aba) = (aba)ba = ba$). Letting a^{-1} denote the inverse of a in this monoid we have $b = a^{-2}$. Hence G is the infinite cyclic group generated by a. The argument above shows that each letter that appears in aba is invertible in M, and the group $G = \langle a, b \mid aba = 1\rangle$ has decidable Diophantine problem (since all free groups do). Hence the hypotheses of Theorem 3.1 are satisfied and we conclude that the Diophantine problem is decidable in the monoid M.

Some more complicated examples to which Theorem 3.1 applies will be discussed in Section 4.

The following lemma will be key later when studying systems of equations in some one-relator monoids (Section 3). A definition of weighted length relation can be found in Subsection 2.1.

Lemma 3.3. *Let M be a monoid, let $C = \langle c_0\rangle$ be an infinite one-generated submonoid of M, and let D be a free rank-n submonoid of M freely generated by a set $\{d_1, \ldots, d_n\} \subseteq M$. Assume that both monoids C and D are e-interpretable in M with interpreting map the identity map. Assume also that for each $i = 1, \ldots, n$ there exists $k_i \in \mathbb{N}$ such that $c^{k_i} d_i = 1$. Then the free monoid with weighted length relation $(D, \cdot, 1, =, \mathsf{L}_{\vec{k}})$ is e-interpretable in M, where $\vec{k} = (k_1, \ldots, k_n)$, and $\cdot$ is the usual concatenation operation.*

Proof. Since the free monoid D is e-interpretable in M, it suffices to show that so is the relation $\mathsf{L}_{\vec{k}}$. Let $\Sigma_C(x, \vec{y})$ and $\Sigma_D(z, \vec{w})$ be two systems of equations e-interpreting C and D in M, so that an element $h \in M$ belongs to C (respectively D) if and only if $\Sigma_C(h, \vec{y})$ (resp. $\Sigma_D(h, \vec{w})$) has a solution $\vec{y}_0$ (resp. $\vec{w}_0$) in M. Take arbitrary elements $c \in C$ and $d \in D$. Then $c = c_0^t$ for some $t \in \mathbb{N}$, and $d = d_{i_1} \ldots d_{i_r}$ for some d_{i_j}. Now,

$$cd = \begin{cases} c_0^{t-|d|_{\vec{k}}} & \text{if } t > |d|_{\vec{k}}, \\ 1 & \text{if } t = |d|_{\vec{k}}, \\ d_{i_{\ell+1}} \ldots d_{\ell_r} & \text{if } t < |d|_{\vec{k}}, \text{ and } cd \in D, \\ c_0^s d_{i_{\ell+1}} \ldots d_{\ell_r} & \text{if } t < |d|_{\vec{k}}, \text{ and } cd \notin D, \end{cases} \tag{1}$$

where in the last two cases ℓ is the minimum number such that $|d_{i_1} \ldots d_{i_{\ell+1}}|_{\vec{k}} > t$ (we have $\ell < r$), and in the last case s is some number such that $0 < s < k_{i_{\ell+1}}$. It follows that if $t \geqslant |d|_{\vec{k}}$ then $cd \in C$. The other implication is true as well: if we had $cd \in C$ and $t < |d|_{\vec{k}}$ then $cd = c_0^s d_{i_{\ell+1}} \ldots d_{\ell_r} = c_0^r$ for some $r, s \geqslant 0$ and some $0 < \ell < r$. Let $d' = d_{i_{\ell+1}} \ldots d_{\ell_r}$ and let $p = |d'|_{\vec{k}}$. Note that $s < p$. Then $1 = c_0^p d' = c_0^{p-s} c_0^s d' = c_0^{p-s+r}$, contradicting the assumption that $\langle c_0\rangle$ is infinite.

We have proved that $cd \in C$ if and only if $|c| = t \geqslant |d|_{\vec{k}}$. Due to the e-definability of C, this in turn occurs if and only if $\Sigma_C(cd, \vec{y})$ has a solution $\vec{y}_0$. Moreover, the second case of (1) and the infiniteness of $\langle c_0\rangle$ indicate that $t = |d|_{\vec{k}}$ if and only if $cd = 1$. Hence given two elements $d_1, d_2 \in D$ we have that $|d_1|_{\vec{k}} \leqslant |d_2|_{\vec{k}}$ if and only if there exists an element $c \in C$ such that $cd_2 = 1$ (this ensures $|c| = |d_2|_{\vec{k}}$) and $\Sigma_C(cd_1, \vec{y})$ has a solution $\vec{y}_0$ (this ensures $|d_1|_{\vec{k}} \leqslant |c|$).

Overall, $|d_1|_{\vec{k}} \leqslant |d_2|_{\vec{k}}$ if and only if the following system of equations has a solution $x_0, \vec{y}_0, \vec{z}_0$:

$$\begin{cases} \Sigma_C(x, \vec{y}), \\ \Sigma_C(xd_1, \vec{z}), \\ xd_2 = 1 \end{cases} \tag{2}$$

It follows that the $\vec{k}$-weighted length relation $\mathsf{L}_{\vec{k}}$ is e-interpretable in M. □

Example 3.4. The above result can be applied to the monoid with presentation

$$\langle a, b_1, \ldots, b_n \mid ab_1 = 1, ab_2 = 1, \ldots, ab_n = 1 \rangle, \tag{3}$$

for any $n > 1$, thus we recover the reduction from Example 21 in [20].

Let $\Delta = \{\alpha_i \ (i \in I)\} \subseteq A^+$ be the set of minimal invertible pieces of the defining relator w. So the word w uniquely decomposes as

$$w \equiv \alpha_{i_1}\alpha_{i_2} \ldots \alpha_{i_k}$$

where each $\alpha_{i_j} \in \Delta$, and each of these words is invertible in M and has no proper non-empty prefix which is invertible in M. As mentioned in the introduction, we call the α_{i_j} the minimal invertible pieces of w. In [2] Adjan gives an algorithm for computing the minimal invertible pieces of the defining relator of a one-relator special monoid. In particular, every letter appearing in the relator represents an invertible element of the monoid if and only if all the minimal invertible pieces have size one, and this can be decided using Adjan's algorithm. Hence Adjan's algorithm can be used to test whether the hypothesis of Theorem 3.1 above are satisfied. This algorithm was discussed in Section 1, see Example 1.1 and the paragraph preceding it. As mentioned there, a good description of Adjan's algorithm can be found in [41, Section 1].

Since each piece α_i is minimal invertible, none of them is a prefix of another piece α_j, and so Δ is a prefix code. Hence the submonoid of A^* generated by Δ is free. We shall denote it Δ^*. Let $B = \{b_i \mid i \in I\}$ be an alphabet in bijective correspondence with Δ. Let $\phi : \Delta^* \to B^*$ be the unique homomorphism extending $\alpha_i \mapsto b_i$ for $i \in I$. It follows from Adjan's results [2] that the group of units G of M is isomorphic to the monoid defined by the monoid presentation

$$\langle B \mid \phi(w) = 1 \rangle = \langle B \mid b_{i_1}b_{i_2} \ldots b_{i_k} = 1 \rangle.$$

Theorem 3.5 ([70], Proposition 3.2)**.** *The infinite monoid presentation*

$$\langle A \mid u = v : u, v \in \Delta^*, v <_{sh} u \quad \& \quad \phi(u) =_G \phi(v) \rangle \tag{4}$$

is an infinite complete presentation defining the monoid M.

In the above theorem $\leqslant_{\text{sh}}$ denotes shortlex ordering, and $\phi(u) =_G \phi(v)$ means that $\phi(u)$ and $\phi(v)$ both represent the same element in the group of units G. For the rest of this section, when we say a word w is *reduced* we mean that it is reduced with respect to the above infinite complete presentation (4). Our aim is to show that for a wide class of special one-relator monoids, if we could solve equations for those monoids then that would imply a solution to equation solving with length constraints in free monoids —which is a longstanding open problem; see [13, 17, 26, 42]. Of course, not every special one-relator monoid encodes equation solving with length constraints since, for instance, we have seen above that equations can be solved over the bicyclic monoid. So we will need some conditions on the monoid. We give conditions in terms of certain combinatorial properties on the set of minimal invertible pieces Δ. We suppose that the following conditions are satisfied:

(C1) No word from Δ is a proper subword of any other word from Δ.

(C2) There exist distinct words $\gamma, \delta \in \Delta$ with a common initial letter $a \in A$.

These conditions are easily satisfied and can be used to construct a wide variety of examples as we shall see in the next section. Note, for instance, if all the words from Δ have the same length, then condition (C1) will be satisfied. In particular there are one-relator monoids with torsion whose minimal invertible pieces satisfy these properties. A concrete example is given by the family monoids

$$\langle a, b, c \mid ((ab)(ac)(ab))^k = 1\rangle,$$

for $k > 1$ where, as we already proved in Example 1.2 above, the set of minimal invertible pieces is $\{ab, ac\}$. This gives many examples of special one-relator monoids with hyperbolic undirected Cayley graphs which satisfy the conditions (C1)-(C2). Applications to examples like this will be discussed below.

> *For the rest of this section let M be the one-relator monoid defined by the monoid presentation*
>
> $$\langle A \mid r = 1\rangle$$
>
> *where we suppose that conditions (C1)-(C2) are satisfied, and we let a be a common initial letter of two distinct words from Δ*

Throughout the rest of the section we denote the projection of a word $w \in A^*$ onto M by $[w]$, so $[w]$ is the element of M represented by the word w.

We now give a series of important technical lemmas. We begin with the following observation:

Remark 3.6. The letter a us not invertible in M. Indeed, there are two distinct words $\delta, \gamma \in \Delta$ having a as their first letter. Since δ and γ are subwords of w which are invertible in M, and they minimal with this property, if a was invertible then we would have $a = \gamma = \delta$. This would contradict the fact that γ and δ are distinct words. Hence, a cannot be invertible.

Lemma 3.7. *Suppose that (C1) and (C2) are both satisfied, and let a be a common initial letter of two distinct words from Δ. Then for every reduced word $w \in A^*$, and every positive integer $i > 0$, if $a^i w = 1$ then w has no prefix in Δ.*

Proof. Since $a^i w = 1$ it follows that $a^i w$ is not reduced and since w is assumed to be reduced it follows that we can write

$$a^i w \equiv a^j \alpha_1 \dots \alpha_t w''$$

where $0 \leqslant j < i$, w'' is a suffix of w, $\alpha_1 \dots \alpha_k$ is the left hand side of a rewrite rule from (4), and each $\alpha_i \in \Delta$. Since a is not invertible and w is reduced, we have $\alpha_1 \equiv a^k w'$ where

$k = i - j > 0$, and w' is a non-empty prefix of w. Suppose, seeking a contradiction, that $w \equiv \beta w_2$ with $\beta \in \Delta$. Note that since w' is a suffix of α_1, where α_1 is invertible, it follows that w' is left invertible. Now, if w' were a prefix of β it would follow that w' is also right invertible and hence invertible. But then since α_1 and w' are both invertible it would follow that a^k is invertible and hence a is invertible, which is a contradiction by Remark 3.6. Therefore we must have that β is a prefix of w', but then $\beta \in \Delta$ is a proper subword of $\alpha_1 \equiv a^k w' \in \Delta$, and this contradicts (C1). This completes the proof of the lemma. □

Lemma 3.8 ([70], Lemma 3.1 and Lemma 3.6)**.** *If $u_1, u_2 \in \Delta^*$ then $[u_1] = [u_2]$ in M if and only if $[\phi(u_1)] = [\phi(u_2)]$ in the one-relator group G.*

Lemma 3.9. *Let δ and γ be two distinct words in Δ. Then $[\delta] \neq [\gamma]$ in M.*

Proof. Since the words δ and γ are distinct it follows that $\phi(\delta)$ and $\phi(\gamma)$ are distinct letters of B. This implies that $|B| \geqslant 2$. If $|B| \geqslant 3$ then it follows from Magnus' Freiheitssatz [45, Theorem 5.1] that $\phi(\delta)$ and $\phi(\gamma)$ represent distinct elements of the group G and hence $[\delta] \neq [\gamma]$ in M, by the previous Lemma 3.8.

Now suppose that $|B| = 2$. Set $c = \phi(\delta)$ and $d = \phi(\gamma)$. If $c = d$ in G then $cd^{-1} = 1$ in G. Since $|B| = 2$ and $c = d$ it follows that the group G has torsion, and hence by [45, Theorem 5.2] the defining relator in the presentation of G must be a proper power. Then it follows from Newman's spelling theorem [45, Theorem 5.5] that cd^{-1} contains a subword of the defining relator (which uses no inverse of c or d), or the inverse of such a subword, with length at least 2. This is clearly impossible and thus completes the proof. □

Lemma 3.10 ([70], Lemma 3.3)**.** *Let $u \in A^*$ be reduced. If $[u]$ is invertible then $u \in \Delta^*$.*

We are interested in right inverses of powers of the element a. These elements clearly form a submonoid of M. The following result shows that the set of reduced words representing elements in this submonoid themselves form a submonoid of the free monoid A^*.

Lemma 3.11. *Let $i, j \in \mathbb{N}$. Let $u, v \in A^*$ be reduced words such that $a^i u = 1$ and $a^j v = 1$. Then uv is a reduced word such that $a^{i+j} uv = 1$.*

Proof. We just need to prove that uv is a reduced word. By Lemma 3.7 the word v does not have any prefix in Δ. If uv were reducible then it would follow that there is a non-empty suffix u_1 of u, and a non-empty prefix v_1 of v, such that $u_1 v_1 \in \Delta$. But then u_1 is left invertible, since u is left invertible, and right invertible, since $u_1 v_1$ is right invertible. This contradicts $u_1 v_1 \in \Delta$. □

Let F be the set of all reduced words β such that $a^i \beta = 1$ for some $i \in \mathbb{N}$ with $i > 0$, together with the empty word 1.

Remark 3.12. Note that by definition all the words in $F \subseteq A^*$ are reduced words with respect to the complete presentation for M defined in Theorem 3.5. It follows from Lemma 3.11 that for any words w_1, w_2 from the set F the concatenation of these two words $w_1 w_2$ is again a word in the set F and hence in particular $w_1 w_2$ is again a reduced word (since all the words in F are reduced words). Therefore, F is a submonoid of the free monoid A^*, and all of the words in F are reduced words with respect to the complete presentation for M.

We shall now prove that F is a free submonoid of A^*. For this it will be useful to recall some standard results about submonoids of free monoids. Recall from [44, Subsection 1.2] that given a submonoid P of A^* there is a unique set $\mathcal{B}$ that generates P and is minimal with respect to set-theoretic inclusion; it is the set

$$(P\backslash\{1\})\backslash(P\backslash\{1\})^2.$$

The following nice characterisation of free subsemigroups of free semigroups, from Lothaire, will be useful for us; see [44, Proposition 1.2.3].

Lemma 3.13. *A submonoid P of A^* is free if and only if for any word $w \in A^*$, one has $w \in P$ whenever there exist $p, q \in P$ such that*

$$pw, wq \in P.$$

Lemma 3.14. *F is a free submonoid of A^*.*

Proof. Suppose that $w \in A^*$ is such that there exist $p, q \in F$ such that $pw, wq \in F$. By Lemma 3.13, we need to show that $w \in F$. Since w is a subword of a reduced word (for example, pw), it is reduced. By assumption there are $i, j \geqslant 1$ such that $a^i p = 1$ and $a^j pw = 1$. If $i = j$ then $w = 1$ and since w is reduced it is the empty word and this belongs to F. If $i < j$ then $a^{j-i}w = 1$ and so $w \in F$. Otherwise, if $i > j$ then it would follow that $w = a^k$ for some $k > 0$. But then $a^j pw = 1$ implies $a^j pa^k = 1$. This last equality implies that a is invertible, contradicting (C2) and the definition of Δ. In all cases $w \in F$ so this completes the proof of the lemma. □

Lemma 3.15. *Let $w \in A^*$ be arbitrary. Write $w \equiv w_1 w_2$ where w_1 is the longest prefix of w which is invertible. Suppose that w' may be obtained from w by a single application of a relation from the presentation. Write $w' \equiv w_1' w_2'$ where w_1' is the longest invertible prefix of w'. Then $w_1 = w_1'$ in M. This implies that for any pair of words u, v, if $u = v$ in M then the longest invertible prefix of u is equal to 1 in M if and only if the longest invertible prefix of v is equal to 1 in M.*

Proof. We consider where the relation is applied to the word $w \equiv w_1 w_2$. If the relation is applied within either w_1 or w_2 the result is immediate, so suppose otherwise. Let $\delta_1 \ldots \delta_m \in \Delta^*$ be the subword of w to which the relation is being applied. If there is a non-empty suffix u_1 of w_1, and a non-empty prefix u_2 of w_2 such that $u_1 u_2 \equiv \delta_r$ for some r, then since u_1 is left invertible since it is a suffix of w_1, and u_1 is right invertible since it is a prefix of δ_r, it would follow that u_1 is invertible, which would contradict the fact that δ_r has no proper prefix which is invertible. So we must have $w_1 \equiv \alpha\delta_1 \ldots \delta_r$, and $w_2 \equiv \delta_{r+1} \ldots \delta_m \beta$, but then $w_1 \delta_1 \ldots \delta_m$ is a prefix of w which is invertible and is longer than w_1, contradicting the definition of w_1. □

Let $m \in \mathbb{N}$ be the maximum value m such that there is a minimal invertible piece $\alpha \in \Delta$ such that a^m is a prefix of α. We define a finite set of words X in the following way. For each $1 \leqslant j \leqslant m$ and for every piece $a^j\beta \in \Delta$ (where β might begin with a) let η be the reduced word representing the inverse of $a^j\beta$ and add the word $\beta\eta$ to the set X.

Lemma 3.16. *Every word in the set X is reduced.*

Proof. Let $a^j\beta \in \Delta$ and let η be a reduced word representing the inverse of $a^j\beta$. We claim that $\beta\eta$ is a reduced word as a consequence of assumption (C1). Indeed, suppose for a contradiction that $\beta\eta$ is not reduced. It follows from Lemma 3.10 that $\eta \in \Delta^*$. Then there is a rewrite rule from (4) which can be applied to the word $\beta\eta$. Let λ be the left hand side of such a rule noting that $\lambda \in \Delta^+$. Since β and η are both reduced words we can write $\lambda \equiv \beta_2\eta_1$ where β_2 and η_1 are both non-empty, with $\beta \equiv \beta_1\beta_2$ and $\eta \equiv \eta_1\eta_2$. Let $\alpha_1 \in \Delta$ be the prefix of λ which belongs to Δ. Let $\alpha_2 \in \Delta$ be the prefix of η which belongs to Δ. Since α_1 cannot be a subword of β since by (C1) it is not a subword of $a^j\beta \in \Delta$ it follows that $\alpha_1 \equiv \alpha_1'\alpha_1''$ where α_1'' is a non-empty prefix of η. But since η is invertible this would imply that α_1'' is invertible and thus α_1' is invertible, contradicting the fact that $\alpha_1 \in \Delta$ is a minimal invertible piece. This is a contradiction, and we conclude that $\beta\eta$ is indeed a reduced word. □

Thus X is a finite set of reduced words, each of which is the right inverse of some a^j with $1 \leqslant j \leqslant m$. Note also that X is a finite subset of the free monoid F.

Lemma 3.17. *Let $i \in \mathbb{N}$ and $w \in A^*$ be a reduced word such that $a^iw = 1$ in M. Then there is an integer $0 < j \leqslant i$, with $j \leqslant m$, and a non-empty prefix w_1 of w such that $w_1 \in X$ and $a^jw_1 = 1$ in M. Moreover, with the same value of j, there is a decomposition*

$$a^iw \equiv a^ka^jw'w''$$

where $k + j = i$, $w \equiv w'w''$ and $a^jw' \in \Delta$. In particular, if no word in Δ begins with a^2 then w can be written as $w \equiv w_1w_2\dots w_i$ such that $aw_l = 1$ for all $1 \leqslant l \leqslant i$.

Proof. Let $i \in \mathbb{N}$ and $w \in A^*$ be a reduced word such that $a^iw = 1$ in M. Since a^iw is not reduced it follows that the left hand side λ of one of the relations from (4) arises as a subword of a^iw. In particular λ is a non-empty word with $\lambda \in \Delta^*$. Since a is not invertible, no word from Δ is a subword of a^i, and since w is reduced, λ is not a subword of w. It follows that there is a prefix λ' of λ such that, $\lambda' \equiv a^jw' \in \Delta$ with $j > 0$ and where w' is a non-empty prefix of w. Thus we have the decomposition

$$a^iw \equiv a^ka^jw'w''$$

where $k + j = i$, $w \equiv w'w''$ and $a^jw' \in \Delta$.

If $k = 0$ then $i = j$ and $a^iw \equiv a^jw = 1$. So we can write $a^jw \equiv (a^jw')(w'')$ and since $(a^jw')(w'') = 1$ it follows that in M we have $w \equiv \beta\eta$ where $\beta \equiv w'$, $\eta \equiv w''$, where η is equal to the inverse of $a^j\beta$ in M (note $a^j\beta$ is invertible because it belongs to Δ). Thus in this case the reduced word w belongs to the set X, as required.

Now suppose that $k > 0$. Consider the longest invertible prefix of the word a^jw. It is certainly non-empty since a^jw' is invertible. Set $v \equiv \mathrm{red}(a^jw)$. Then we have $a^kv = 1$ with $k > 0$ and v a reduced word. It follows from Lemma 3.7 that v cannot begin with a word from Δ. Hence v has no invertible prefix. Now by the last part of Lemma 3.15, since $v = a^jw$ in M, it follows that the longest invertible prefix p of a^jw is equal to 1 in M. So now we can write

$$a^iw \equiv a^ka^jw_1w_2$$

where $k + j = i$, $w \equiv w_1w_2$ and $p \equiv a^jw_1 = 1$ in M, and a^jw_1 has prefix $a^jw' \in \Delta$. It then follows that in M we have $w_1 = \beta\eta$ where $\beta \equiv w'$ and η is equal to the inverse of a^jw' in M. Also, w_1 is a reduced word because w is reduced. It follows that $w_1 \in X$, as required. This completes the proof of the lemma. □

Lemma 3.18. *X is a finite generating set for the monoid F.*

Proof. Let $i \in \mathbb{N}$ and $w \in A^*$ be a reduced word such that $a^i w = 1$ in M. It follows from Lemma 3.17 that there is an integer $0 < j \leqslant i$, with $j \leqslant m$, and a non-empty prefix w_1 of w such that $w_1 \in X$ and $a^j w_1 = 1$ in M. The lemma now follows by induction. □

Let $\mathcal{B}$ be the unique subset of F that generates F and is minimal with respect to set-theoretic inclusion, that is $\mathcal{B}$ is equal to the set

$$(F\backslash\{1\})\backslash(F\backslash\{1\})^2.$$

Since $X \subseteq F$ is a finite generating set for F it follows that $\mathcal{B} \subseteq X$.

Lemma 3.19. *The basis $\mathcal{B}$ has size at least two. Thus the submonoid F of A^* is a free monoid of rank at least two.*

Proof. By assumption (C2) there are distinct words $\gamma, \delta \in \Delta$ with common initial letter $a \in A$. Write $\gamma \equiv a\gamma'$ and $\delta \equiv a\delta'$. Note that either γ' or δ' can begin with the letter a. By Lemma 3.9 the words γ and δ represent different elements of the monoid M. This in turn implies that $[\gamma'] \neq [\delta']$. Let $(a\gamma')^{-1}$ be a reduced word representing the inverse of $a\gamma'$ in M, and let $(a\delta')^{-1}$ be a reduced word representing the inverse of $a\delta'$ in M. In particular $(a\gamma')^{-1}, (a\delta')^{-1} \in \Delta^*$. Then by definition we have $\gamma_2 \equiv \gamma'(a\gamma')^{-1} \in X$ and $\delta_2 \equiv \delta'(a\delta')^{-1} \in X$, and both of these words are reduced words. Suppose, seeking a contradiction, that F is a free monoid of rank 1. It follows that there is a word $\nu \in A^+$ such that each of γ_2 and δ_2 is, in A^+, equal to some power of the word ν. But this would imply that γ' is a prefix of δ', or vice versa. Suppose without loss of generality γ' is a proper prefix of δ'. Then $a\gamma'$ is a proper prefix of $a\delta'$. But this contradicts condition (C1) since both of these words belong to Δ. This completes the proof of the lemma. □

The free submonoid F of the free monoid A^* defined above may also naturally be viewed as a free submonoid of the monoid M. This is because, as explained in Remark 3.12, all the words in F are reduced words and the concatenation of any two words from F is again a reduced word. In particular, since distinct reduced words represent distinct elements of M, the map $[\cdot] : A^* \to M$ defined by $w \mapsto [w]$ induces an embedding $[\cdot] : F \hookrightarrow M$. Thus we have identified a free submonoid of M of rank at least two, namely the image $[F]$ of F under this embedding.

Lemma 3.20. *Let $w \in A^*$ be a word. If $a^i w = 1$ and $a^j w = 1$ then $i = j$.*

Proof. Seeking a contradiction suppose that $a^i w = a^j w = 1$ with $j < i$. Then $a^{i-j} = a^{i-j}a^j w = a^i w = 1$. But this contradicts the fact that a is not invertible. □

Define a mapping $\omega : F \to \mathbb{Z}^{\geqslant 1}$ where $w \mapsto i$ if and only if $a^i w = 1$. This is a well-defined mapping by the previous lemma. Also, it is easy to see that ω is a homomorphism to $(\mathbb{Z}, +)$. The mapping ω assigns a weight to every element of the free monoid F. Abusing the notation, we also use ω to denote the map $\omega : [F] \to \mathbb{Z}^{\geqslant 1}$ defined by $\omega([w]) = \omega(w)$. This is well defined by the comments preceding Lemma 3.20.

The following result is now an immediate consequence of the previous results proved in this section.

Lemma 3.21. *Let $w \in A^*$ be a non-empty reduced word with $w \in F$, and suppose that $a^i w = 1$ with $i \geqslant 1$. Then the word w can be written uniquely as*

$$w \equiv w_1 w_2 \dots w_k$$

where $w_j \in \mathcal{B}$ for all $1 \leqslant j \leqslant k$, and

$$\omega(w_1) + \omega(w_2) + \dots + \omega(w_k) = i.$$

In the special case that Δ contains no word beginning with a^2 then $\omega(w_j) = 1$ for all $1 \leqslant j \leqslant k$, i.e. the statement above holds with $k = i$.

Note that in particular condition (C1) is satisfied if all the pieces have the same length. We note that Adjan [2] gives an algorithm for computing the set Δ by analysing overlaps of the relator with itself.

The following lemma will allow us to express membership in $\{a\}^*$ in terms of equations.

Lemma 3.22. *Let $u \in A^*$ be reduced. Then $u \in \{a\}^*$ if and only if $[ua] = [au]$ in M.*

Proof. Clearly if $u \in \{a\}^*$ then $[ua] = [au]$ in M.

For the converse, suppose that $u \in A^*$ is such that $[ua] = [au]$ in M. Since a is right invertible and a is not invertible, it follows that for all $\delta \in \Delta$ the last letter of δ is not equal to a. (Note this is true for all $\delta \in \Delta$ including those δ in Δ where δ does not begin with the letter a.)

Seeking a contradiction, suppose that $u \notin \{a\}^*$ and write $u \equiv u_1 a^y$ where $u_1 \in A^+$ and the last letter of u_1 is not equal to a, and $y \geqslant 0$. Consider $\mathrm{red}(ua) = \mathrm{red}(u_1 a^{y+1})$. Since for every rewrite rule $\alpha = \beta$ from (4) neither α nor β ends in the letter a, it follows that $\mathrm{red}(ua) \equiv w_1 a^{y+1}$ where w_1 does not end in the letter a.

In contrast, consider $\mathrm{red}(au) = \mathrm{red}(au_1 a^y)$. Reasoning in the same way as in the previous paragraph $\mathrm{red}(au) \equiv w_2 a^y$ where w_2 does not end in the letter a (note it may start with the letter a). In particular this implies that $\mathrm{red}(ua) \not\equiv \mathrm{red}(au)$ which implies $[ua] \neq [au]$. This contradicts our original assumption, and completes the proof of the lemma. □

The main result we shall prove in this section is the following.

Theorem 3.23. *Let $M = \langle A \mid r = 1 \rangle$ and let $\Delta \subseteq A^*$ be the set of minimal invertible pieces of r. Suppose that:*

(C1) no word from Δ is a proper subword of any other word from Δ, and

(C2) there exist distinct words $\gamma, \delta \in \Delta$ with a common first letter.

Then there exists a free submonoid D of M of finite rank $n \geqslant 2$ and a tuple of weights $\vec{\lambda} = (\lambda_1, \dots, \lambda_n)$ such that the free monoid with weighted length relation $(D, \cdot, 1, L_{\vec{\lambda}})$ is interpretable in M by systems of equations and one coefficient.

Proof. Let F be the free monoid from our previous arguments (see Remark 3.12), and consider the embedding $[F]$ of F in M via the map $[] : A^* \to M$ (see the discussion above Lemma 3.20). Let $\vec{\omega}$ be the tuple $(\omega(w_1), \dots, \omega(w_n))$ where $w_1, \dots, w_n$ freely generate F and $\omega : F \to \mathbb{Z}^{\geqslant 1}$

is the homomorphism defined after Lemma 3.20, so that $a^{\omega(w_i)}w_i = 1$ for all i (by Lemma 3.21). Recall that ω also then defines a map $\omega : [F] \to \mathbb{Z}^{\geqslant 1}$.

We claim that a generates an infinite submonoid of M. Indeed, if it did not, we would have $a^k = a^{k+\ell}$ for some $k, \ell \geqslant 0$. Since a is right invertible (due to condition (C2)), this implies that $a^\ell = 1$, from where it follows that a is invertible, a contradiction. This proves the claim.

By Lemma 3.22 the submonoid $\langle a \rangle$ is interpretable in M by the equation $ax = xa$ (Lemma 3.22). Since $[F] = \{x \in M \mid a^t x = 1 \text{ for some } t \in \mathbb{N}\backslash\{0\}\}$, it follows that $[F]$ is e-interpretable in M by the system of two equations $ay = ya, yx = 1$. □

The following two results follow immediately from the above Theorem 3.23 and from Proposition 2.5 regarding reducibility of decision problems.

Corollary 3.24. *Let M be a monoid satisfying the hypothesis of Theorem 3.23. Then there exists a free monoid with a weighted length relation $(D, \cdot, 1, \mathsf{L}_{\vec{\omega}})$ such that the Diophantine problem in $(D, \cdot, 1, \mathsf{L}_{\vec{\omega}})$ is reducible to the Diophantine problem in M. In particular, if the latter is decidable, then systems of word equations with $\vec{\omega}$-weighted length constraints are decidable as well.*

Theorem 3.25. *Any one-relator monoid of the form $\langle A \mid w = 1\rangle$ satisfying conditions (C1) and (C2) has undecidable positive AE-theory with coefficient. In particular, its first-order theory with coefficients is undecidable.*

Proof. It is an immediate consequence of Theorem 3.23, of the fact that the AE-theory with coefficients of free monoids is undecidable [24, 49] and of reducibility of theories (Proposition 2.5). □

If we add to Theorem 3.23 the extra condition that no word in Δ starts with a^2, then the same result holds with all weights being 1, i.e. $\vec{\lambda} = (1, \ldots, 1)$. In this case $\mathsf{L}_{\vec{\lambda}}$ is the standard length relation L:

Theorem 3.26. *Let $M = \langle A \mid r = 1\rangle$ and let $\Delta \subseteq A^*$ be the set of minimal invertible pieces of r. Suppose that:*

(C1) no word from Δ is a proper subword of any other word from Δ,

(C2) there exist distinct words $\gamma, \delta \in \Delta$ with a common first letter, say a,

(C3) no word in Δ starts with a^2.

Then there exists a free monoid D of finite rank $n \geqslant 2$ such that the free monoid with length relation $(D, \cdot, 1, =, L)$ is interpretable in M by systems of equations.

Proof. The proof works in the same way as in Theorem 3.23, with the addition that the last part of Lemma 3.21 now ensures that $\omega(w_i) = 1$ for all $i = 1, \ldots, n$. Then $\vec{\omega} = (1, \ldots, 1)$ and $(D, \cdot, 1, \mathsf{L}_{\vec{\omega}}) = (D, \cdot, 1, =, \mathsf{L})$, where $(D, \cdot, 1, \mathsf{L}_{\vec{\omega}})$ is the free monoid with weighted length relation given by Theorem 3.23. Hence $(D, \cdot, 1, =, \mathsf{L})$ is interpretable in M by systems of equations and one coefficient. □

We obtain an analogue of Corollary 3.24

Corollary 3.27. *Let M be a monoid satisfying the hypothesis of Theorem 3.26. Then there exists a free monoid with (non-weighted) length relation $(D, \cdot, 1, =, \mathsf{L})$ such that the Diophantine problem in $(D, \cdot, 1, =, \mathsf{L})$ is reducible to the Diophantine problem in M. In particular, if the latter is decidable, then systems of word equations with length constraints are decidable as well.*

We further prove that the monoids from Theorems 3.23 and 3.26 naturally embed the monoids from Example 3.4.

Theorem 3.28. *Let $M = \langle A \mid r = 1\rangle$ and let $\Delta \subseteq A^*$ be the set of minimal invertible pieces of r. Suppose conditions (C1), (C2), (C3) are satisfied, i.e.:*

(C1) no word from Δ is a proper subword of any other word from Δ,

(C2) there exist distinct words $\gamma, \delta \in \Delta$ with a common first letter, say a,

(C3) no word in Δ starts with a^2.

Let

$$\Sigma_a = \{w \in A^* : \ w \text{ is reduced and } [aw] = 1 \ \}.$$

Then

(i) Σ_a is a finite set with $|\Sigma_a| \geqslant 2$;

(ii) the submonoid of M generated by Σ_a is free with basis Σ_a.

Let $\Sigma_a = \{\gamma_1, \ldots, \gamma_q\}$. Then the submonoid of M generated by $\{[a]\} \cup [\Sigma_a]$ is naturally isomorphic to the monoid defined by the presentation

$$\langle a, d_1, d_2, \ldots, d_q \mid ad_1 = 1, \ldots, ad_q = 1\rangle.$$

Proof. We claim that Σ_a is equal to the set $X = \{\beta(a^j\beta)^{-1} \mid a^j\beta \in \Delta\}$ defined above; see Lemma 3.18. It is immediate from the definition of X that $X \subseteq \Sigma_a$. For the converse, let $\gamma \in \Sigma_a$. This means that γ is a reduced word and $[a\gamma] = 1$ in M. By Lemma 3.17 we can write $a\gamma \equiv a\gamma'\gamma''$ with $\gamma' \in X$ and $a\gamma' = 1$ in M. Then $\gamma'' = (a\gamma')\gamma'' = a\gamma = 1$ in M. Since γ is a reduced word it follows that $\gamma'' \equiv \epsilon$ and thus $\gamma \equiv \gamma' \in X$. This completes the proof that $X = \Sigma_a$.

Since $X = \Sigma_a$, part (i) now follows from Lemmas 3.18 and 3.19.

To prove part (ii) it will suffice to prove that $\Sigma_a = \mathcal{B}$, where $\mathcal{B}$ is the unique basis of the free monoid generated by $\Sigma_a = X$. To prove this it will suffice to prove that no $\gamma \in \Sigma_a$ can be written as a product of other γ from Σ_a. Suppose that

$$\gamma = \gamma_1\gamma_2 \cdots \gamma_m$$

where $\gamma_i \in \Sigma_a$ for all $1 \leqslant i \leqslant m$. Then

$$1 = a\gamma = a\gamma_1\gamma_2 \ldots \gamma_m = \gamma_2 \cdots \gamma_m$$

By Lemma 3.11, $\gamma_2 \ldots \gamma_m$ is a reduced word and hence it follows that it must equal the empty word. Hence $m = 1$ and $\gamma \equiv \gamma_1$ since they are both reduced words and they are equal in M. This completes the proof that $\Sigma_a = \mathcal{B}$, and hence completes the proof of (ii).

For the last part, let $w \in (\{a\} \cup \Sigma_a)^* = \{a, \gamma_1, \ldots, \gamma_q\}^*$. Since $a\gamma_i = 1$ for all i, this word is equal in M to a word w' where w' has the form $w' \equiv w_1 a^j$, where $w_1 \in \{\gamma_1, \ldots, \gamma_q\}^*$. We claim that in fact $\mathrm{red}(w) \equiv w_1 a^j$. Indeed, since none of the words appearing in the rewrite rules in (4) ends in a (because otherwise together with condition (C2) this would imply that a is invertible) to show that $w_1 a^j$ is reduced it suffices to prove that w_1 is reduced, and this was proved in Lemma 3.11. Therefore, each element of the submonoid of M generated by $\{[a]\} \cup [\Sigma_a]$ may be uniquely written in the form αa^j for some $j \geqslant 0$ and some word $\alpha \in \{\gamma_1, \ldots, \gamma_q\}^*$. Now consider the monoid N defined by the presentation

$$\langle a, d_1, d_2, \ldots, d_q \mid ad_1 = 1, \ldots, ad_q = 1\rangle$$

This is a finite complete presentation, and the reduced words are precisely those of the form βa^j where $j \geqslant 0$ and $\beta \in \{d_1, \ldots, d_q\}^*$.

Let $\phi : \{a, d_1, \ldots, d_q\}^* \to A^*$ be the homomorphism induced by the map $a \mapsto a$, and $d_i \mapsto \gamma_i$ for $1 \leqslant i \leqslant q$. Since each relation in the presentation for $\langle a, d_1, \ldots, d_q\rangle$ is preserved by this homomorphism it follows that ϕ induces a homomorphism $\phi : \langle a, d_1, \ldots, d_q\rangle \to M$. Moreover, this homomorphism maps $\langle a, d_1, \ldots, d_q\rangle$ bijectively to the submonoid of M generated by $\{[a]\} \cup [\Sigma_a]$ since it clearly defines a bijection between the normal forms described above. This completes the proof of the theorem. □

Remark 3.29. We follow the notation of the previous Theorem 3.28. In the proof of Theorem 3.23 we showed that both $\langle a\rangle$ and $\langle \Sigma_a\rangle$ are e-interpretable in M. It is natural to ask whether the submonoid $\langle a, \Sigma_a\rangle$, which by Theorem 3.28 is isomorphic to the monoid from Example 3.4, is itself e-interpretable in M. The answer to this question is not clear and we leave it open.

3.1 One-relator monoids with hyperbolic undirected Cayley graph and hyperbolic group of units

In this subsection we prove some sufficient conditions for one-relator monoids to have hyperbolic undirected Cayley graph and to have hyperbolic group of units. These are of interest to the paper given our question in the introduction regarding the reducibility of the Diophantine problem in a special one-relator monoid to the same problem in its group of units: since the Diophantine problem in hyperbolic groups is decidable [16, 64], such a reduction would imply the decidability of the Diophantine problem in the one-relator monoid.

Before presenting the main result of this section we first define what we mean by the undirected Cayley graph of a monoid, and what it means for this graph to be hyperbolic. For more background on the theory of hyperbolic metric spaces and hyperbolic groups we refer the reader to [12].

Let (X, d) be a metric space. For $x, y \in X$ a *geodesic path* from x to y is a map $f : [0, l] \to X$ from the closed interval $[0, l] \subseteq \mathbb{R}$ to X such that $f(0) = x$, $f(l) = y$ and $d(f(a), f(b)) = |a - b|$ for all $a, b \in [0, l]$. Note in particular this implies that $d(x, y) = l$. The image α of the map f is called a *geodesic segment* with endpoints x and y. A *geodesic metric space* is one in which there exist geodesic segments between all pairs of points. Note that in general there can be more than one geodesic segment between a given pair of points. A *geodesic triangle* Ω in X is a union of three geodesic segments from x to y, y to z and z to x, where $x, y, z \in X$. These three geodesic segments are called the *sides* of the geodesic triangle Ω.

Definition 3.30. Let X be a geodesic metric space and let Ω be a geodesic triangle in X with sides α, β and γ. The triangle Ω is called *δ-slim* if for every point a on α the distance from a to $\beta \cup \gamma$ is less than δ, and similarly every point b on β is within distance δ of $\alpha \cup \gamma$, and every point c on γ is within distance δ of $\alpha \cup \beta$. If every geodesic triangle in X is δ-slim then we say that the geodesic metric space X is *δ-hyperbolic*. If X is δ-hyperbolic for some $\delta > 0$ then we say X is *hyperbolic*.

Let M be a monoid generated by a set A. Then by the *undirected Cayley graph* $\Gamma(M, A)$ *of M with respect to the generating set A* we mean the graph with vertex set M and where there is an undirected edge connecting $m \in M$ to $n \in M$ if and only if $ma = n$ or $na = m$ for some $a \in A$. Note that here we have opted to work with the right Cayley graph, but the results we prove in this subsection are also true working with the left Cayley graph instead. The graph $\Gamma(M, A)$ is a metric space with the usual distance metric on graphs where for $a, b \in M$ we define $d(a, b)$ to be the shortest length of a path in $\Gamma(M, A)$ from a to b. This is not a geodesic metric space, but can be made into one in a natural way by making each edge isometric to the unit interval $[0, 1]$ and extending the metric to the points of these edges in the obvious way. This is called the *geometric realisation* of the Cayley graph. We say that the *undirected Cayley graph of a monoid M is hyperbolic* if the geometric realisation of $\Gamma(M, A)$ is a hyperbolic metric space. If M is a finitely generated monoid, it may be shown that this property is independent of the choice of finite generating set for M, so it makes sense to talk about a finitely generated monoid having a hyperbolic undirected Cayley graph, without reference to any specific finite generating set.

Proposition 3.31. *Let $M = \langle A \mid w = 1 \rangle$. Let G be the group of units of M. If G is a hyperbolic group then the undirected Cayley graph of M is hyperbolic.*

Proof. As usual, let Δ be the set of minimal invertible pieces of the relator w (see the discussion above Theorem 3.5 for further details). Let I be the set of all non-empty prefixes of the words from Δ, that is

$$I = \{x \in A^+ \mid xy \in \Delta \text{ for some } y \in A^*\}.$$

Let $Y = \{[u] : u \in I\}$. Then, by Zhang [70, Lemma 3.3], Y is a finite generating set for the submonoid of right units R of M. Note that R is the $\mathcal{R}$-class of the identity element of M, where $\mathcal{R}$ is Green's $\mathcal{R}$-relation on M defined by saying $m\mathcal{R}n$ if and only if $mM = nM$. Clearly Δ is a subset of I. Let $\mathcal{G}$ be the underlying undirected graph of the right Cayley graph of the monoid R, with respect to the generating set Y. So $\mathcal{G}$ has vertex set R and edges $\{[u], [ux]\}$ where $u \in A^*$, $x \in I$ and $\{[u], [ux]\}$ is a subset of R. Note that $\mathcal{G}$ is a connected infinite graph and its vertices have bounded degree since R is a right cancellative monoid. We use $\mathcal{S}$ to denote the undirected Schützenberger graph of the $\mathcal{R}$-class R. So $\mathcal{S}$ also has vertex set R but has edges $\{[u], [ua]\}$ where $u \in A^*$, $a \in A$ and $\{[u], [ua]\}$ is a subset of R.

We claim that the identity map on R defines a quasi-isometry between the graph $\mathcal{G}$ and the graph $\mathcal{S}$.

To prove this claim, let $d_{\mathcal{S}}$ and $d_{\mathcal{G}}$ denote the distances in each of these graphs. Consider an arbitrary edge $\{[u], [ux]\}$ in the graph $\mathcal{G}$. Let D be the maximum length of a word in Δ. Then $d_{\mathcal{S}}([u], [ux]) \leqslant D$. For the converse, let $\{[u], [ua]\}$ be an arbitrary edge in the graph $\mathcal{S}$. We claim that $d_{\mathcal{G}}([u], [ua]) \leqslant 2$. We may assume without loss of generality that u is a reduced word. There are now two cases to consider.

First suppose that ua is a reduced word. It then follows from [70, Lemma 3.3] that $ua \in I^*$ (i.e. is a graphical product of words from I). Note that ua may admit several different decompositions in I^*. Write $ua = u'\gamma$ where $\gamma \in I$ and $u' \in I^*$. If $|\gamma| = 1$ then $a = \gamma \in I$ and so $d_{\mathcal{G}}([u],[ua]) = 1$. Now suppose that $|\gamma| > 1$. Write $\gamma = \gamma' a$ with $\gamma' \in I$. Then we have $u = u'\gamma'$ and both $\{[u'],[u'\gamma']\}$ and $\{[u'],[u'\gamma]\}$ are edges in the graph $\mathcal{G}$. It follows that $d_{\mathcal{G}}([u],[ua]) = d_{\mathcal{G}}([u'\gamma'],[u'\gamma]) \leqslant 2$.

Now suppose that ua is not a reduced word. Since u is reduced, it follows that we can write $ua = u'\gamma$ where $\gamma \in \Delta$ is a non-empty word. Then arguing as in the previous paragraph, either $|\gamma| = 1$ and $d_{\mathcal{G}}([u],[ua]) = 1$, or else $|\gamma| > 1$ and $d_{\mathcal{G}}([u],[ua]) \leqslant 2$. This completes the proof of the claim that the identity mapping on R induces a quasi-isometry between the graph $\mathcal{G}$ and the graph $\mathcal{S}$.

It follows from [70, Theorem 4.5] that the submonoid of right units R of M is isomorphic to a monoid free product $T * G$ where T is a free monoid of finite rank, and G is the group of units of the monoid M. Since the Cayley graph of a free monoid is a tree, it then follows that the undirected Cayley graph $\mathcal{G}$ of $R \cong T * G$ is hyperbolic. Since $\mathcal{S}$ is quasi-isometric to $\mathcal{G}$ we conclude that the undirected Schützenberger graph of the $\mathcal{R}$-class of the identity element is hyperbolic.

It follows from the results in [30, Section 3] that (i) the Schützenberger graphs of any pair of $\mathcal{R}$-classes of M are isomorphic to each other, and (ii) for every $\mathcal{R}$-class R' of M there is at most one edge $\{m, ma\}$ in the Cayley graph of M such that $m \in M$, $a \in A$, with $ma \in R'$ but $m \notin R'$, and (iii) the quotient graph with vertex set the $\mathcal{R}$-classes of M and edges all edges $\{m,n\}$ from the Cayley graph of M such that $(m,n) \notin \mathcal{R}$ is a rooted tree.

Combining these observations we see that the Cayley graph of M has the structure of a "regular tree of copies of" the hyperbolic graph $\mathcal{S}$. From this it then quickly follows (e.g. by applying [29, Theorem 5.4]) that the undirected Cayley graph of M is hyperbolic. □

In fact, using a similar argument, it may be shown that Proposition 3.31 holds more generally for any finitely presented monoid M defined by a presentation of the form

$$\langle A \mid w_1 = 1, \ldots, w_k = 1 \rangle.$$

Proposition 3.32. *Let $M = \langle A \mid w^k = 1 \rangle$ ($k \geqslant 2$). Then the group of units of M is a one-relator group with torsion. It follows that the group of units of M is a hyperbolic group, and the undirected Cayley graph of M is a hyperbolic metric space.*

Proof. It follows from results of Adjan [2] that the group of units G of M is a one-relator group with torsion (see [30, Section 3] for a proof of this). By the Newman Spelling Theorem [45, Theorem 5.5] we have that G is a hyperbolic group. This and Proposition 3.31 imply that the undirected Cayley graph of M is hyperbolic. □

4 Applications, examples and open problems

In this section we list some examples, and classes of examples, of monoids to which the main results of this paper apply. We shall also collect together a selection of open problems, and possible future research directions, which naturally arise from our results. As part of this we will identify the simplest examples of one-relator monoids for which we do not yet know whether or not the Diophantine problem is decidable. In general, we do not know if there is an example of a one-relator monoid of the form $\langle A \mid r = 1\rangle$ with undecidable Diophantine problem.

Let us begin by recording some examples of one-relator monoids of the form $\langle A \mid r = 1\rangle$ where we have shown that the Diophantine problem is decidable. Consider, in particular the case of 2-generated one-relator monoids $\langle a, b \mid r = 1\rangle$. Let M denote the monoid defined by this presentation. Very often questions about one-relator monoids can be reduced to just considering the 2-generator case e.g. this is the case for the word problem.

By Makanin [47] the Diophantine problem is decidable for the free monoid $\langle a, b \mid \rangle$, while in [20, Example 21] it is proved that it is decidable for the bicyclic monoid $\langle a, b \mid ab = 1\rangle$. Now consider the general case $\langle a, b \mid r = 1\rangle$ and let $r = r_1 r_2 \ldots r_k$ be the decomposition of r into minimal invertible pieces as described in Section 3 and in Example 1.1 and the paragraph preceding it. If $r \in \{a\}^*$ or $r \in \{b\}^*$ then the monoid is a free product of a free monoid of rank one and a finite cyclic group, and thus the Diophantine problem is decidable by [20]. Now suppose that both the letters a and b appear in the defining relator r. There are then two cases to consider. If there are minimal invertible pieces r_i and r_j such that the first letter of r_i equals the last letter of r_j, then applying the Adjan overlap algorithm it follows that both a and b both represent invertible elements of M and hence M is a group. In this case, M is the group defined by the same one-relator group presentation, and hence M is a so-called *positive* one-relator group. Such groups have been studied e.g. by Baumslag [9] and Wise [67]. This motivates the question of whether the Diophantine problem is decidable for positive one-relator groups. Up to symmetry the case that remains is when all the invertible pieces r_i $(1 \leqslant i \leqslant k)$ begin with the letter a and end with the letter b. This case then divides into two subcases, either (i) all of the pieces r_i are equal to each other as words, or (ii) there is some pair of minimal invertible pieces r_i and r_j with $r_i \not\equiv r_j$. Note that subcase (i) includes in particular the case where there is a single invertible piece. This is precisely the case where the relator r is self-overlap free meaning that no proper non-empty prefix is equal to a proper non-empty suffix of r. This in turn is equivalent to saying that the group of units of the monoid is the trivial group. Also note that many of the examples in (ii) will satisfy the conditions (C1) and (C2) (and (C3)) from Section 3, and thus the main theorems of that section, Theorem 3.23 and Theorem 3.26, will apply to them. Some examples of these are listed in the introduction after Theorem A.

A similar division into cases can also be done for one-relator monoids $\langle A \mid r = 1\rangle$ with more than two generators. For instance, as we already saw in Example 3.2, the monoid $\langle a, b, c, d \mid aba = 1\rangle$ has decidable Diophantine problem by Theorem 3.1 above, since all the letters in the relator are invertible, and the group of units is the infinite cyclic group which has decidable Diophantine problem. Similarly the monoid $\langle a, b, c, d, e, f \mid abcddcbbaa = 1\rangle$ has decidable Diophantine problem, again applying Theorem 3.1. Indeed, applying the Adjan overlap algorithm we deduce that all the letters a, b, c and d appearing in the defining relator are invertible, and the group of units of this monoid is defined by the group presentation

$$\mathrm{Gp}\langle a, b, c, d \mid abcddcbbaa = 1\rangle.$$

To apply Theorem 3.1 we need to show that this group has decidable Diophantine problem. To show this, note that this group can be written

$$\mathrm{Gp}\langle a, b, c, d \mid cddc = b^{-1}a^{-1}a^{-1}a^{-1}b^{-1}b^{-1}\rangle.$$

The words $cddc$ and $b^{-1}a^{-1}a^{-1}a^{-1}b^{-1}b^{-1}$ are non-primitive since the words $cddc$ and $bbaaab$ are not Christoffel words (see e.g. [59]), and neither are any of the conjugates of these words, since the first word have the same number of cs and ds, and similarly for the second word. It is known, see [11, 34, 36], that a cyclically pinched one-relator group defined by a presentation $\mathrm{Gp}\langle A|u = v\rangle$, where u and v are non-primitive words written over disjoint sets of letters, and it is not the case that both u and v are proper powers, is hyperbolic. Hence the group of units of $\langle a, b, c, d, e, f \mid abcdcbba = 1\rangle$ is a hyperbolic group and thus by Theorem 3.1 above this monoid has decidable Diophantine problem. Many other examples similar to this can be written down. This gives a reasonably rich source of examples of one-relator monoids $\langle A \mid r = 1\rangle$ which have solvable Diophantine problem as a consequence of the fact that their groups of units are hyperbolic. We do not know in general whether having a hyperbolic group of units is enough to imply that a one-relator monoid of the form $\langle A \mid r = 1\rangle$ has solvable Diophantine problem. As explained in the introduction, this was one of the original motivating questions for the work done in this paper. By Proposition 3.32 and Theorem 3.26, a positive answer to this questions implies decidability of word equations with length constraints.

In light of this discussion, it is sensible to identify the simplest examples of one-relator monoids of the form $\langle A \mid r = 1\rangle$ for which we neither know that the Diophantine problem is decidable, but we also do not know of a reduction theorem (like the theorems from Section 3 above) of a known difficult open problem. Thus we ask whether either of the monoids $\langle b, c \mid b^2c = 1\rangle$ or $\langle a, b, c \mid abc = 1\rangle$ has decidable Diophantine problem? Initial investigations indicate that this might relate to solving word equations with a variation on the notion of twisting, in the sense of [22]. More generally we ask the following

Question 4.1. *If the word $w \in A^*$ has no self overlaps, i.e. there is no non-empty word which is both a proper prefix of w and a proper suffix of w, then is the Diophantine problem for the one-relator monoid $\langle A \mid w = 1\rangle$ decidable?*

Note that the condition that w has no self overlaps is equivalent to saying the group of units of this monoid is trivial (this follows from the discussion immediately before the statement of Theorem 3.5).

The corresponding class of monoids with torsion are also not covered by any of the theorems in this paper. Thus we ask whether $\langle b, c \mid bcbc = 1\rangle$ has decidable Diophantine problem? More generally, of course, we can ask whether the Diophantine problem is decidable for monoids $\langle A \mid w^n = 1\rangle$ where w has no self overlaps.

Finally, we restate some natural questions which have arisen in this work. As already mentioned above, if any of these problems has a positive answer, then as a corollary this would give a positive solution to the open problem of solving word equations with length constraints.

Question 4.2. *Is the Diophantine problem decidable for one-relator monoids of the form $\langle A \mid w^n = 1\rangle$ where $n > 1$?*

Question 4.3. *Let M be the monoid defined by $\langle A \mid w = 1\rangle$ and let G be the group of units of M. If the Diophantine problem is decidable in G, then does it follow that it is decidable in M?*

It follows from the results in the present paper that the positive AE-theory is in general undecidable in the classes of monoids from Questions 4.1 through Question 4.3 (due to Theorem 3.25, Proposition 3.32, and Remark 3.4).

5 Acknowledgements

We would like to thank the anonymous reviewer who carefully read our paper, and whose comments and suggestions helped improve it.

We are also thankful to Olga Kharlampovich and Alexei Miasnikov for useful conversations.

This research was initiated during the Heilbronn Focused Research Workshop "Equations in One-Relator Groups and Semigroups", in September 2018, at the ICMS in Edinburgh organised by Laura Ciobanu of Heriot-Watt University.

6 Bibliography

[1] P. A. Abdulla, M. F. Atig, Y.-F. Chen, L. Holík, A. Rezine, P. Rümmer, and J. Stenman. Norn: An SMT solver for string constraints. In D. Kroening and C. S. Păsăreanu, editors, *Computer Aided Verification*, pages 462–469, Cham, 2015. Springer International Publishing.

[2] S. I. Adjan. Defining relations and algorithmic problems for groups and semigroups. *Trudy Mat. Inst. Steklov.*, 85:123, 1966.

[3] J. Araújo, M. Kinyon, J. Konieczny, and A. Malheiro. Decidability and independence of conjugacy problems in finitely presented monoids. *Theoret. Comput. Sci.*, 731:88–98, 2018.

[4] J. Araújo, J. Konieczny, and A. Malheiro. Conjugation in semigroups. *J. Algebra*, 403:93–134, 2014.

[5] S. Arora and B. Barak. *Computational Complexity: A Modern Approach.* Cambridge University Press, New York, NY, USA, 1st edition, 2009.

[6] A. Aydin, L. Bang, and T. Bultan. Automata-based model counting for string constraints. In D. Kroening and C. S. Păsăreanu, editors, *Computer Aided Verification*, pages 255–272, Cham, 2015. Springer International Publishing.

[7] C. Barrett, C. L. Conway, M. Deters, L. Hadarean, D. Jovanović, T. King, A. Reynolds, and C. Tinelli. Cvc4. In G. Gopalakrishnan and S. Qadeer, editors, *Computer Aided Verification*, pages 171–177, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[8] C. Barrett, R. Sebastiani, S. Seshia, and C. Tinelli. *Satisfiability modulo theories*, volume 185 of *Frontiers in Artificial Intelligence and Applications*, pages 825–885. 1 edition, 2009.

[9] G. Baumslag. Positive one-relator groups. *Trans. Amer. Math. Soc.*, 156:165–183, 1971.

[10] M. Berzish, V. Ganesh, and Y. Zheng. Z3str3: A string solver with theory-aware heuristics. In *2017 Formal Methods in Computer Aided Design (FMCAD)*, pages 55–59, Oct 2017.

[11] M. Bestvina and M. Feighn. A combination theorem for negatively curved groups. *J. Differential Geom.*, 35(1):85–101, 1992.

[12] M. R. Bridson and A. Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[13] J. R. Büchi and S. Senger. Definability in the existential theory of concatenation and undecidable extensions of this theory. *Z. Math. Logik Grundlag. Math.*, 34(4):337–342, 1988.

[14] M. Casals-Ruiz and I. Kazachkov. On systems of equations over free products of groups. *Journal of Algebra*, 333(1):368 – 426, 2011.

[15] L. Ciobanu Radomirovic, V. Volker Diekert, and M. Elder. Solution sets for equations over free groups are EDT0L languages. *International Journal of Algebra and Computation*, 26(5):843–886, 8 2016.

[16] F. Dahmani and V. Guirardel. Foliations for solving equations in groups: free, virtually free, and hyperbolic groups. *Journal of Topology*, 3(2):343–404, 2016.

[17] J. D. Day, V. Ganesh, P. He, F. Manea, and D. Nowotka. The satisfiability of word equations: Decidable and undecidable theories. In I. Potapov and P.-A. Reynier, editors, *Reachability Problems*, pages 15–29, Cham, 2018. Springer International Publishing.

[18] L. De Moura and N. Bjørner. Satisfiability modulo theories: Introduction and applications. *Commun. ACM*, 54(9):69–77, September 2011.

[19] T. Deis, J. Meakin, and G. Sénizergues. Equations in free inverse monoids. *Internat. J. Algebra Comput.*, 17(4):761–795, 2007.

[20] V. Diekert and M. Lohrey. Word equations over graph products. *International Journal of Algebra and Computation*, 18(03):493–533, 2008.

[21] V. Diekert. More than 1700 years of word equations. In A. Maletti, editor, *Algebraic Informatics*, pages 22–28, Cham, 2015. Springer International Publishing.

[22] V. Diekert and M. Elder. Solutions to twisted word equations and equations in virtually free groups. *International Journal of Algebra and Computation*, 30(04):731–819, 2020.

[23] V. Diekert, F. Martin, G. Sénizergues, and P. V. Silva. Equations over free inverse monoids with idempotent variables. *Theory Comput. Syst.*, 61(2):494–520, 2017.

[24] V. G. Durnev. Undecidability of the positive $\forall\exists^3$-theory of a free semigroup. *Siberian Mathematical Journal*, 36(5):917–929, Sep 1995.

[25] V. Ganesh, A. Kieżun, S. Artzi, P. J. Guo, P. Hooimeijer, and M. Ernst. HAMPI: a string solver for testing, analysis and vulnerability detection. In *Computer aided verification*, volume 6806 of *Lecture Notes in Comput. Sci.*, pages 1–19. Springer, Heidelberg, 2011.

[26] V. Ganesh, M. Minnes, A. Solar-Lezama, and M. Rinard. Word equations with length constraints: What's decidable? In A. Biere, A. Nahir, and T. Vos, editors, *Hardware and Software: Verification and Testing*, pages 209–226, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[27] A. Garreta, A. Miasnikov, and D. Ovchinnikov. Diophantine problems in rings and algebras: undecidability and reductions to rings of algebraic integers. *arXiv e-prints*, May 2018.

[28] A. Garreta, A. Miasnikov, and D. Ovchinnikov. Diophantine problems in solvable groups. *Bulletin of Mathematical Sciences*, 10(01):2050005, 2020.

[29] R. D. Gray, P. V. Silva, and N. Szakács. Algorithmic properties of inverse monoids with hyperbolic and tree-like schützenberger graphs. *arXiv e-prints*, 2019.

[30] R. D. Gray and B. Steinberg. Topological finiteness properties of monoids. Part 2: special monoids, one-relator monoids, amalgamated free products, and HNN extensions. *arXiv e-prints*, May 2018.

[31] W. Hodges, S. Hodges, H. Wilfrid, G. Rota, B. Doran, P. Flajolet, T. Lam, E. Lutwak, and M. Ismail. *Model Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1993.

[32] J. M. Howie. *Fundamentals of semigroup theory*, volume 12 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1995. Oxford Science Publications.

[33] A. Jez. Word Equations in Nondeterministic Linear Space. In I. Chatzigiannakis, P. Indyk, F. Kuhn, and A. Muscholl, editors, *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, volume 80 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 95:1–95:13, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[34] A. Juhász and G. Rosenberger. On the combinatorial curvature of groups of F-type and other one-relator free products. In *The mathematical legacy of Wilhelm Magnus: groups, geometry and special functions (Brooklyn, NY, 1992)*, volume 169 of *Contemp. Math.*, pages 373–384. Amer. Math. Soc., Providence, RI, 1994.

[35] O. Kharlampovich and A. G. Miasnikov. Model theory and algebraic geometry in groups, non-standard actions and algorithmic problems. *Proceedings of the Intern. Congress of Mathematicians, Seoul, v. 2, invited lectures, 223-244*, 2014.

[36] O. Kharlampovich and A. Myasnikov. Hyperbolic groups and free constructions. *Trans. Amer. Math. Soc.*, 350(2):571–613, 1998.

[37] O. Kharlampovich and L. López. Bi-interpretability of some monoids with the arithmetic and applications. *Semigroup Forum*, 99(1):126–139, Aug 2019.

[38] O. Kharlampovich, L. Lopez Cruz, and A. Miasnikov. The Diophantine problem in some metabelian groups. *Mathematics of Computation*, 89:1, 02 2020.

[39] O. Kharlampovich and A. Myasnikov. Elementary theory of free non-abelian groups. *J. Algebra*, 302(2):451–552, 2006.

[40] O. Kharlampovich and A. G. Myasnikov. Equations and fully residually free groups. In O. Bogopolski, I. Bumagin, O. Kharlampovich, and E. Ventura, editors, *Combinatorial and Geometric Group Theory*, pages 203–242, Basel, 2010. Birkhäuser Basel.

[41] G. Lallement. On monoids presented by a single relation. *J. Algebra*, 32:370–388, 1974.

[42] A. W. Lin and R. Majumdar. Quadratic word equations with length constraints, counter systems, and presburger arithmetic with divisibility. In S. K. Lahiri and C. Wang, editors, *Automated Technology for Verification and Analysis*, pages 352–369, Cham, 2018. Springer International Publishing.

[43] M. Lohrey and G. Sénizergues. Theories of HNN-extensions and amalgamated products. In M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, editors, *Automata, Languages and Programming*, pages 504–515, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[44] M. Lothaire. *Combinatorics on Words*. Cambridge Mathematical Library. Cambridge University Press, 2 edition, 1997.

[45] R. Lyndon and P. Schupp. *Combinatorial Group Theory*. Classics in Mathematics. Springer Berlin Heidelberg, 2015.

[46] G. S. Makanin. On the identity problem in finitely defined semigroups. *Dokl. Akad. Nauk SSSR*, 171:285–287, 1966.

[47] G. S. Makanin. The problem of solvability of equations in a free semigroup. *Mathematics of the USSR-Sbornik*, 32(2):129–198, feb 1977.

[48] G. S. Makanin. Equations in a free group. *Mathematics of the USSR-Izvestiya*, 21(3):483, 1983.

[49] S. S. Marchenkov. Undecidability of the positive ∀∃-theory of a free semigroup. *Sibirsk. Mat. Zh.*, 23(1):196–198, 223, 1982.

[50] D. Marker. *Model Theory : An Introduction*. Graduate Texts in Mathematics. Springer New York, 2013.

[51] J. V. Matijasevič. The Diophantineness of enumerable sets. *Dokl. Akad. Nauk SSSR*, 191:279–282, 1970.

[52] Y. Matiyasevich. Word equations, Fibonacci numbers, and Hilbert's Tenth Problem. 02 2019.

[53] P. Narendran and F. Otto. Complexity results on the conjugacy problem for monoids. *Theoret. Comput. Sci.*, 35(2-3):227–243, 1985.

[54] P. Narendran, F. Otto, and K. Winklmann. The uniform conjugacy problem for finite Church-Rosser Thue systems is NP-complete. *Inform. and Control*, 63(1-2):58–66, 1984.

[55] F. Otto. Conjugacy in monoids with a special Church-Rosser presentation is decidable. *Semigroup Forum*, 29(1-2):223–240, 1984.

[56] W. Plandowski. Satisfiability of word equations with constants is in PSPACE. *J. ACM*, 51(3):483–496, 2004.

[57] W. V. Quine. Concatenation as a basis for arithmetic. *Journal of Symbolic Logic*, 11(4):105–114, 1946.

[58] A. A. Razborov. On systems of equations in a free group. *Mathematics of the USSR-Izvestiya*, 25(1):115, 1985.

[59] C. Reutenauer. *From Christoffel words to Markoff numbers.* Oxford University Press, Oxford, 2019.

[60] E. Rips and Z. Sela. Canonical representatives and equations in hyperbolic groups. *Inventiones mathematicae*, 120(1):489–512, 1995.

[61] V. A. Roman′kov. Equations over groups. *Groups Complexity Cryptology*, 4:191–239, 2012.

[62] B. V. Rozenblat. Diophantine theories of free inverse semigroups. *Sibirsk. Mat. Zh.*, 26(6):101–107, 190, 1985.

[63] Z. Sela. Diophantine geometry over groups. VI. The elementary theory of a free group. *Geom. Funct. Anal.*, 16(3):707–730, 2006.

[64] Z. Sela. Diophantine geometry over groups. VII. The elementary theory of a hyperbolic group. *Proc. Lond. Math. Soc. (3)*, 99(1):217–273, 2009.

[65] Z. Sela. Word Equations I: Pairs and their Makanin-Razborov diagrams. *arXiv e-prints*, Jul 2016.

[66] M.-T. Trinh, D.-H. Chu, and J. Jaffar. Progressive reasoning over recursively-defined strings. In S. Chaudhuri and A. Farzan, editors, *Computer Aided Verification*, pages 218–240, Cham, 2016. Springer International Publishing.

[67] D. T. Wise. The residual finiteness of positive one-relator groups. *Comment. Math. Helv.*, 76(2):314–338, 2001.

[68] F. Yu, M. Alkhalaf, and T. Bultan. Stranger: An automata-based string analysis tool for php. In J. Esparza and R. Majumdar, editors, *Tools and Algorithms for the Construction and Analysis of Systems*, pages 154–157, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[69] L. Zhang. Conjugacy in special monoids. *J. Algebra*, 143(2):487–497, 1991.

[70] L. Zhang. Applying rewriting methods to special monoids. *Math. Proc. Cambridge Philos. Soc.*, 112(3):495–505, 1992.