Word equations and the exponent of periodicity

Volker Diekert
University of Stuttgart, Germany
diekert@fmi.uni-stuttgart.de

Silas Natterer
University of Stuttgart, Germany
natterersilas@gmail.com

Alexander Thumm
University of Siegen, Germany
alexander.thumm@uni-siegen.de

Abstract

In this article, we study word equations in free semigroups and the conjecture that the existence of infinitely many solutions entails the existence of solutions with arbitrarily large exponent of periodicity. We examine this question in the broader framework of word equations with regular constraints and establish new positive results: the conjecture holds for all quadratic word equations with constraints in finite semigroups from the variety **DLG** and its left-right dual **DRG**, encompassing, in particular, all finite groups, commutative semigroups, and $\mathcal{J}$-trivial semigroups.

2012 ACM Subject Classification Theory of computation → Formal languages and automata theory; Computing methodologies → Symbolic and algebraic algorithms

Keywords and phrases Mathematics Subject Classification 2020: 03D05, 68Q45, 20F70.

1 Introduction

Word equations form a rich area of study. A natural question in this general setting asks whether the existence of infinitely many solutions implies the existence of solutions with arbitrarily large exponent of periodicity. In this article, we examine a variant of this question for word equations with regular constraints, uncovering interesting connections between word equations and structural properties of finite semigroups along the way.

1.1 A Conjecture

Consider a word equation $U = V$ over constants and variables, and let σ be a solution, that is, a substitution mapping variables to words over constants in such a way that $\sigma(U) = \sigma(V)$.

The *exponent of periodicity* of σ, denoted by $\exp(\sigma)$, is defined as the largest power of a nonempty word that appears as a factor in some $\sigma(X)$, where X is a variable. (The formal definition is given in Definition 4.) This notion naturally leads to the following question.

Provided that a word equation admits infinitely many solutions,
must there exist solutions with arbitrarily large exponent of periodicity?

The bold conjecture is that the answer to the above question is positive.[1] The authors support this conjecture, for which several partial results provide evidence.

First, the conjecture holds for *quadratic equations*, that is, equations in which each variable occurs at most twice; this is the content of Proposition 13, formulated in different terms. Furthermore, the conjecture also holds for the case of *two-variable equations*, as can be

[1] Unfortunately, bold conjectures share the fate of pilots: there are many old pilots and many bold pilots, but not so many who are both old and bold.

derived from the description of their solution sets by Lucien Ilie and Wojciech Plandowski [22, Thm. 36]; see Proposition 7 for a restatement in our framework.

Additional support arises from group theory. In their work on Tarski's conjectures on the elementary theory of free groups, Olga Kharlampovich and Alexei Miasnikov [24], as well as Zlil Sela [36], employ induction arguments based on the JSJ-decomposition of limit groups. According to these authors (personal communication), the same techniques show that the conjecture holds for equations in free groups, although this result has not been published yet.

Finally, a weaker form of support comes from computational exploration: so far we have not been able to construct a counterexample, neither manually nor with computer assistance.

1.2 Historical Background

The *satisfiability problem for word equations* asks whether a given equation $U = V$ admits a solution. Interest in this problem first arose in the Russian school of mathematics, when it was observed that this problem can be reduced to *Hilbert's Tenth Problem*. Thus, proving the undecidability of the former would have simultaneously resolved the latter in the negative.

This program ultimately failed greatly with two major milestones in mathematics. In 1970, Yuri Matiyasevich demonstrated the undecidability of Hilbert's Tenth Problem using number-theoretic methods building on the earlier work of Martin Davis, Hilary Putnam, and Julia Robinson; see [28] for a detailed exposition. The resulting theorem is now referred to as the DPRM theorem, in accordance with Matiyasevich's preference.

The second milestone was achieved in 1977 by Gennadiĭ Makanin [26], who proved that the satisfiability problem for word equations is, in fact, decidable. A key ingredient in his proof is a result by Hmelevskiĭ [21], showing that a shortest solution of a word equation $U = V$ has an exponent of periodicity which is bounded in the size of the equation. Moreover, there is a computable number κ, depending only on $|UV|$, such that if a solution σ satisfies $\exp(\sigma) \geq \kappa$, then there are solutions σ_n with $\exp(\sigma_n) \geq n$ for all $n \in \mathbb{N}$; we then write $\exp(U = V) = \infty$.

Clearly, $\exp(U = V) = \infty$ implies the existence of infinitely many solutions. But what about the converse direction? This question behaves, to quote a phrase, like '*a complete unknown, like a rolling stone*': it is easy to state and easy to grasp, and – with numerous new tools for analyzing word equations having emerged since Makanin's result in 1977 – it might now be within reach to answer this question. However, it still tends to '*roll away*'.[2]

What are these new tools? First and foremost, *compression techniques* have greatly clarified why the satisfiability problem for word equations is decidable. This approach was initiated by Wojciech Plandowski and Wojciech Rytter [31] and simplified by Artur Jeż [23]. Moreover, compression techniques extend naturally to *word equations with regular constraints*, as studied by Klaus Schulz in his Habilitationsschrift [34].

Broadening the setting to include regular constraints has made the field more attractive to research that builds upon word equations. For instance, many modern works in *string solving* and *formal verification* incorporate word equations as part of their underlying framework.

Solving equations with regular constraints also opens the possibility of finding an equation that has infinitely many constrained solutions, yet – solely because of the constraints – the exponent of periodicity remains bounded by some fixed constant.

For example, the word equation $XabY = YbaX$ (over constants a, b and variables X, Y) is a promising candidate. Lucien Ilie and Wojciech Plandowski [22] devoted an entire section to describing the full solution set of this short, two-variable quadratic equation. See also the

[2] Or, in Salinger's terms, it resembles *The Catcher in the Rye.*

article of Weinbaum [38] for the description of the solution set of equations $XAY = YBX$ with $A \neq B$; and Problem 4 for more details.

On the other hand, the hope for positive results was kindled during a discussion with Wojciech Plandowski at a meeting held in Wittenberg, Germany, in September 2018. In that discussion, Plandowski referred to the equation $XabY = YbaX$ as the *'one and only'* known equation with infinitely many solutions that conceals its property $\exp(XabY = YbaX) = \infty$. His joke alluded to the fact that typical examples of word equations with infinitely many solutions arise from systems of integer linear equations. This *'one and only'* equation is, of course, the source of infinitely many similar examples.[3]

1.3 Our Contribution

In this article, we address the question posed above: whether a word equation with infinitely many solutions necessarily admits solutions of arbitrarily large exponent of periodicity. We study this problem in the setting of *word equations with regular constraints*.

To obtain positive results, we approach the problem by considering restricted classes of word equations and restricted classes of finite semigroups. Given a class of word equations $\mathcal{E}$, we call a finite semigroup S *nice* for $\mathcal{E}$ if every equation in $\mathcal{E}$ with constraints in S has infinitely many solutions if and only if it has solutions with arbitrarily large exponent of periodicity; see Definition 8. We show that, for every fixed class $\mathcal{E}$, the corresponding class of all finite semigroups that are nice for $\mathcal{E}$ enjoys several natural closure properties. Indeed, this class is always closed under taking subsemigroups and homomorphic images. Moreover, under mild assumptions on $\mathcal{E}$, it is also closed under taking left-right duals and forming retractive nilpotent extensions; see Lemma 9, Lemma 10, and Theorem 11 in Section 4.

As an immediate application of these closure properties, we obtain that all finite nilpotent semigroups are nice for the class of *two-variable word equations*; see Corollary 12.

Our main result concerns the class of all *quadratic word equations* – that is, equations in which each variable occurs at most twice – and members of the variety **DLG** of finite semigroups whose regular $\mathcal{D}$-classes are right groups (see Section 7). The following statement, the proof of which occupies Section 6, is formulated in Corollary 27.

▶ **Theorem 1.** *Let $S \in$ **DLG**. Then S is nice for all quadratic word equations.*

Our results thus supports the conjecture in a nontrivial setting, where additional regular constraints are permitted.

2 Notation and Preliminaries

By $\mathbb{N}$ we denote the set of natural numbers and by $\mathbb{N}_+$, the set of positive natural numbers.

We use standard notation of set theory. Frequently, we identify an element with the singleton containing this element. The restriction of a mapping $\psi : B \to C$ to some $A \subseteq B$ is still denoted ψ. The set of mappings from A to B is denoted by B^A. If $\varphi : A \to B$ and $\psi : B \to C$ are mappings then the composition $\psi\varphi : A \to C$ is defined by $(\psi\varphi)(a) = \psi(\varphi(a))$.

2.1 Semigroups and Monoids

Recall that a monoid is a semigroup with a neutral element, which we usually denote by 1. Every homomorphism between monoids has to map the neutral element to the neutral

[3] Similarly, the language $\{a^n b^n \mid n \in \mathbb{N}\}$ is the *'one and only'* context-free language which is not regular.

element. This is not required for semigroups. Thus, for monoids S and T there can be more semigroup homomorphisms from S to T than monoid homomorphisms.

The *empty* semigroup is a semigroup with no elements. A semigroup with exactly one element is called *trivial*. For a given semigroup S, we let $S^1 = S$ provided that S is a monoid. If S is not a monoid, then S^1 is obtained by adjoining a neutral element to S. Thus, $S^1 = S \cup \{1\}$. For every semigroup $S = (S, \cdot)$ there is a left-right dual semigroup $S^{\mathrm{op}} = (S, \circ)$, using the same underlying set S, but with its multiplication reversed, that is, $x \circ y = y \cdot x$. Note that, in contrast to groups, the semigroups S and S^{op} are not isomorphic, in general.

A *pseudovariety* is a class of semigroups which is is closed under forming subsemigroups, homomorphic images, and finite direct products.[4] Following Pin [30], we call a pseudovariety of finite semigroups simply a *variety of finite semigroups*. This facilitates the notation and there is no risk of confusion because if a variety contains a semigroup with at least two elements, it contains an infinite semigroup.

The *free monoid* over a set Γ is denoted by Γ^*. In this context Γ is called an *alphabet*, its elements are called *letters*, and the elements of Γ^* are called *words*. The empty word is denoted by 1 (as in other monoids). The *free semigroup* over a set Γ is $\Gamma^+ = \Gamma^* \setminus \{1\}$.

If $w \in \Gamma^*$ is a word, then $|w|$ denotes its length and $|w|_a$ denotes the number of occurrences of the letter $a \in \Gamma$ within w so that, in particular, $|w| = \sum_{a \in \Gamma} |w|_a$.

A word $u \in \Gamma^*$ is a *factor* of $w \in \Gamma^*$ if there are $p, q \in \Gamma^*$ such that $w = puq$. Moreover, if p or q can be chosen as the empty word, then u is a *prefix* or *suffix* of w, respectively. These notions generalize to arbitrary semigroups in the form of Green's relations.

2.2 Green's Relations

We assume that most of the readers are familiar with Green's relations. For convenience we still recall some basic facts using the same notation as, for example, in [18, Sec. 7.6].

Let S be a semigroup. For each $x \in S$ the left ideal S^1x, the right ideal xS^1, and the two-sided ideal S^1xS^1 are subsemigroups of S containing x. The concept of ideals leads to natural preorders $\leq_{\mathcal{L}}$, $\leq_{\mathcal{R}}$, $\leq_{\mathcal{J}}$ on S and their induced equivalence relations $\sim_{\mathcal{L}}$, $\sim_{\mathcal{R}}$, $\sim_{\mathcal{J}}$:

$$\begin{array}{llll} x \leq_{\mathcal{L}} y \Leftrightarrow & S^1x \subseteq S^1y & \quad x \sim_{\mathcal{L}} y \Leftrightarrow & S^1x = S^1y \\ x \leq_{\mathcal{R}} y \Leftrightarrow & xS^1 \subseteq yS^1 & \quad x \sim_{\mathcal{R}} y \Leftrightarrow & xS^1 = yS^1 \\ x \leq_{\mathcal{J}} y \Leftrightarrow & S^1xS^1 \subseteq S^1yS^1 & \quad x \sim_{\mathcal{J}} y \Leftrightarrow & S^1xS^1 = S^1yS^1 \end{array}$$

These are three out of five relations referred to as *Green's relations*. Instead of $\sim_{\mathcal{L}}$, $\sim_{\mathcal{R}}$, and $\sim_{\mathcal{J}}$ we also write $\mathcal{L}$, $\mathcal{R}$, and $\mathcal{J}$ for the corresponding relations on S. The other two relations are then defined by $\mathcal{H} = \mathcal{L} \cap \mathcal{R}$ and $\mathcal{D} = \mathcal{L} \circ \mathcal{R} = \mathcal{R} \circ \mathcal{L}$. For any $\mathcal{K} \in \{\mathcal{H}, \mathcal{L}, \mathcal{R}, \mathcal{D}, \mathcal{J}\}$ we denote the $\mathcal{K}$-class of an element $x \in S$ by $\mathcal{K}(x) = \{y \in S \mid x \sim_{\mathcal{K}} y\}$.

An important fact about finite semigroups is that they satisfy $\mathcal{D} = \mathcal{J}$; thus, all $\mathcal{D}$-classes are $\mathcal{J}$-classes and vice versa. Moreover, if S is finite, then there is some $\omega \in \mathbb{N}_+$ such that that x^ω is idempotent for all $x \in S$, where an element $e \in S$ is called *idempotent* if $e^2 = e$.

An element x in a semigroup S is called *regular* if there exists $y \in S$ with $xyx = x$, and an $\mathcal{L}$-, $\mathcal{R}$, or $\mathcal{D}$-class is called *regular* if it contains an idempotent element. The following is a well-known fact; for a standard proof (for finite semigroups) see e.g. [18, Prop. 7.49].

▶ **Proposition 2.** *Let L, R, and D be $\mathcal{L}$-, $\mathcal{R}$, and $\mathcal{D}$-classes of a semigroup S, respectively. Further, suppose that $L, R \subseteq D$. Then the following assertions are equivalent.*

[4] Consequently, every variety of finite semigroups contains the trivial semigroup and the empty one.

(1) *The $\mathcal{D}$-class D is regular.*
(2) *The $\mathcal{L}$-class L is regular.*
(3) *The $\mathcal{R}$-class R is regular.*
(4) *Every element in D is regular.*

2.3 Rational, Recognizable, and Regular Sets

We now turn to rational, recognizable, and regular sets. All facts discussed about them in this subsection admit simple proofs, which can be found, for example, in Eilenberg's classical textbook [19] or in other textbooks, sometimes stated for monoids, as in [18, Sec. 7].

Let S be an arbitrary semigroup. Its family of *rational sets* is the family $\mathrm{Rat}(S)$ of subsets of S that is defined inductively as follows.

- If $L \subseteq S$ is finite, then L belongs to $\mathrm{Rat}(S)$.
- If $L_1, L_2 \in \mathrm{Rat}(S)$, then $L_1 \cup L_2$ and $L_1 \cdot L_2 = \{u_1 \cdot u_2 \mid u_1 \in L_1, u_2 \in L_2\}$ are in $\mathrm{Rat}(S)$.
- If $L \in \mathrm{Rat}(S)$, then the subsemigroup $L^+ = L \cup L \cdot L \cup \dots$ generated by L is in $\mathrm{Rat}(S)$.

There is a convenient alternative definition of $\mathrm{Rat}(S)$ by using nondeterministic automata.

A *nondeterministic S-automaton*, or simply an *S-automaton* for short, is a directed graph $\mathcal{A}$ with sets of initial and final vertices, and where arcs are labeled by elements of S.

- The vertices of $\mathcal{A}$ are called *states*, and the set of all states of $\mathcal{A}$ is typically denoted by Q. There are distinguished sets $I \subseteq Q$ of *initial states*, and $F \subseteq Q$ of *final states*.
- The directed labeled arcs of $\mathcal{A}$ are called *transitions*. Each transition can be written as a triple $(p, s, q) \in Q \times S \times Q$ and depicted as $p \xrightarrow{s} q$. Given any path $p_0 \xrightarrow{s_1} p_1 \cdots \xrightarrow{s_n} p_n$ of transitions in $\mathcal{A}$, we define its label to be the element $s_1 \cdots s_n \in S^1$. Note that this label is an element of S unless $n = 0$, that is, unless the path is an *empty path*.
- The *accepted language* $L(\mathcal{A})$ of $\mathcal{A}$ is defined as follows.

$$L(\mathcal{A}) = \{s \in S \mid \text{there is a path labeled } s \text{ from an initial to a final state in } \mathcal{A}\}$$

An S-automaton $\mathcal{A}$ with finitely many transitions is called a *nondeterministic finite S-automaton*, or *S-NFA* for short. The family $\mathrm{Rat}(S)$ coincides with the family of all languages which are accepted by S-NFA.

It is sometimes convenient to only consider those states of an S-automaton that directly contribute to its accepted language. An S-automaton is called *trim* if every state is on some directed path from an initial to a final state.

For a semigroup S there is also the family $\mathrm{Rec}(S)$ of *recognizable sets* of S. It consists of those subsets $L \subseteq S$ that are *recognized* by some homomorphism $\mu : S \to T$ to a finite semigroup T, which means that $L = \mu^{-1}(\mu(L))$.

- A semigroup S is finitely generated if and only if $\mathrm{Rec}(S) \subseteq \mathrm{Rat}(S)$.
- If $\varphi : S \to S'$ is a homomorphism, then $L' \in \mathrm{Rec}(S')$ implies $\varphi^{-1}(L') \in \mathrm{Rec}(S)$.
- The family $\mathrm{Rec}(S)$ is a Boolean algebra with respect to the standard set operations.
- If S contains the free product $\mathbb{N}_+ * (\mathbb{N}_+ \times \mathbb{N}_+)$, then $\mathrm{Rat}(S)$ is not closed under finite intersection.[5] In particular, this implies $\mathrm{Rat}(S) \neq \mathrm{Rec}(S)$.

The descriptions of full solution sets for word equations (in our and several other papers) use rational sets in the monoid of endomorphisms $\mathrm{End}(\Gamma^+)$. For $|\Gamma| \geq 2$ this monoid is not finitely generated, it is not torsion free, and it contains $\mathbb{N}_+ * (\mathbb{N}_+ \times \mathbb{N}_+)$ as a subsemigroup.[6]

[5] In fact, closure of $\mathrm{Rat}(S)$ under finite intersection should be viewed as a rare exception.
[6] A possible realization of $\mathbb{N}_+ * (\mathbb{N}_+ \times \mathbb{N}_+)$ uses three endomorphisms α, β and γ. Here, γ is defined by $a \mapsto ab$ and $b \mapsto ba$, α is defined by $a \mapsto a^2$ and $b \mapsto b$, and β interchanges the role of a and b in α.

Kleene's subset construction [25] shows that for every finitely generated free semigroup $S = \Gamma^+$ (resp. for $S = \Gamma^*$) it holds that $\mathrm{Rec}(S) = \mathrm{Rat}(S)$. In this case, the sets in $\mathrm{Reg}(S) = \mathrm{Rat}(S)$ are called *regular sets*, and we also write $\mathrm{Reg}(S)$ for the family of all such sets. We use this terminology only when Γ is finite. The regular sets of finitely generated free semigroups and monoids are Boolean algebras and related by

$$\mathrm{Reg}(\Gamma^+) = \{L \in \mathrm{Reg}(\Gamma^*) \mid 1 \notin L\},\ \mathrm{Reg}(\Gamma^*) = \mathrm{Reg}(\Gamma^+) \cup \{L \cup \{1\} \subseteq \Gamma^* \mid L \in \mathrm{Reg}(\Gamma^+)\}.$$

3 Word Equations

In this paper, we deal with systems of word equations with regular constraints over finitely generated free monoids and semigroups. We begin with the monoid case.

Word Equations over Free Monoids. Let Σ and $\mathcal{X}$ be two disjoint finite sets. We say that Σ is the *alphabet of constants* and $\mathcal{X}$ is the set of *variables*; by $\Omega = \Sigma \cup \mathcal{X}$ we mean their (disjoint) union. A *system of word equations (with regular constraints)* over $\Omega^* = (\Sigma \cup \mathcal{X})^*$ is a pair $(\mathcal{S}, \mu)$ where $\mathcal{S} = \{(U_1, V_1), \ldots, (U_n, V_n)\}$ is a finite set of *word equations* (U_i, V_i) – that is, $U_i, V_i \in \Omega^*$ – and $\mu : \Omega^* \to N$ is a morphism to a finite monoid N.

A *solution* at $E = (\mathcal{S}, \mu)$ is given by a mapping $\sigma : \mathcal{X} \to \Sigma^*$ such that its extension to a morphism $\sigma : \Omega^* \to \Sigma^*$, leaving the constants invariant, satisfies the following conditions.

(1) We have $\sigma(U) = \sigma(V)$ for all $(U, V) \in \mathcal{S}$.

(2) We have $\mu\sigma(X) = \mu(X)$ for all $X \in \mathcal{X}$; that is, solutions have to respect the constraints.

The set of all solutions at $E = (\mathcal{S}, \mu)$ is denoted by $\mathrm{Sol}(E) = \mathrm{Sol}(\mathcal{S}, \mu)$.

A solution σ is called *singular* if $\sigma(X) = 1$ for some variable $X \in \mathcal{X}$. Otherwise σ is called *nonsingular*. For each $\sigma \in \mathrm{Sol}(E)$ there is a subset $\mathcal{X}_{\mathrm{sing}}$ such that $\sigma(X) = 1 \Leftrightarrow X \in \mathcal{X}_{\mathrm{sing}}$. By removing all variables in $\mathcal{X}_{\mathrm{sing}}$ from the system E (that is, replacing all occurrences in each U_i and V_i by the empty word) the solution σ can be made nonsingular. This leads to the following nondeterministic procedure: we guess a subset $\mathcal{X}' \subseteq \mathcal{X}$, and then we remove all variables in $\mathcal{X}'$ from the equations and the set of variables; simultaneously, we strengthen the requirement that all solutions at the resulting system $E_{\mathcal{X}'}$ must be nonsingular. The procedure is sound and complete: we have $\mathrm{Sol}(E) \neq \emptyset$ (resp. $|\mathrm{Sol}(E)| = \infty$) if and only if $\mathrm{Sol}(E_{\mathcal{X}'}) \neq \emptyset$ (resp. $|\mathrm{Sol}(E_{\mathcal{X}'})| = \infty$) holds for one of these guesses $E_{\mathcal{X}'}$. Moreover, E has infinite exponent of periodicity (according to Definition 4 below) if and only if one of these systems does. Thus, for the purposes of this article it is enough to work with systems of word equations over free semigroups instead.

Word Equations over Free Semigroups. A *system of word equations with regular constraints* over $\Omega^+ = (\Sigma \cup \mathcal{X})^+$ is a pair $(\mathcal{S}, \mu)$ where $\mathcal{S} \subseteq \Omega^+ \times \Omega^+$ is a finite set of *word equations*, and where $\mu : \Omega^+ \to S$ a morphism to a finite semigroup S. A *solution* at $E = (\mathcal{S}, \mu)$ is given by a mapping $\sigma : \mathcal{X} \to \Sigma^+$ with the same properties (1) and (2) as in the monoid case.

As just explained, the monoid setting can be reduced to the semigroup setting, but the semigroup setting has the advantage that we can use constraints in finite semigroups without neutral elements, which will allow us to state some of our results in a more natural way.[7]

In order to reduce equations over monoids to equations over semigroups, we have already implicitly used one of the central tools in the study of word equations, namely the following.

[7] In some sense, the theory of finite semigroups is indeed richer than the one of finite monoids.

▶ **Definition 3.** *A* substitution *for* $\Omega = \Sigma \cup \mathcal{X}$ *is an endomorphism* τ *of* Ω^+ *leaving the constants in* Σ *invariant, and therefore defined by its restriction to* $\mathcal{X}$. *We say that a substitution is* defined by *a mapping* $X \mapsto \tau(X)$ *if, for some* $X \in \mathcal{X}$, $\tau(\alpha) = \alpha$ *for all* $\alpha \in \Omega \smallsetminus \{X\}$. *A* basic substitution *is a substitution defined by* $X \mapsto \tau(X)$ *with* $\tau(X) \in \Omega X \cup \Sigma$. ◇

Note that (compositions of) basic substitutions are powerful enough to compute every solution σ, but *cave canem*: in general, basic substitutions increase the length of the equation.

3.1 Reducing Systems of Equations

Frequently it is convenient to encode a system of word equations $\mathcal{S} = \{(U_1, V_1), \dots, (U_n, V_n)\}$ with constraint $\mu : \Omega^+ \to S$ by a single word equation. In order to do so, we first extend the set of constants Σ by introducing a new constant $\# \notin \Sigma$. Let $\tilde{\Sigma} = \Sigma \cup \{\#\}$ be this new set of constants, and let $\tilde{\Omega} = \tilde{\Sigma} \cup \mathcal{X}$. By adjoining a new zero element 0 to S, we obtain a semigroup $\tilde{S} = S \cup \{0\}$ and we extend μ to a homomorphism $\tilde{\mu} : \tilde{\Omega}^+ \to \tilde{S}$ with $\tilde{\mu}(\#) = 0$. Having this, we finally replace $(\mathcal{S}, \mu)$ by $((U_1 \# \cdots U_n \#, V_1 \# \cdots V_n \#), \tilde{\mu})$, which has the same solutions.

Note that the above encoding does not create any new variables, and it makes sure that the total number of occurrences of each variable is the same as before. Thus, this construction allows us to subsequently focus on the case of a single word equation.

Henceforth, we also write a word equation (U, V) in the more readable form as $U = V$.

3.2 The Exponent of Periodicity

The exponent of periodicity is a well-known concept in combinatorics on words. Formally, the *exponent of periodicity* $\exp(w)$ of a word $w \in \Sigma^*$, is the greatest number such that there is a factorization $w = u\, p^{\exp(w)}\, v$ with $p \in \Sigma^+$. Note that $\exp(w) = 0$ if and only if $w = 1$.

▶ **Definition 4.** *Let* $E = (U = V, \mu)$ *be a word equation (over a free monoid or semigroup) with a regular constraint* μ. *The* exponents of periodicity *of a solution* σ *at* E, *which we denote by* $\exp(\sigma)$, *and of* E *itself, denoted by* $\exp(E) = \exp(U = V, \mu)$, *are defined as follows.*

- *We let* $\exp(\sigma)$ *be the maximum* $\max\{\exp(\sigma(X)) \mid X \in \mathcal{X}\} \in \mathbb{N}$.
- *We let* $\exp(E)$ *be the supremum* $\sup\{\exp(\sigma) \mid \sigma \in \mathrm{Sol}(E)\} \in \mathbb{N} \cup \{\infty\}$. ◇

The next proposition generalizes a classical result of Hmelevskiĭ [21]. His proof is based on p-stable normal forms, which were also used by Makanin [26] and Schulz [34] for solving word equations with regular constraints. The description of Makanin's algorithm in [13] contains a proof of Proposition 5 with explicit bounds, even in case S is part of the input.

▶ **Proposition 5.** *Let* $E = (U = V, \mu)$ *be a word equation with a regular constraint* μ. *Then there is a constant* κ *such that we have* $\exp(E) > \kappa$ *if and only if there is some* $X \in \mathcal{X}$ *and a nonempty word* $p \in \Sigma^+$ *such that for each* $n \in \mathbb{N}$ *there is a solution* σ_n *at* E *where* p^n *appears as a factor in* $\sigma_n(X)$. *In particular, if* $\exp(E) > \kappa$, *then* $\exp(E) = \infty$ *and* $|\mathrm{Sol}(E)| = \infty$. *Moreover, if* μ *maps to a fixed finite semigroup* S, *then we can choose* $\kappa \in 2^{\mathcal{O}(|UV| \cdot |\Omega|)}$.

The above result has the following noteworthy application.

▶ **Corollary 6** (Plandowski, Schubert [32])**.** *Let* $E = (\mathcal{S}, \mu)$ *be any systems of word equations with a regular constraint* $\mu : \Omega^+ \to S$. *If the finite semigroup* S *is fixed and not part of the input, then there is a* PTIME*-reduction of the problem to decide on input of* E *whether* $\exp(E) = \infty$ *to the satisfiability problem of word equations, that is, to decide on input of* E *whether there exists a solution at* E.

Proof. As explained in Section 3.1, we can encode a system by a single equation $(U = V, \mu)$. It is shown in [13] that that we can choose a concrete value $m \in \mathcal{O}(|UV| \cdot |\Omega|)$ so that $\kappa < 2^m$ holds for the constant κ in Proposition 5. Knowing m, we consider the variables $X \in \mathcal{X}$ one after another. For each $X \in \mathcal{X}$, we create a system of word equations with fresh variables $X_1, \dots, X_m, Y, Z$, and additional equations $X = YX_1Z$ and $X_{i-1} = X_iX_i$ for $1 \le i \le m$.

Having this new system we decide whether it has a solution. If the answer is negative for every variable $X \in \mathcal{X}$, then $\exp(E) < \infty$. Otherwise, $\exp(E) = \infty$ by Proposition 5.[8] ◀

3.3 Two-Variable Word Equations

A *two-variable word equation* is a word equation $U = V$ over a set of variables $\mathcal{X} = \{X, Y\}$. They were studied deeply in e.g. [22] and further in [11]. The main structural result about the full solution set for two-variable word equations is Theorem 36 in [22]. Its proof is a careful case analysis. Inspecting these cases shows the following.

▶ **Proposition 7** (Ilie, Plandoswski). *If a two-variable word equation $U = V$ has infinitely many solutions, then* $\exp(U = V) = \infty$.

4 Nice Semigroups

As mentioned in the introduction, we are interested in when $|\mathrm{Sol}(U = V, \mu)| = \infty$ implies an infinite exponent of periodicity for a given word equation with constraints $(U = V, \mu)$. This naturally leads to the following definition.

▶ **Definition 8.** *A finite semigroup S is* nice *for a family of word equations $\mathcal{E}$, if the following holds for all equations $(U = V) \in \mathcal{E}$ over $\Omega^+ = (\Sigma \cup \mathcal{X})^+$ and constraints $\mu : \Omega^+ \to S$.*

If $(U = V, \mu)$ has infinitely many solutions, then $\exp(U = V, \mu) = \infty$. ◇

An immediate consequence of this definition is that a finite semigroup S is nice for the family of word equations $\mathcal{E}$ if and only if its left-right dual S^{op} is nice for the family $\mathcal{E}^{\mathrm{rev}}$ consisting of all reversed equations $(U^{\mathrm{rev}} = V^{\mathrm{rev}})$ with $(U = V) \in \mathcal{E}$. In particular, if $\mathcal{E} = \mathcal{E}^{\mathrm{rev}}$, as for quadratic or two-variable equations, then S is nice for $\mathcal{E}$ if and only if S^{op} is.

4.1 Divisors of Nice Semigroups

The following oberservations, Lemma 9 and Lemma 10, show that the class of all finite semigroups which are nice for any given family of word equations $\mathcal{E}$ is closed under taking *divisors* (that is, homomorphic images of subsemigroups).

▶ **Lemma 9.** *If S is a nice semigroup for $\mathcal{E}$, then so is every subsemigroup S' of S.*

Proof. Let S' be a subsemigroup of S, and let $\iota : S' \to S$ be the corresponding inclusion. Then, for all word equation $(U = V)$ over $\Omega^+ = (\Sigma \cup \mathcal{X})^+$ and all constraints $\mu' : \Omega^+ \to S'$, the solution sets $\mathrm{Sol}(U = V, \mu')$ and $\mathrm{Sol}(U = V, \iota\mu')$ coincide. Hence, $(U = V, \mu')$ has infinitely many solutions (resp. infinite exponent of periodicity) if and only if $(U = V, \iota\mu')$ does. ◀

▶ **Lemma 10.** *If S is a nice semigroup for $\mathcal{E}$, then so is every homomorphic image S' of S.*

[8] The standard way to extend the result to free monoids is by guessing a constant a and introducing a new equation $X_1 = aX_0$ where X_0 is another fresh variable.

Proof. Let $\rho : S \to S'$ be an epimorphism. Further, suppose that $(U = V) \in \mathcal{E}$ is a word equation over $\Omega^+ = (\Sigma \cup \mathcal{X})^+$ such that $|\mathrm{Sol}(U = V, \mu')| = \infty$ where $\mu' : \Omega^+ \to S'$. Let us first note that for every $\sigma \in \mathrm{Sol}(U = V, \mu')$ there is some constraint $\mu_\sigma : \Omega^+ \to S$ with $\mu' = \rho\mu_\sigma$ such that $\sigma \in \mathrm{Sol}(U = V, \mu_\sigma)$. Indeed, choosing $\mu_\sigma(a) \in \rho^{-1}(\mu'(a))$ arbitrarily for all $a \in \Sigma$ and setting $\mu_\sigma(X) = \mu_\sigma(a_1)\mu_\sigma(a_2)\cdots\mu_\sigma(a_n)$ where $\sigma(X) = a_1a_2\cdots a_n$ with $a_i \in \Sigma$ for all $X \in \mathcal{X}$ defines such a constraint μ_σ. Since there exist only a finite number of constraints $\mu : \Omega^+ \to S$, we must have $|\mathrm{Sol}(U = V, \mu)| = \infty$ for some μ with $\rho\mu = \mu'$ by the pigeonhole principle.

Assuming that S is nice for $\mathcal{E}$, we deduce that $\exp(U = V, \mu) = \infty$. In turn, this implies $\exp(U = V, \mu') = \infty$, as required, since $\mathrm{Sol}(U = V, \mu)$ is contained in $\mathrm{Sol}(U = V, \mu')$. ◀

4.2 Extensions of Nice Semigroups

An *extension* of a semigroup S is a semigroup T that contains S as an ideal. Such an extension is *retractive* if there exists a homomorphism $\rho : T \to S$ with $\rho(x) = x$ for all $x \in S$; that is, if S is a retract of T. It is *nilpotent* if $T^k \subseteq S$ for some k or, equivalently, if the Rees quotient T/S (in which all elements of S are identified) is a nilpotent semigroup.

Recall the notion of (basic) substitution in Definition 3. A substitution τ, thus an endomorphism of Ω^+ leaving Σ invariant, is called *trivial* if for all $X \in \mathcal{X}$ we have $\tau(X) \in \Sigma^* X \cup \Sigma^+$. Note that every trivial substitution can be written as a sequence of basic substitutions.

▶ **Theorem 11.** *Let $\mathcal{E}$ be a family of word equations which is closed under trivial substitutions. Further, let T be a finite semigroup which is a retractive nilpotent extension of a semigroup S. Then S is a nice semigroup for $\mathcal{E}$ if and only if T is a nice semigroup for $\mathcal{E}$.*

Proof. If T is a nice for $\mathcal{E}$, then its subsemigroup (and homomorphic image) S is nice for $\mathcal{E}$ by Lemma 9 (or Lemma 10). Thus, it suffices to show that T nice for $\mathcal{E}$ assuming that S is so.

Suppose that the word equation $(U = V) \in \mathcal{E}$ with constraints $\mu : \Omega^+ \to T$ has infinitely many solutions; that is, $|\mathrm{Sol}(U = V, \mu)| = \infty$. We will show that this implies $\exp(U = V, \mu) = \infty$.

Let us call a variable X *inessential* if, for some fixed $w \in \Sigma^+$, the equation $E = (U = V, \mu)$ has infinitely many solutions σ with $\sigma(X) = w$. Note that we can simply eliminate such a variable X by following the trivial substitution defined by $X \mapsto w$, since E has infinite exponent of periodicity whenever the new word equation does so. Hence, we may as well assume that none of the variables are inessential. For every $k \in \mathbb{N}$, all but finitely many of the solutions σ of the word equation E must therefore satisfy $|\sigma(X)| > k$ for all $X \in \mathcal{X}$.

Next, let us fix a number k large enough so that $T^{k+1} \subseteq S$. From the above we see that $\mu(X) \in S$ for every $X \in \mathcal{X}$. Let us also fix a retraction $\rho : T \to S$, and consider the word equation $E_\rho = (U = V, \rho\mu)$ with constraints in S. As every solution of E is a solution of E_ρ, the latter has infinitely many solutions σ with $|\sigma(X)| > k$ for all $X \in \mathcal{X}$. Note that, conversely, every solution σ of E_ρ with $|\sigma(X)| > k$ for all $X \in \mathcal{X}$ is also a solution of E. We claim that there are such solutions with arbitrarily large exponent of periodicity.

Indeed, by the pigeonhole principle, there exist words $w_X \in \Sigma^k$ such that $\sigma(X) \in w_X\Sigma^+$ holds for all $X \in \mathcal{X}$ for infinitely many solutions σ. Hence, some $E'_\rho = (U' = V', \mu') \in \mathcal{E}$ with constraints in S and infinitely many solutions can be obtained from E_ρ by the trivial substitution $X \mapsto w_X X$. By assumption on S there are solutions σ' of E'_ρ with arbitrarily large exponent of periodicity. The corresponding solutions σ of E_ρ with $\sigma(X) = w_X\sigma'(X)$ have the desired property since $|\sigma(X)| = |w_X| + |\sigma'(X)| \geq k + 1$ holds for all $X \in \mathcal{X}$. ◀

In order to apply Theorem 11 we need at least one nonempty finite semigroup S which is nice for a family of word equations $\mathcal{E}$. By Lemma 10 (or Lemma 9) the trivial semigroup would then be nice for $\mathcal{E}$ and, thus, so would all finite nilpotent semigroups by the above. A first example are two-variable word equations which we briefly discussed in Section 3.3. Based on Proposition 7 we can state a more general form of this proposition by using Theorem 11.

▶ **Corollary 12.** *Every finite nilpotent semigroup is nice for two-variable word equations.*

Unfortunately we were not able to prove that the trivial semigroup is nice for the class of all word equations over free semigroups. However, we will show that this is indeed also the case for the family of quadratic word equations, which we consider in the next section.

5 Quadratic Word Equations

A system $\mathcal{S} = \{(U_1, V_1), \ldots, (U_n, V_n)\}$ of word equations is called *quadratic* if every variable occurs at most twice in its equations, that is, $|U_1 \cdots U_n V_1 \cdots V_n|_X \leq 2$ holds for all $X \in \mathcal{X}$. As explained in Section 3.1, it will suffice for our purposes to deal with (single) quadratic word equations $U = V$. Throughout, we denote the class of all quadratic word equations by $\mathcal{QE}$.

The following Proposition 13 states that the trivial semigroup is nice for the family of quadratic word equations. This result was initially worked out in 2019 by Bastien Laboureix in his (unpublished) stage report for the Laboratoire Spécification et Vérification (now at the ENS Paris-Saclay, France) when he was on leave as part of a collaboration between Benedikt Bollig and the first author. The approach of Laboureix was based on ideas of Olga Kharlampovich (personal communication). The following proof, however, is from the bachelor's thesis of the second author [29].

▶ **Proposition 13.** *The trivial semigroup is nice for the family of quadratic word equations.*

Proof. Let $U = V$ be an equation without constraints and with infinitely many solutions. This implies $|U| > 1$ and $|V| > 1$. Moreover, we may assume that there infinitely many solutions σ where, for all $X \in \mathcal{X}$, we have $|\sigma(X)| > 1$. Indeed, if this is not the case, then there is some $X \in \mathcal{X}$ such that there are infinitely many solutions σ with $\sigma(X) = c_X$ for some fixed constant $c_X \in \Sigma$. Substituting $X \mapsto c_X$ then reduces the number of variables which appear in the equation, and the claim therefore follows by induction.

We now distinguish between several types of variables. The set of variables $X \in \mathcal{X}$ which appear i times in UV are denoted by $\mathcal{X}_i$; thus, $\mathcal{X} = \mathcal{X}_0 \cup \mathcal{X}_1 \cup \mathcal{X}_2$. Moreover, we say that a variable $X \in \mathcal{X}_2$ is *balanced* if it appears on both sides of the equation, that is, $|U|_X = |V|_X = 1$. Otherwise, the variable $X \in \mathcal{X}_2$ is called *unbalanced*.

In order to prove the claim we may assume without restriction that $\mathcal{X}_0 = \emptyset$. By symmetry we may also assume that $U = XU'$ and $V = \alpha V'$ with $X \in \mathcal{X}$. If $\alpha \in \Sigma$, then we are essentially forced to continue with a substitution $X \mapsto \alpha X$. Doing so repeatedly, while keeping the invariant that infinitely many solutions σ satisfy $|\sigma(Z)| > 1$ for all $Z \in \mathcal{X}$, we eventually arrive at the situation where $U = XU'$ and $V = YV'$ with $X, Y \in \mathcal{X}$. If $X = Y$, then we can remove both occurrences of X from the equation. Similarly, if $X \neq Y$, but $|\sigma(X)| = |\sigma(Y)|$ for infinitely many solutions σ, then the substitution defined by $X \mapsto Y$, and subsequently removing the leading occurrences of Y, allows us to make the equation shorter. In either case, we the claim then follows by induction. By symmetry we may thus assume that there are infinitely many solutions σ with $|\sigma(Y)| < |\sigma(X)|$. Therefore, there are also infinitely many solutions σ at $X\tau(U') = \tau(V')$ where τ is defined by $\tau(X) = YX$.

The crucial observation is that we have $\exp(U = V) = \infty$ as soon as X or Y is balanced. Indeed, fix any solution σ for $XU' = YV'$, and assume that Y is balanced. Then we have the equality $\sigma(XU') = v\sigma(Y)w = \sigma(YV')$ for some $v \in \Sigma^+$ and $w \in \Sigma^*$. Hence, for all $n \in \mathbb{N}$, there is a solution σ_n of $XU' = YV'$ by $\sigma_n(Y) = v^n\sigma(Y)$ and $\sigma_n(Z) = \sigma(Z)$ for $Z \neq Y$. As $|v| \geq 1$, this shows that $\exp(U = V) = \infty$. The case in which X is balanced is similar. It thus remains to show that we always reach an equation where one sides begins with a balanced variable.

If $X \in \mathcal{X}_1$, then we have $|X\tau(U')\tau(V')| < |XU'YV'|$, and we are done by induction. Therefore, let $X \in \mathcal{X}_2$ and, moreover, let X be unbalanced. As long as the right side of the equation begins $\alpha \in \Omega$ with $\alpha \in \Sigma \cup \mathcal{X}_1$, we are forced to follow the substitution $X \mapsto \alpha X$, for otherwise we would be done by induction on the length of the equation. However, this can only continue for at most $|V|$ steps, since $|V|$ decreases after every such substitution as both occurrences of X are in U. Thus, at some point, V must begin with a variable $Y \in \mathcal{X}_2$. If Y is balanced, then we are done. However, if Y is unbalanced, then both possible substitutions $X \mapsto YX$ and $Y \mapsto XY$ decrease the number of unbalanced variables, since they make Y (respectively X) balanced. The number of unbalanced variables cannot increase again, unless we reach an equation where one side begins with a balanced variable and the other side with an unbalanced variable. This finishes the proof. ◀

An immediate application of Theorem 11 yields the following.

▶ **Corollary 14.** *Every finite nilpotent semigroup is nice for the family $\mathcal{QE}$.*

5.1 The Matiyasevich Solution Graph

In 1968, Matiyasevich [27] defined a solution graph of linear size in the input for showing that the satisfiability problem for a quadratic equations without constraints and without involution is decidable in PSPACE for free semigroups and for free monoids. The same approach copes with an involution and regular constraints. It also provides a simple method to decide whether or not a quadratic equation has infinitely many solutions.

Let $n \in \mathbb{N}$ and let S be a fixed semigroup. The *Matiyasevich Solution Graph* for the given n is defined as an edge-labeled directed graph $\mathcal{A}_n$. The vertices are called *states* and the edges, labeled by basic substitutions according to Definition 3, are called *transitions*.

Each state is a triple $(U = V, \mathcal{X}, \mu)$ with $|UV| \le n$ and $|\mathcal{X}| \le n$. The state is called *final*, if $U = a = V$ for some $a \in \Sigma$ and if $\mathcal{X} = \varnothing$. Next we describe the transitions.

We start with state of the form $(\alpha U = \alpha V, \mathcal{X}, \mu)$ with $\alpha \in \Omega$. We then have a unique outgoing ε-transition, labeled by the identity, to the state $(U = V, \mathcal{X}, \mu)$.

Every other state $(U = V, \mathcal{X}, \mu)$ always has the following two outgoing transitions for every $X \in \mathcal{X}$ which does not appear in UV and for every $a \in \Sigma$.

- A transition labeled $\tau(X) = aX$ to the state $(U = V, \mathcal{X}, \mu')$.
- A transition labeled $\tau(X) = a$ to the state $(U = V, \mathcal{X} \setminus \{X\}, \mu')$

Moreover, if a state can be written as $(XU = \alpha V, \mathcal{X}, \mu)$ then we have the following two additional outgoing transitions.

- A transition labeled $\tau(X) = \alpha X$ to the state $(X\tau(U) = \tau(V), \mathcal{X}, \mu')$.
- A transition labeled $\tau(X) = \alpha$ to the state $(\tau(U) = \tau(V), \mathcal{X} \setminus \{X\}, \mu')$.

The constraints μ' are always chosen such that $\mu = \mu'\tau$. If this is not possible, the corresponding transition does not exist. If μ' can be chosen in several different ways, then multiple copies of the respective transition are introduced.

Given a quadratic equation $(U = V, \mu)$ with constraints and $|UV| \le n$, we let $\mathcal{X}$ be the set of variables occurring in UV and set $E = (U = V, \mathcal{X}, \mu)$. Note that E is a state in $\mathcal{A}_n$. We can construct an NFA $\mathcal{A}(E)$ from $\mathcal{A}_n$ as follows. First, we declare E to be the unique initial state, keeping the final states already defined, and then we trim the automaton accordingly. The resulting automaton $\mathcal{A}(E)$ is called the *(Matiyasevitch) solution graph* of E.

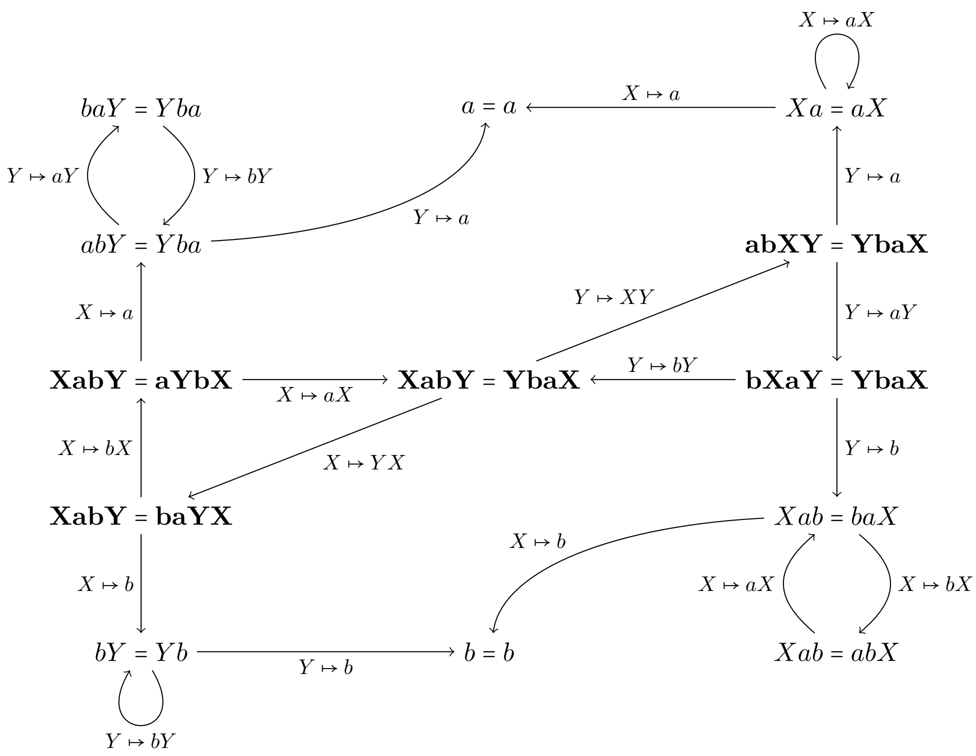

Figure 1 The solution graph of the equation $XabY = YbaX$ without constraints.

▶ **Example 15.** The diagram in Figure 1 shows the solution graph of the quadratic equation $XabY = YbaX$ without constraints where some ε-transitions are omitted for clarity. Unsolvable states do not appear as the automaton is trim. The equations in boldface form the strongly connected component where both X and Y appear in the equation.

▶ **Proposition 16.** *On input $n \in \mathbb{N}$ we can effectively construct its Matiyasevich Solution Graph, that is the graph $\mathcal{A}_n$. Let E be a state of $\mathcal{A}_n$. Then the following two properties hold.*
- *The automaton $\mathcal{A}(E)$ accepts the language* $\mathrm{Sol}(E)$*.*
- *We have $|\mathrm{Sol}(E)| = \infty$ if and only if $\mathcal{A}(E)$ contains some nontrivial connected component.*

Proof. Let π be an accepting path in $\mathcal{A}(E)$. The path π is labeled an endomorphism $\tau_1 \cdots \tau_m$ where each τ_i is a substitution; and π defines a solution σ such that $\sigma(X) = \tau_m \cdots \tau_1(X)$ for all $X \in \mathcal{X}$. By induction of the length of a solution, it is easy to see that every solution has such a factorization. Since $\mathcal{A}(E)$ is finite, we obtain that $|\mathrm{Sol}(E)| = \infty$ if and only if there is an accepting path using a transition within some strongly connected component of $\mathcal{A}(E)$. ◀

▶ Remark 17. Both problems related to the questions in the next Corollary 18 remain decidable for all word equations with regular constraints in free monoids or semigroups (with or without involution) and extend to free groups with rational constraints in [17] and [8]. There are also extensions to virtually free groups [14] and, even further, to hyperbolic groups [9]. All these papers are quite technical. Dealing with quadratic equations makes life nicer, as the proof follows directly from properties of the Matiyasevich Solution Graph $\mathcal{A}_n$. ◇

The next corollary is a direct consequence of Proposition 16 since we can encode a system of quadratic word equations as a single quadratic equation as in Section 3.1.

▶ **Corollary 18.** *On input of a system $\mathcal{S}$ of quadratic word equations over $\Omega = \Sigma \cup \mathcal{X}$, it is decidable whether $(\mathcal{S}, \mu)$ has some solution, and whether $(\mathcal{S}, \mu)$ has infinitely many solutions.*

6 The Main Result

Using the Matiyasevitch solution graph defined in Section 5, we will prove our main result on the exponent of periodicity of quadratic equations with certain constraints (which are defined in Section 6.1, and discussed in more detail in Section 7). To do so, we first introduce a special type of equation for which the desired claim is particularly easy to establish. The general case of quadratic equations is then reduced to this special case. In order to achieve this reduction, we establish a set of invariants of the strongly connected components of the corresponding solution graph.

6.1 $\mathcal{L}$-Stabilizers

The following notation plays a key role in our construction.

▶ **Definition 19.** *Let S be a finite semigroup. We say that an element $u \in S^1$ is an* $\mathcal{L}$-stabilizer *of an element $x \in S$ provided that $u^\omega x = x$, and we denote the set of all such elements by*

$$\operatorname{stab}_{\mathcal{L}}(x) = \{u \in S^1 \mid u^\omega x = x\}. \qquad \diamond$$

This terminology stems from the observation that the orbit of an element x under the action of multiplication on the left by an $\mathcal{L}$-stabilizer u of x is completely contained in $\mathcal{L}(x)$. Indeed, if $u^\omega x = x$, then $x = u^\omega x \leq_{\mathcal{L}} u^n x \leq_{\mathcal{L}} x$ holds for all $n \in \mathbb{N}$.

Subsequently, we will be interested in the following class of finite semigroups for which this notion of $\mathcal{L}$-stabilizers is particularly well-behaved.

▶ **Definition 20.** *We say that a finite semigroup S belongs to the class* **DLG** *if the following conditions are satisfied for all $x \in S$ and $u, v \in S^1$.*

(1) *If $u^\omega x = x$ and $v^\omega x = x$, then $(uv)^\omega x = x$; that is, $\operatorname{stab}_{\mathcal{L}}(x)$ is a submonoid of S^1.*
(2) *If $x \sim_{\mathcal{J}} y$, then $u^\omega x = x$ if and only if $u^\omega y = y$; that is, $x \sim_{\mathcal{J}} y \Rightarrow \operatorname{stab}_{\mathcal{L}}(x) = \operatorname{stab}_{\mathcal{L}}(y)$.*
(3) *If $ux \sim_{\mathcal{L}} x$, then we have $u^\omega x = x$.* ◇

We will show in Section 7 that **DLG** is precisely the variety of finite semigroups whose regular $\mathcal{D}$-classes are right groups; see Lemma 33 and Lemma 34. We also show that various other well studied varieties of finite semigroups are subvarieties of **DLG**. We postpone this discussion, since none of these facts are used for the proof of our main result.

6.2 Nicely Balanced Equations

In Section 4 we defined when a finite semigroup is nice. The counterpart for equations is the following definition, which is not restricted to quadratic word equations.

▶ **Definition 21.** *Let $U = V$ be an equation with regular constraints μ and with a set of variables $\mathcal{X}$. It is called* nicely balanced *with regard to $X \in \mathcal{X}$ if $|\mu^{-1}(\mu(X)) \cap \Sigma^+| = \infty$ and if, up to symmetry, one of the following conditions holds.*

- *We have $|UV|_X = 0$. That is, X does not appear in $U = V$.*
- *We have $U = Xu$ and $V = vXv'$ with $u, v, v' \in (\Omega \setminus \{X\})^*$ and $\mu(v) \in \operatorname{stab}_{\mathcal{L}}(\mu(X))$.* ◇

▶ **Lemma 22.** *Let $(U = V, \mu)$ be a solvable word equation which is nicely balanced with regard to $X \in \mathcal{X}$ (whether or not it is quadratic). Then $\exp(U = V, \mu) = \infty$.*

Proof. Starting with some solution $\sigma \in \mathrm{Sol}(U = V, \mu)$, we will create new solutions with high exponents of periodicity by only changing the value of $\sigma(X)$. There are two cases to consider.

- Suppose that $|UV|_X = 0$. Then any word $w \in \mu^{-1}(\mu(X))$ induces a solution σ_w by setting $\sigma_w(X) = w$ and $\sigma_w(Y) = \sigma(Y)$ for all $Y \in \mathcal{X} \setminus \{X\}$. The claim now follows from the pumping lemma for regular languages, since $|\mu^{-1}(\mu(X)) \cap \Sigma^+| = \infty$ by hypothesis.
- Suppose that $U = Xu$ and $V = vXv'$ with $u, v, v' \in (\Omega \setminus \{X\})^*$ and $\mu(v) \in \mathrm{stab}_{\mathcal{L}}(\mu(X))$. If $v = 1$, then the equation is equivalent to the equation $u = v'$ with $|uv'|_X = 0$ and we are done by the first case. Otherwise, we fix a number $\omega \in \mathbb{N}_+$ such that $\mu(v)^\omega$ is idempotent. For any $m \in \mathbb{N}$, let σ_m be defined by $\sigma_m(X) = \sigma(v^{m\omega}X)$ and $\sigma_w(Y) = \sigma(Y)$ for all $Y \in \mathcal{X} \setminus \{X\}$. Indeed, this solves the equation $U = V$ and has high exponent of periodicity due to $\sigma(v) \neq 1$. The constraint is also satisfied for σ_m since we have $\mu(\sigma_m(X)) = \mu(\sigma(v))^{m\omega}\mu(\sigma(X)) = \mu(v)^\omega \mu(X) = \mu(X)$, where the second equality uses that σ is a solution of $(U = V, \mu)$ and the third uses $\mu(v) \in \mathrm{stab}_{\mathcal{L}}(\mu(X))$.

In either case, $(U = V, \mu)$ has solutions with arbitrarily high exponent of periodicity. ◀

6.3 Strongly Connected Components and Leaders

Let $\mathcal{A}(E_0)$ be the Matiyasevich Solution Graph for some quadratic equation E_0 with constraints in a finite semigroup S. By $\mathcal{C}$ we denote a strongly connected component of $\mathcal{A}(E_0)$ which contains at least one transition.[9] This implies that each state in $\mathcal{C}$ is solvable since $\mathcal{A}(E_0)$ is defined to be trim. In particular, each state in $\mathcal{C}$ has at least one outgoing transition. The next lemma characterizes those transitions.

▶ **Lemma 23.** *Let $E = (U = V, \mathcal{X}, \mu) \to E' = (U' = V', \mathcal{X}', \mu')$ be a transition within $\mathcal{C}$. Then:*

(1) *The transition is labeled $X \mapsto \alpha X$ for some $X \in \mathcal{X}$ and $\alpha \in \Omega$.*
(2) *We have $\mu(X) \leq_{\mathcal{R}} \mu(\alpha)$ and, in particular, $\mu(X) \leq_{\mathcal{J}} \mu(\alpha)$.*
(3) *We have $\mu(X) \sim_{\mathcal{L}} \mu'(X)$ and, in particular, $\mu(X) \sim_{\mathcal{J}} \mu'(X)$.*
(4) *We have $|L_{\mu(X)}| = \infty$ where $L_s = \mu^{-1}(s) \cap \Sigma^+$.*

Proof. Item (1) follows since all other transitions reduce the size of $\mathcal{X}$. Item (2) then follows from $\mu(X) = \mu(\alpha)\mu'(X)$. Since the transition is within $\mathcal{C}$, the latter also implies (3).

To see claim (4), observe that $L_{\mu(\alpha)} L_{\mu'(X)} \subseteq L_{\mu(X)}$ and that $L_{\mu(\alpha)} \neq \emptyset$ since the state E is solvable. If $L_{\mu'(X)}$ is finite, then we can choose $w \in L_{\mu'(X)}$ of maximal length. Choosing any $u \in L_{\mu(\alpha)}$ now yields $uw \in L_{\mu(X)}$. This is however impossible within $\mathcal{C}$, since $|uw| > |w|$. We conclude that $L_{\mu'(X)}$ and, thus, $L_{\mu(X)}$ are infinite. ◀

▶ **Lemma 24.** *Let $E = (\alpha U, \beta V, \mu)$ be any equation in $\mathcal{C}$. Then the minimum of $\mathcal{J}$-classes $J_E = \min\{\mathcal{J}(\mu(\alpha)), \mathcal{J}(\mu(\beta))\}$ is well-defined and independent of the chosen state E in $\mathcal{C}$. In particular, the* leading $\mathcal{J}$-class $J_{\mathcal{C}} = J_E$ *of $\mathcal{C}$ is well-defined.*

Proof. Since $\mathcal{C}$ is nontrivial, there is at least one outgoing transition, which is not an ε-transition. Therefore, using item (1) in Lemma 23 and by symmetry, we may assume that $\beta = X \in \mathcal{X}$ and that E has an outgoing transition to some state E' in $\mathcal{C}$ with label $\tau : X \mapsto \alpha X$. Then item (2) in Lemma 23 implies that $J_E = \mathcal{J}(\mu(X))$ is well-defined. For any such transition, we then have $E' = (\tau(U), X\tau(V), \mu')$ so that $J_{E'} \leq_{\mathcal{J}} \mathcal{J}(\mu'(X)) = \mathcal{J}(\mu(X)) = J_E$ by item (3) in Lemma 23. Since we remain within $\mathcal{C}$, this implies $J_E = J_{E'}$. ◀

[9] Such a component $\mathcal{C}$ exists as soon as we have $n \geq 4$. To see this, let $a \in \Sigma$ and $X \in \mathcal{X}$ with $\mu(X) = \mu(a)^\omega$. Then the NFA $\mathcal{A}_n$ contains the state $(Xa = aX, \mu)$ which is part of a cycle of transitions labeled $X \mapsto aX$.

6.4 Playgrounds and Players

This section makes heavy use of the properties from Definition 20. From now on, we thus assume that all constraints $\mu : \Omega^+ \to S$ are with respect to semigroups $S \in \mathbf{DLG}$.

Combining property (2) in Definition 20 with Lemma 24 allows us to define $S_{\mathcal{C}} = \operatorname{stab}_{\mathcal{L}}(J_{\mathcal{C}})$, which we call the *leading $\mathcal{L}$-stabilizer* of the strongly connected component $\mathcal{C}$. Given an equation $E = (U = V, \mu) \in \mathcal{C}$ we can now determine maximal prefixes $\widetilde{U}$ and $\widetilde{V}$ of U and V, respectively, such that every proper prefix consists only of symbols $\alpha \in \Omega$ with $\mu(\alpha) \in S_{\mathcal{C}}$. We call the resulting pair $(\widetilde{U}, \widetilde{V})$ the *playground* of E. The *size* of this playground is $\widetilde{n}_E = |\widetilde{U}\widetilde{V}|$, and the *players* of E are the variables belonging to the set

$$\widetilde{\mathcal{X}}_E = \{X \in \mathcal{X} \mid \mathcal{J}(\mu(X)) = J_{\mathcal{C}},\ |L_{\mu(X)}| = \infty, \text{ and } X \text{ occurs twice in } \widetilde{U}\widetilde{V}\},$$

where $L_s = \mu^{-1}(s) \cap \Sigma^+$. The following establishes that both $\widetilde{n}_E$ and $\widetilde{\mathcal{X}}_E$ are invariants of $\mathcal{C}$.

▶ **Lemma 25.** *Let $E = (U = V, \mathcal{X}, \mu) \to E' = (U' = V', \mathcal{X}', \mu')$ be a transition in $\mathcal{C}$ which is labeled by the substitution defined by $X \mapsto \alpha X$. We have $\widetilde{\mathcal{X}}_E = \widetilde{\mathcal{X}}_{E'}$ and $\widetilde{n}_E = \widetilde{n}_{E'}$. Moreover, if $|UV|_X \geq 1$, then $X \in \widetilde{\mathcal{X}}_E$ is a player of E and $\mu(\alpha) \in S_{\mathcal{C}}$.*

Proof. The claim is trivial in case $|UV|_X = 0$. Hence, we may assume that $U = Xu = U'$ and $V = \alpha V'$. Item (3) in Lemma 23 yields $\mu'(X) \sim_{\mathcal{L}} \mu(\alpha)\mu'(X)$ and, hence, $\mu(\alpha) \in \operatorname{stab}_{\mathcal{L}}(\mu'(X))$ by property (3) in Definition 20. Thus $\mu(\alpha) \in S_{\mathcal{C}} = \operatorname{stab}_{\mathcal{L}}(J_{\mathcal{C}})$ as $\mathcal{J}(\mu'(X)) = \mathcal{J}(\mu(X)) = J_{\mathcal{C}}$.

In particular, the symbol α is part of the playground of E, and therefore contributes to the size $\widetilde{n}_E$ and possibly to the players $\widetilde{\mathcal{X}}_E$. Considering all possible ways in which the variable X might occur in E leads us to the following case distinction.

- Suppose that X occurs only once in E. Then we can deduce that $|E'| < |E'|$.
- Suppose that X occurs again, but outside of the playground of E. Then the transition removes α from the playground of E' and thus $\widetilde{n}_{E'} = \widetilde{n}_E - 1$ and $\widetilde{\mathcal{X}}_{E'} = \widetilde{\mathcal{X}}_E \setminus \{\alpha\}$.
- Suppose that X occurs again inside the playground of E. Then α remains part of the playground of E' after the substitution, and we have $\widetilde{n}_{E'} = \widetilde{n}_E$ and $\widetilde{\mathcal{X}}_{E'} = \widetilde{\mathcal{X}}_E$. Moreover, $X \in \widetilde{\mathcal{X}}_E$ is a player since $\mathcal{J}(\mu(X)) = J_{\mathcal{C}}$ and since $|L_{\mu(X)}| = \infty$ by item (4) in Lemma 23.

Finally, since the above case distinction is exhaustive, and since the transition takes place within the strongly connected component $\mathcal{C}$, we can rule out the first two cases. ◀

By Lemma 25 we can now talk about the players of the strongly connected component $\mathcal{C}$, the set of which we denote by $\widetilde{\mathcal{X}}_{\mathcal{C}}$. The lemma also informally states that, within $\mathcal{C}$, we may restrict our analysis to the respective playgrounds and treat all non-players as if they were constants. Be aware however that, even though the set of players $\widetilde{\mathcal{X}}_{\mathcal{C}}$ depends only on $\mathcal{C}$, the playground itself and the arrangement of players within it may vary between states.

For a given quadratic equation $E = (U = V, \mu) \in \mathcal{C}$ with playground $(\widetilde{U}, \widetilde{V})$, we now split the set of players $\widetilde{\mathcal{X}}_E = \widetilde{\mathcal{X}}_E^b \cup \widetilde{\mathcal{X}}_E^u$ into a disjoint union where

$$\widetilde{\mathcal{X}}_E^b = \{X \in \widetilde{\mathcal{X}}_{\mathcal{C}} \mid |\widetilde{U}|_X = 1 \text{ and } |\widetilde{V}|_X = 1\} \quad \text{and} \quad \widetilde{\mathcal{X}}_E^u = \{X \in \widetilde{\mathcal{X}}_{\mathcal{C}} \mid |\widetilde{U}|_X = 2 \text{ or } |\widetilde{V}|_X = 2\}.$$

Accordingly, we call the players $X \in \widetilde{\mathcal{X}}_E^b$ *balanced* and the players $X \in \widetilde{\mathcal{X}}_E^u$ *unbalanced*.

The following observation is now crucial. If any side of the equation E begins with a balanced player X, then E is nicely balanced with regard to X in the sense of Definition 21: We have $J_{\mathcal{C}} = \mathcal{J}(\mu(X))$ and $S_{\mathcal{C}} = \operatorname{stab}_{\mathcal{L}}(\mu(x))$ since X is a player. Writing $U = Xu$ and $V = vXv'$, we note that vX is a prefix of $\widetilde{V}$ and, hence, every symbol $\alpha \in \Omega$ occurring in v satisfies $\mu(\alpha) \in S_{\mathcal{C}}$. By property (1) in Definition 20, this implies $\mu(v) \in S_{\mathcal{C}} = \operatorname{stab}_{\mathcal{L}}(\mu(x))$. Finally, $|\mu^{-1}(\mu(X)) \cap \Sigma^+| = \infty$ holds by definition, since X is a player.

▶ **Theorem 26.** *Every cycle in $\mathcal{C}$ contains an equation, which is nicely balanced with regard to some variable.*

Proof. Take some equation E on the cycle. The claim is trivial, if there exists a variable $X \in \mathcal{X}$ with $|L_{\mu(X)}| = \infty$ such that X does not appear on either side of E. By claim (4) in Lemma 23 we may thus assume that X appears in E for every transition $X \to \alpha X$ within $\mathcal{C}$. We can then write $E = (XU, \alpha V, \mu)$. If E is not nice, then neither side begins with a balanced player by the above discussion. In particular, $X \in \widetilde{\mathcal{X}}_E^u$ is unbalanced. If $\alpha \notin \widetilde{\mathcal{X}}_{\mathcal{C}}$ is not a player, then the length of the right side of the equation diminishes after the substitution. Hence after following at most $|V|$ transitions of this type, we may assume that $\alpha \in \widetilde{\mathcal{X}}_{\mathcal{C}}$ is a player and, moreover, unbalanced. Following another substitution the number $|\widetilde{\mathcal{X}}_E^u|$ is reduced since α is now balanced. The only way this amount can increase again is if one of the sides begins with a balanced player; but then E is nicely balanced. ◀

Letting now E be any quadratic equation with infinitely many solutions, Proposition 16 tells us that $\mathcal{A}(E)$ contains some nontrivial connected component. Using this and Theorem 26, we can reduce E to a solvable nicely balanced equation. By combining this with Lemma 22, we deduce our main result, which is stated as Theorem 1 in the introduction.

▶ **Corollary 27.** *Let $(U = V, \mu)$ be a quadratic equation with regular constraints in a semigroup from $\mathbf{DLG}$. Then $(U = V, \mu)$ has infinitely many solutions if and only if $\exp(U = V, \mu) = \infty$.*

6.5 The Brandt Semigroup B_2

The *(combinatorial) Brandt semigroup B_2* is the semigroup $\{a, b, ab, ba, 0\}$ with multiplication given by the equations $\{aba = a, bab = b, a^2 = b^2 = 0\}$. It does not belong to the class $\mathbf{DLG}$ because we have $ba \sim_{\mathcal{L}} a$ but $b^\omega a = 0 \neq a$, so that property (3) in Definition 20 is violated.[10] However, we can still show the following.

▶ **Proposition 28.** *The Brandt semigroup B_2 is nice for the family of quadratic word equations with at most two constants.*

Our proof of Proposition 28 relies on the following lemma, which itself can be easily verified by routine calculations. We omit the details of these calculations.

▶ **Lemma 29.** *For every proper subsemigroup S of B_2, it holds that $S \in \mathbf{DLG}$.*

Proof of Proposition 28. Let $E = (U = V, \mu)$ be a quadratic equation with $\mu : \Omega^+ \to B_2$ over a set of constants Σ with $|\Sigma| \leq 2$. If $\mu(\Sigma)$ generates a proper subsemigroup, then the result follows from Lemma 29 together with Corollary 27. Hence, we may assume that $\mu(\Sigma)$ generates B_2. Up to symmetry this leaves only one possibility, namely that $\Sigma = \{a, b\}$ gets mapped to B_2 as $\mu(a) = a$ and $\mu(b) = b$. Accordingly, we have

$$\mu^{-1}(a) = a(ba)^*, \quad \mu^{-1}(b) = b(ab)^*, \quad \mu^{-1}(ab) = (ab)^*, \quad \mu^{-1}(ba) = (ba)^*.$$

Since arbitrarily long words in these languages clearly have arbitrarily high exponent of periodicity, we may assume that $\mu(X) \notin \{a, b, ab, ba\}$ and, hence, $\mu(X) = 0$ for all $X \in \mathcal{X}$.

---

[10] Note that $ba \sim_{\mathcal{L}} a$ implies that the $\mathcal{J}$-class of a contains the idempotent ba. Hence, $\mathcal{J}$-class of a is regular, but not a subsemigroup as $a^2 = 0 \not\sim_{\mathcal{J}} a$. This means that B_2 is not contained in the variety $\mathbf{DS}$ consisting of all finite semigroups where the regular $\mathcal{D}$-classes are semigroups. In fact, it holds that a finite semigroup S belongs to $\mathbf{DS}$ if and only if B_2 does not divide $S \times S$ [1, Exercise 8.1.6].

On the other hand, we have $\mu^{-1}(0) = \Sigma^*\{a^2, b^2\}\Sigma^*$. Phrased differently, if $\sigma \in \mathrm{Sol}(E)$, then $\sigma(X)$ contains a factor c_X^2 with $c_X \in \Sigma$ for every variable $X \in \mathcal{X}$. We can guess these constants c_X for all variables $X \in \mathcal{X}$ simultaneously in the form of a mapping $c : \mathcal{X} \to \Sigma$. Applying the substitution $X \mapsto \tau_c(X) = X_1 c_X^2 X_2$ with fresh variables X_1, X_2 for all $X \in \mathcal{X}$, we arrive at a quadratic equation $E_c = (\tau_c(U) = \tau_c(V))$ *without* constraints.

Allowing for singular solutions at the E_c, this nondeterministic procedure is both sound and complete: every solution σ' at any E_c gives rise to a solution $\sigma = \sigma'\tau_c$ at E, and every solution σ at E factorizes as $\sigma = \sigma'\tau_c$ for some solution σ' at E_c for some suitable $c : \mathcal{X} \to \Sigma$.

By the pigeonhole principle, if E has infinitely many solutions, then so does some E_c. Hence, $\exp(E) \geq \exp(E_c) = \infty$ holds by Proposition 13. ◀

For $\Sigma = \{a, b, c\}$ the prove fails to handle mappings like $\mu(a) = a, \mu(b) = b$, and $\mu(c) = ab$. These images generate the entire semigroup B_2, so we are not able to reduce to a proper subsemigroup. At the same time, $\mu^{-1}(ab) = \{ab, c\}^*$ contains arbitrarily long words with bounded exponent of periodicity, so we cannot hope to eliminate those variables $X \in \mathcal{X}$ with $\mu(X) = ab$ as we did in the above proof.

7 The Variety DLG

Our main result, Corollary 27, concerns semigroups from a certain subvariety of **DS**, which itself consists of all finite semigroups whose regular $\mathcal{D}$-classes are semigroups, and which was first considered in Schützenberger's article '*Sur le produit de concaténation ambigu*' [35]. The subvariety in question, which is also the subject of this section, is the variety **DLG** comprising all finite semigroups whose regular $\mathcal{D}$-classes are right groups. Throughout this section, we adopt this as a definition of **DLG**, and show in Section 7.3 this definition agrees with the earlier Definition 20 – that is, both define the same class of finite semigroups.

Before we examine this variety in more detail in this section, let us remark that **DLG** (and indeed any variety of the form **DV**) has certain useful closure properties: It is closed under adjunction of a neutral element (hence, the distinction between treating it as a variety of finite semigroups or as one of monoids is marginal) and, adjunction of a zero element. Due to the latter, we can reduce any system of word equations with constraints in **DLG** to a single such equation as described in Section 3.1. The variety is also closed under forming nilpotent extensions, so that nothing is to be gained from combining Theorem 11 and Corollary 27.

7.1 Right Groups

A finite semigroup S is called a *right group* if it satisfies the identity $x = y^\omega x$. If a monoid N is right group, then N is a group. However, there are infinitely many finite semigroups satisfying $x = y^\omega x$ without being groups. To see this, let G be a finite group and I be a nonempty finite set. Then the set $S = G \times I$ becomes a right group when endowed with the multiplication defined by $(g, i) \cdot (h, j) = (gh, j)$, but S is a group only if I is a singleton.

It is a well-known fact, which however will not be used explicitly here, that every right group is of the above form [10, Thm. 1.27]. Instead, we content ourselves with the following.

▶ **Lemma 30.** *A finite semigroup S is a right group if and only if it is right simple, meaning that S has no proper right ideals or, equivalently, that all elements of S are $\mathcal{R}$-equivalent.*[11]

[11] The empty set has no proper subsets, so the empty set is indeed a right group.

Proof. If S is right group, then $y \geq_{\mathcal{R}} y^\omega x = x$ for all $x, y \in S$ and, hence, $x \geq_{\mathcal{R}} y \geq_{\mathcal{R}} x$.

To see the converse, consider $x, y \in S$ and note that $x \sim_{\mathcal{R}} y^\omega$ by assumption. We may thus write $x = y^\omega z$ for some $z \in S^1$ and, therefore, $x = y^\omega z = y^\omega y^\omega z = y^\omega x$. ◀

Right groups are regular semigroups, meaning that every element of a right group is regular (since it is contained in a subgroup). Moreover, the idempotent elements of a right group always form a subsemigroup and, in fact, a right zero semigroup (as witnessed by the identity $y^\omega x^\omega = x^\omega$). Therefore, right groups are examples of *orthodox* semigroups, which are precisely those regular semigroup in which the idempotents form a subsemigroup.

7.2 Other Characterizations

We now turn to the variety **DLG**. As mentioned in the introduction to this section, it comprises all finite semigroups whose regular $\mathcal{D}$-classes are right groups. In particular, **DLG** is contained in the variety **DO** consisting of all finite semigroups whose regular $\mathcal{D}$-classes are orthodox semigroups. The latter is, in turn, contained in the variety **DS** consisting of all finite semigroups whose regular $\mathcal{D}$-classes are semigroups. Thus, $\mathbf{DLG} \subseteq \mathbf{DO} \subseteq \mathbf{DS}$.

Even though the variety of interest is defined in terms of right groups, its name **DLG** is more aptly explained by the following characterization.[12]

▶ **Lemma 31.** *The following two assertions are equivalent for a finite semigroup S.*
(1) *Every regular $\mathcal{D}$-class of S is a right group.*
(2) *Every regular $\mathcal{L}$-class of S is a group.*

Proof. *Ad* (1) ⇒ (2). Let $\mathcal{L}(e)$ be a regular $\mathcal{L}$-class. Then $\mathcal{D}(e)$ is regular and thus a right group by assumption. As in the proof of Lemma 30, this implies that all elements of $\mathcal{D}(e)$ are $\mathcal{R}$-equivalent. But then $\mathcal{L}(e)$ is a regular $\mathcal{H}$-class and, in particular, a group.

Ad (2) ⇒ (1). Let $\mathcal{D}(e)$ be a regular $\mathcal{D}$-class. Then $\mathcal{L}(x)$ is regular for every $x \in \mathcal{D}(e)$. By assumption, each $\mathcal{L}(x)$ is now a group and thus a $\mathcal{H}$-class, implying $\mathcal{D}(e) = \mathcal{R}(e)$. To see that $\mathcal{D}(e)$ is closed under multiplication, note that $xy \sim_{\mathcal{R}} x^2 \sim_{\mathcal{L}} x$ for $x, y \in \mathcal{D}(e)$. The first equivalence follows since $x \sim_{\mathcal{R}} y$ and the second, since $\mathcal{L}(x)$ is a group. In particular, we have $x \sim_{\mathcal{R}} y^\omega$ and we may thus conclude $x = y^\omega x$ as in the proof of Lemma 30. ◀

The variety **DLG**, along with its left-right dual **DRG**, has been studied extensively by Almeida and Weil [3]. Various other results on **DLG** and its generalizations can be found in Célia Borlido's thesis [7] and the articles [2, 5, 6]. The following characterization of **DLG** and its proof were kindly provided by Jorge Almeida (personal communication).

▶ **Lemma 32.** *For every finite semigroup S, the following assertions are equivalent.*
(1) *The semigroup S satisfies the identity $(xy)^\omega = y^\omega (xy)^\omega$.*
(2) *The semigroup S satisfies the identity $(xyz)^\omega = y^\omega (xyz)^\omega$.*
(3) *The semigroup S satisfies the identity $(xy)^\omega = (yx)^\omega (xy)^\omega$.*
(4) *Every regular $\mathcal{D}$-class of S is a right group.*

Proof. To see the implication (1) ⇒ (2), let us first note that (1) implies the identity

$$(xy)^\omega = x(yx)^\omega y(xy)^{\omega-1} = xx^\omega (yx)^\omega y(xy)^{\omega-1} = x^\omega x(yx)^\omega y(xy)^{\omega-1} = x^\omega (xy)^\omega.$$

Using this, we obtain $(xyz)^\omega = (yz)^\omega (xyz)^\omega = y^\omega (yz)^\omega (xyz)^\omega = y^\omega (xyz)^\omega$.

[12] Proposition 2 permits reading the second statement as '*in each regular $\mathcal{D}$-class, each $\mathcal{L}$-class is a group*', thereby giving rise to a memo for the notation: **DLG** = **D**-class **L**-class **G**roup.

Assuming (2), the semigroup S satisfies $(xy)^\omega = (xyxy)^\omega = (yx)^\omega(xyxy)^\omega = (yx)^\omega(xy)^\omega$, which shows that (2) $\Rightarrow$ (3). To see (3) $\Rightarrow$ (1), first note that $(xy)^\omega = (yx)^\omega(xy)^\omega$ implies the defining identity $(xy)^\omega = ((xy)^\omega(yx)^\omega(xy)^\omega)^\omega$ for the variety **DS** [1, Sec. 8.1]. Now, for all $x, y \in S$, the idempotents $e = (y^\omega(xy)^\omega)^\omega$ and $f = (xy)^\omega$ satisfy $e \sim_{\mathcal{J}} f$ by [1, Lemma 8.1.4]. Indeed, we have $f \leq_{\mathcal{J}} y^n$ for all $n \geq 1$ since this implies $f \sim_{\mathcal{J}} fy^n \leq_{\mathcal{J}} y^{n+1}$ by the cited lemma and the assertion clearly holds for $n = 1$. So $f \leq_{\mathcal{J}} y^\omega$ and, therefore, $f \sim_{\mathcal{J}} y^\omega f \sim_{\mathcal{J}} (y^\omega f)^\omega = e$.

Moreover, since $e \leq_{\mathcal{L}} f$, this shows $e \sim_{\mathcal{L}} f$ so that $e = ef$ and $f = fe$. We then deduce that $f = fe = (ef)(fe) = ef = e$ using $(xy)^\omega = (yx)^\omega(xy)^\omega$ for the second equality. Hence, it follows that $(xy)^\omega = (y^\omega(xy)^\omega)^\omega = y^\omega(y^\omega(xy)^\omega)^\omega = y^\omega(xy)^\omega$ as required.

Still assuming (3), we know from the above that $S \in$ **DS**. Hence, if $x, y \in S$ are elements of the same regular $\mathcal{D}$-class, then $xy \sim_{\mathcal{J}} yx \sim_{\mathcal{J}} x$. Thus, $x = uyxv = (uy)^\omega xv^\omega$ for some $u, v \in S^1$. Applying the identity from (1) yields $x = (uy)^\omega xv^\omega = y^\omega(uy)^\omega xv^\omega = y^\omega x$, which is a defining identity for right groups; hence, (3) $\Rightarrow$ (4). Finally, to see that (4) $\Rightarrow$ (3), note that $(xy)^\omega$ and $(yx)^\omega$ are always contained in the same (necessarily regular) $\mathcal{D}$-class. Due to the identity $x = y^\omega x$ for right groups, this implies the equation $(xy)^\omega = (yx)^\omega(xy)^\omega$. ◀

7.3 Properties of $\mathcal{L}$-Stabilizers

We now turn to the $\mathcal{L}$-stabilizers discussed in Section 6.1, which are particularly well-behaved for finite semigroups belonging to **DLG**. Indeed, the following results show that the definition of the class **DLG** used throughout this section is consistent with Definition 20.

▶ **Lemma 33.** *Let $S \in$ **DLG**. Then the following hold for all $x, y \in S$ and $u, v \in S^1$.*
(1) *If $u^\omega x = x$ and $v^\omega x = x$, then $(uv)^\omega x = x$; that is, $\mathrm{stab}_{\mathcal{L}}(x)$ is a submonoid of S^1.*
(2) *If $x \sim_{\mathcal{J}} y$, then $u^\omega x = x$ if and only if $u^\omega y = y$; that is, $x \sim_{\mathcal{J}} y \Rightarrow \mathrm{stab}_{\mathcal{L}}(x) = \mathrm{stab}_{\mathcal{L}}(y)$.*
*Conversely, if the second item is satisfied for some finite semigroup S, then $S \in$ **DLG**.*

Proof. *Ad* (1). Suppose that $u^\omega x = x$ and $v^\omega x = x$. Then $x = u^\omega x = u^\omega v^\omega x = (u^\omega v^\omega)^\omega x$ and, since $uv \geq_{\mathcal{J}} u^\omega v^\omega$, item (2) of Lemma 32 yields $(u^\omega v^\omega)^\omega x = (uv)^\omega(u^\omega v^\omega)^\omega x = (uv)^\omega x$.

Ad (2). Suppose that $x \sim_{\mathcal{J}} y$, that is, $y = rxs$ and $x = r'ys'$ for some $r, r', s, s' \in S^1$. If, moreover, $u^\omega x = x$, then $y = ru^\omega r'ys's = (ru^\omega r')^\omega y(s's)^\omega = u^\omega(ru^\omega r')^\omega y(s's)^\omega = u^\omega y$ where the second to last equality is due to item (2) of Lemma 32. Conversely, if S satisfies (2), then we have $yx \in \mathrm{stab}_{\mathcal{L}}((yx)^\omega) = \mathrm{stab}_{\mathcal{L}}((xy)^\omega)$ for all $x, y \in S$ since $(yx)^\omega \sim_{\mathcal{J}} (xy)^\omega$. Hence, the semigroup S satisfies the identity $(xy)^\omega = (yx)^\omega(xy)^\omega$ so that $S \in$ **DLG** by Lemma 32. ◀

▶ **Lemma 34.** *Let $S \in$ **DLG**. Then $u, x \in S$ satisfy $u^\omega x = x$ if and only if $ux \sim_{\mathcal{L}} x$. Conversely, if this equivalence holds for all elements u, x of a finite semigroup S, then $S \in$ **DLG**.*

Proof. If $u^\omega x = x$, then $x = u^\omega x \leq_{\mathcal{L}} ux \leq_{\mathcal{L}} x$ so that $ux \sim_{\mathcal{L}} x$. Moreover, $ux \sim_{\mathcal{L}} x$ holds if and only if $x = vux$ and, hence, $x = (vu)^\omega x$ for some $v \in S^1$. Hence, we also have $(vu)^\omega = u^\omega(vu)^\omega$ by the first identity from Lemma 32 and, therefore, $x = (vu)^\omega x = u^\omega(vu)^\omega x = u^\omega x$.

Now suppose that $ux \sim_{\mathcal{L}} x \Rightarrow u^\omega x = x$ holds in a finite semigroup S. Then, for all $x, y \in S$, we have $y^\omega(xy)^\omega = (xy)^\omega$ since $y(xy)^\omega \sim_{\mathcal{L}} (xy)^\omega$. Hence, $S \in$ **DLG** by Lemma 32. ◀

7.4 Examples

The following non-exhaustive list shows that the variety **DLG** is rather rich in the sense that it contains various interesting families of semigroups and can recognize interesting languages.

- For aperiodic semigroups, the variety **DLG** contains the variety **Sl** of finite semilattice, which allows for the implementation of *alphabetical constraints*, as well as the larger variety **J** of finite $\mathcal{J}$-trivial semigroups. The latter allows for *piecewise testable constraints* in the sense of Imre Simon [37] (see also [30, Thm. 4] or [15, Thm. 4]).

- More generally, **DLG** contains the variety **RRB** of finite right regular bands, as well as the larger variety **L** of finite $\mathcal{L}$-trivial semigroups, which coincides with the class of all finite semigroups whose regular $\mathcal{D}$-classes are right zero semigroups. On the other hand, the (nontrivial) left zero semigroups are *not* contained in **DLG**.
- As a subvariety of **J**, the variety **N** of all finite nilpotent semigroups is contained in **DLG**, and so is the variety $\mathbf{D} \subseteq \mathbf{L}$ consisting of all finite semigroups in which all idempotents are right zeros. Using constraints in **N** it is possible to, for example, enforce (or prohibit) the length of a solution to exceed a specified threshold, or to require that the solution is a specific word; using constraints in **D** on can require a specific word to appear as a suffix.
- The variety **G** of finite groups is also clearly contained in **DLG**, which allows for *group constraints*, such as modular counting. Note also that this type of constraints may be quite liberally combined with those discussed in the previous items; see [3, Prop. 2.2.4].

Another interesting family of finite semigroups contained in **DLG** is the family of *duo* semigroups comprising semigroups S in which $xS^1 = S^1x$ holds for all $x \in S$. This class of semigroups, which is not a variety but contains all groups and commutative semigroups, was defined by Pondělíček [33] as a semigroup-theoretic generalization of Feller's notion of duo rings [20]. Note that in a duo semigroup S every product of the form $(x_1y)(x_2y)\cdots(x_ny)$ with $x_1, \ldots, x_n, y \in S$ can be written as $y^n z$ for some $z \in S^1$. Hence, if S is finite, then we obtain $(xy)^\omega = y^\omega z = y^\omega y^\omega z = y^\omega (xy)^\omega$ for some $z \in S^1$. Thus, $S \in \mathbf{DLG}$ by Lemma 32.

Apart from the nontrivial left zero semigroups mentioned above, we have seen in Section 6.5 that the Brandt semigroup B_2 does *not* belong to **DLG** and, in fact, not even to **DS**.

8 Conclusion and Open Problems

The guiding question of this work is whether a word equation with infinitely many solutions must have an infinite exponent of periodicity. We do not yet know the answer. Nevertheless, regardless of the ultimate resolution of this question, the results obtained here – focused on equations with regular constraints – are of independent interest.

Our approach has two complementary aspects, like Yin and Yang or two sides of the same medal: on one side, a class $\mathcal{E}$ of word equations; on the other, a class **V** of semigroups. When **V** is the variety generated by the trivial semigroup (that is, for equations without constraints), it was already known that a word equation with infinitely many solutions has an infinite exponent of periodicity in two cases: two-variable and quadratic equations.

The quadratic case extends naturally to constraints in finite groups. More generally, if $U = V$ is quadratic and the constraint μ maps variables to units of a finite monoid M, the machinery from Section 6.2 applies because finite monoids are Dedekind finite, meaning that their non-units form an ideal. In this setting, the playground of a word $W \in \Omega^+$ is the maximal prefix $\widetilde{W}$ that is mapped to a unit $\mu(\widetilde{W})$. Similar reasoning works for alphabetic constraints, where playgrounds are defined in terms of sets of constants. This approach extends to duo semigroups and to the variety **DLG** (and its left-right dual **DRG**), but seems to reach its limit there. Indeed, **DLG** is up to duality the largest variety of finite semigroups known (to us) whose members are nice for all quadratic word equations. Whether an even larger variety shares this property remains open. Beyond this, a few special cases are known: for example, we have seen that the Brandt semigroup B_2 is nice for quadratic equations with at most two constants.

Still, even without constraints, the central question remains far from settled, and there is ample room for partial progress. Below we outline several open problems. In what follows, it is natural to first restrict attention to $\mathcal{E}$ consisting of two-variable or quadratic equations.

▷ Problem 1. Does the family of nice semigroups for $\mathcal{E}$ form a variety of finite semigroups?

▷ Problem 2. Given a finite semigroup S, is it decidable whether S is nice for $\mathcal{E}$?

▷ Problem 3. Is every finite semigroup in **DO** or **DS** nice for $\mathcal{E}$?

The equation $XabY = YbaX$ admits solutions where $\sigma(X)$ and $\sigma(Y)$ are arbitrarily long factors of Fibonacci words [38, Sec. 6], hence with bounded exponent of periodicity, making it a candidate for a potential counterexample. Regular constraints alone might not suffice, so one could also consider length constraints – for instance, requiring $p \cdot |\sigma(Y)| \le |\sigma(X)| \le q \cdot |\sigma(Y)|$ for rationals $p < q$ close to the golden ratio $\varphi = 1.618\ldots$. While the decidability of word equations with length constraints is a long-standing open problem, Day and Konefa [12] recently proved it decidable for two-variable quadratic equations, including $XabY = YbaX$.

▷ Problem 4. Consider an equation $XAY = YBX$ over $\Omega = \{a,b\} \cup \{X,Y\}$ with $A, B \in \Omega^+$, for example $XabY = YbaX$. Does there exist a constant κ, and combination of regular constraints and length constraints such that, under these constraints, the equation has infinitely many solutions, and every solution has exponent of periodicity less than κ?

The problems above about word equation in free semigroups also arise naturally in the setting of free groups. Here, recognizable constraints correspond to homomorphisms to finite groups. A much larger and much more interesting class of constraints for a free group F is the family of rational subsets $\mathrm{Rat}(F)$. It is known that $\mathrm{Rat}(F)$ forms an effective Boolean algebra [4], and that the full solution set of an equation with a rational constraint can be effectively described by an EDT0L language, using results from [16], [17], and [8]. However, even for quadratic equations, the following question remains open.

▷ Problem 5. Are all rational constraints nice for quadratic word equations in free groups?

References

1 J. Almeida. *Finite Semigroups and Universal Algebra*. World Scientific, Singapore, 1994.

2 J. Almeida and C. Borlido. Complete κ-reducibility of pseudovarieties of the form DRH. *Internat. J. Algebra Comput.*, 27:189–235, 2017.

3 J. Almeida and P. Weil. Free profinite $\mathcal{R}$-trivial monoids. *Internat. J. Algebra Comput.*, 7:625–671, 1997.

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

5 C. Borlido. The κ-word problem over DRH. *Theoretical Computer Science*, 702:1–22, 2017.

6 C. Borlido. Some reducibility properties for pseudovarieties of the form DRH. *Comm. Algebra*, 46:1813–1830, 2018.

7 C. M. R. Borlido. *The word problem and some reducibility properties for pseudovarieties of the form* DRH. ProQuest LLC, Ann Arbor, MI, 2016. Thesis (Ph.D.)–Universidade do Porto (Portugal).

8 L. Ciobanu, V. Diekert, and M. Elder. Solution sets for equations over free groups are EDT0L languages. *Int. J. Algebra Comp.*, 26:843–886, 2016. Conference version in Proc. ICALP 2015, LNCS 9135.

9 L. Ciobanu and M. Elder. The complexity of solution sets to equations in hyperbolic groups. *Israel J. Math.*, 245(2):869–920, 2021.

10 A. H. Clifford and G. B. Preston. *The Algebraic Theory of Semigroups, Volume I.* Number 7 in Mathematical Surveys. American Mathematical Society, 1961.

11 R. Dabrowski and W. Plandowski. Solving two-variable word equations. In J. Díaz, J. Karhumäki, A. Lepistö, and D. Sannella, editors, *31st International Colloquium ICALP 2004 Proceedings*, volume 3142 of *Lecture Notes in Comput. Sci.*, pages 408–419, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

12 J. D. Day and M. Konefal. Word equations with length constraints via weak arithmetics and matrix reachability problems. In P. Ganty and A. Mansutti, editors, *Reachability Problems*, pages 51–67, Cham, 2026. Springer Nature Switzerland.

13 V. Diekert. Makanin's Algorithm. In M. Lothaire, editor, *Algebraic Combinatorics on Words*, volume 90 of *Encyclopedia of Mathematics and Its Applications*, chapter 12, pages 387–442. Cambridge University Press, 2002.

14 V. Diekert and M. Elder. Solutions to twisted word equations and equations in virtually free groups. *Internat. J. Algebra Comput.*, 30:731–819, 2020. Based on the conference in LIPIcs.ICALP.2017.96:1–96:14.

15 V. Diekert, P. Gastin, and M. Kufleitner. A survey on small fragments of first-order logic over finite words. *Int. J. Found. Comput. Sci.*, 19:513–548, 2008. Special issue DLT 2007.

16 V. Diekert, C. Gutiérrez, and Christian Hagenah. The existential theory of equations with rational constraints in free groups is PSPACE-complete. *Information and Computation*, 202:105–140, 2005. Conference version in STACS 2001, LNCS 2010, 170–182, 2001.

17 V. Diekert, A. Jeż, and W. Plandowski. Finding all solutions of equations in free groups and monoids with involution. *Inform. and Comput.*, 251:263–286, 2016. Conference version in Proc. CSR 2014, LNCS 8476 (2014).

18 V. Diekert, M. Kufleitner, G. Rosenberger, and U. Hertrampf. *Discrete Algebraic Methods. Arithmetic, Cryptography, Automata and Groups.* De Gruyter, 2016.

19 S. Eilenberg. *Automata, Languages, and Machines*, volume A. Academic Press, New York and London, 1974.

20 E. H. Feller. Properties of primary noncommutative rings. *Trans. Amer. Math. Soc.*, 89:79–91, 1958.

21 Ju. I. Hmelevskiĭ. *Equations in Free Semigroups.* Number 107 in Proc. Steklov Institute of Mathematics. American Mathematical Society, 1976. Translated from the Russian original: Trudy Mat. Inst. Steklov. 107, 1971.

22 L. Ilie and W. Plandowski. Two-variable word equations. *RAIRO Theor. Informatics Appl.*, 34:467–501, 2000.

23 A. Jeż. Recompression: a simple and powerful technique for word equations. *Journal of the ACM*, 63(1):4:1–4:51, 2016. Conference version in Proc. STACS 2013.

24 O. Kharlampovich and A. Myasnikov. Elementary theory of free non-abelian groups. *J. of Algebra*, 302:451–552, 2006.

25 S. C. Kleene. Representation of events in nerve nets and finite automata. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, number 34 in Annals of Mathematics Studies, pages 3–40. Princeton University Press, 1956.

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

27 J. V. Matijasevič. A connection between systems of word and length equations and Hilbert's tenth problem. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 8:132–144, 1968.

28 Yu. V. Matiyasevich. *Hilbert's Tenth Problem.* MIT Press, Cambridge, Massachusetts, 1993.

29 S. Natterer. Quadratic equations in free groups and free monoids. Bachelor thesis, University of Stuttgart, 2024.

30 J.-É. Pin. *Varieties of Formal Languages.* North Oxford Academic, London, 1986.

31 W. Plandowski and W. Rytter. Application of Lempel-Ziv encodings to the solution of word equations. In K. G. Larsen et al., editors, *Proc. 25th International Colloquium Automata,*

Languages and Programming (ICALP'98), Aalborg (Denmark), 1998, volume 1443 of *Lecture Notes in Comput. Sci.*, pages 731–742, Heidelberg, 1998. Springer-Verlag.

32 W. Plandowski and A. Schubert. On the complexity of computation maximal exponent of periodicity of word equations and expressible relations (note). *Theor. Comput. Sci.*, 792:62–68, 2019.

33 B. Pondělíček. On weakly commutative semigroups. *Czechoslovak Math. J.*, 25(100):20–23, 1975.

34 K. U. Schulz. Makanin's algorithm for word equations — Two improvements and a generalization. In K. U. Schulz, editor, *Word Equations and Related Topics*, volume 572 of *Lecture Notes in Comput. Sci.*, pages 85–150, Heidelberg, 1990. Springer-Verlag.

35 M.-P. Schützenberger. Sur le produit de concaténation non ambigu. *Semigroup Forum*, 13:47–75, 1976.

36 Z. Sela. Diophantine geometry over groups VIII: Stability. *Ann. of Math.*, 177:787–868, 2013.

37 I. Simon. Piecewise testable events. In H. Barkhage, editor, *Automata Theory and Formal Languages, 2nd GI Conference, Kaiserslautern, May 22–23, 1975*, volume 33 of *Lecture Notes in Comput. Sci.*, pages 214–222. Springer-Verlag, 1975.

38 C. M. Weinbaum. Word equation $ABC = CDA$, $B \neq D$. *Pacific J.Math.*, 213:157–162, 2004.