The Complexity of Checking Quasi-Identities over Finite Algebras with a Mal'cev Term

Erhard Aichinger
Institute for Algebra, Johannes Kepler Universität Linz, Austria

Simon Grünbacher
Institute for Algebra, Johannes Kepler Universität Linz, Austria

Abstract

We consider finite algebraic structures and ask whether every solution of a given system of equations satisfies some other equation. This can be formulated as checking the validity of certain first order formulae called *quasi-identities*. Checking the validity of quasi-identities is closely linked to solving systems of equations. For systems of equations over finite algebras with finitely many fundamental operations, a complete P/NPC dichotomy is known, while the situation appears to be more complicated for single equations. The complexity of checking the validity of a quasi-identity lies between the complexity of term equivalence (checking whether two terms induce the same function) and the complexity of solving systems of polynomial equations. We prove that for each finite algebra with a Mal'cev term and finitely many fundamental operations, checking the validity of quasi-identities is coNP-complete if the algebra is not abelian, and in P when the algebra is abelian.

2012 ACM Subject Classification Mathematics of computing → Combinatorial algorithms; Theory of computation → Complexity classes

Keywords and phrases quasi-identities, conditional identities, systems of equations

Digital Object Identifier 10.4230/LIPIcs.STACS.2023.4

Funding Supported by the Austrian Science Fund FWF P33878: Equations in Universal Algebra.

Acknowledgements The authors thank M. Behrisch for discussions on this problem, the referees of a previous version for their criticism that helped to improve the result, and the referees of the present version for several suggestions improving the presentation.

1 Introduction

The computational complexity of solving equations over some fixed finite algebra has been an active field of research for the last two decades, and several different problems related to solving equations have been studied in the literature. The most general problem that we consider is that of solving systems of polynomial equations (PolSysSat). For this problem, Goldmann and Russel [9] have proved a P/NPC dichotomy for groups, which was later generalized to algebras in congruence modular varieties by Larose and Zádori [15, 22]. For arbitrary finite algebras with finitely many fundamental operations, a dichotomy follows from [3, 23] because solving polynomial systems can be seen as a constraint satisfaction problem [15, Theorem 2.2]. If the input is restricted to a single equation (PolSat), it becomes easier for some algebras, including nilpotent rings and groups [12]. No dichotomy theorem is known for PolSat and in fact, recent results suggest that such a dichotomy might not exist [21, 14]. The situation is similar for the problem of checking whether two polynomials induce the same function (PolEqv).

In this paper, we ask whether all solutions of a system of term equations over an algebraic structure $\mathbf{A} = (A, (f_i)_{i \in I})$ satisfy some other equation. Solving this problem, we can determine whether a set $S = \{\boldsymbol{x} \in A^n \mid \bigwedge_{i \in \underline{k}} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})\}$ defined as the solution set of a system of term equations is contained in another set $U = \{\boldsymbol{x} \in A^n \mid u(\boldsymbol{x}) = v(\boldsymbol{x})\}$.

40th International Symposium on Theoretical Aspects of Computer Science (STACS 2023).
Editors: Petra Berenbrink, Patricia Bouyer, Anuj Dawar, and Mamadou Moustapha Kanté;
Article No. 4; pp. 4:1–4:12
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

(For $k \in \mathbb{N}_0$, we use $\underline{k}$ as an abbreviation for $\{1, 2, \dots, k\}$.) This problem arises naturally in algebraic geometry and has therefore motivated considerable mathematical insights: for algebraically closed fields, Hilbert's Nullstellensatz reduces this question to the radical membership problem of a multivariate polynomial ring, and in a finite field $\mathbb{F}_q$, a polynomial $u \in \mathbb{F}_q[x_1, \dots, x_n]$ vanishes at all solutions of $s_1(x_1, \dots, x_n) = \cdots = s_k(x_1, \dots, x_n) = 0$ if and only if there are $a_1, \dots, a_k, b_1, \dots, b_n \in \mathbb{F}_q[x_1, \dots, x_n]$ such that $u = \sum_{i=1}^{k} a_i\, s_i + \sum_{i=1}^{n} b_i\, (x_i^q - x_i)$ ([19], [6, Theorem 7]). In the present note, we consider this problem for finite algebras from the viewpoint of universal algebra [5, 17], and we seek to determine its computational complexity. For an algebraic structure $\mathbf{A} = (A, (f_i)_{i \in I})$, an *equation* is a formula $s(x_1, \dots, x_n) = t(x_1, \dots, x_n)$, where s and t are terms built from the operation symbols f_i and the variables $x_1, \dots, x_n$. A *solution* to this equation is a tuple $(a_1, \dots, a_n) \in A^n$ such that $s(a_1, \dots, a_n) = t(a_1, \dots, a_n)$. Given $k, n \in \mathbb{N}$, equations $s_i(x_1, \dots, x_n) = t_i(x_1, \dots, x_n)$ $(i \in \{1, \dots, k\})$ and an equation $u(x_1, \dots, x_n) = v(x_1, \dots, x_n)$, the solutions of $\bigwedge_{i \in \underline{k}} s_i(x_1, \dots, x_n) = t_i(x_1, \dots, x_n)$ are contained in the solutions of $u(x_1, \dots, x_n) = v(x_1, \dots, x_n)$ if and only if the first order formula

$$\forall \boldsymbol{x} \,:\, \big(\bigwedge_{i \in \underline{k}} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})\big) \Rightarrow u(\boldsymbol{x}) = v(\boldsymbol{x})$$

holds in $\mathbf{A}$. Such a formula is called a *conditional identity* or *quasi-identity.* For a given algebra $\mathbf{A}$, the problem QuasiIdVal($\mathbf{A}$) is to decide whether a given quasi-identity holds in $\mathbf{A}$. For example, in the group S_3, $\forall x_1, x_2 \,:\, (x_1{\cdot}(x_1{\cdot}x_1) = 1 \wedge x_2{\cdot}(x_2{\cdot}x_2) = 1) \Rightarrow x_1{\cdot}x_2 = x_2{\cdot}x_1$ is a valid quasi-identity expressing that all elements of order dividing 3 commute. For semigroups, this decision problem has been investigated in [20].

▶ **Definition 1.1.** Let $\mathbf{A}$ be an algebra. Then the *quasi-identity validity* problem over $\mathbf{A}$, QuasiIdVal($\mathbf{A}$), is defined as follows: Given terms $s_1, t_1, \dots, s_k, t_k, u, v$ over the variables $(x_i)_{i \in \underline{n}}$ in the language of the algebra $\mathbf{A}$, determine whether

$$\forall \boldsymbol{a} \in A^n : \big(\bigwedge_{i \in \underline{k}} s_i(\boldsymbol{a}) = t_i(\boldsymbol{a})\big) \Rightarrow u(\boldsymbol{a}) = v(\boldsymbol{a})$$

holds.

We refer to $\bigwedge_{i \in \underline{k}} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})$ as the *precondition* and to $u(\boldsymbol{x}) = v(\boldsymbol{x})$ as the *conclusion* of the quasi-identity. The length of the input is defined by $l := (\sum_{i=1}^{k} ||s_i|| + ||t_i||) + ||u|| + ||v||$, where $||t||$ denotes the *length* of t as defined, e.g., in [2]. There are constants $c_1, c_2 \in \mathbb{R}$ such that over a finite alphabet with at least two letters, we can encode the input into a string of length x with $c_1\, l \leq x \leq c_2\, l\, \log(l)$; the $\log(l)$ factor comes from the fact that the input terms may contain at most l different variables, for which we find names of length $\leq c_3 \log(l)$. Since our results are a mere P/coNPC-distinction, measuring computation time in terms of l is a sufficient degree of precision. We will use the notions from complexity theory as they are defined in [18]; in particular, coNP-completeness is to be understood with respect to polynomial time many-one reductions. For a finite algebra of finite type, a tuple $\boldsymbol{a}$ from A^n for which the precondition is true and the conclusion is false can serve as a certificate for the answer "no", and therefore the problem is in coNP. Since QuasiIdVal is closely related to solving polynomial systems and checking identities, the complexity of QuasiIdVal is often determined by the connection to these problems.

2 Complexity determined by the relation to other problems on terms and polynomials

For a finite algebra $\mathbf{A}$, we first compare $\textsc{QuasiIdVal}(\mathbf{A})$ to the problem $\textsc{PolSysSat}(\mathbf{A})$ of solving systems of polynomial equations over $\mathbf{A}$, which was studied, e.g., in [15]. Here, a *polynomial equation* over $\mathbf{A}$ is a formula $s(x_1, \dots, x_n, b_1, \dots, b_m) = t(x_1, \dots, x_n, b_1, \dots, b_m)$, where s, t are terms and $b_1, \dots, b_m \in A$; a *solution* is an $\boldsymbol{a} \in A^n$ with

$$s(a_1, \dots, a_n, b_1, \dots, b_m) = t(a_1, \dots, a_n, b_1, \dots, b_m).$$

We first observe that $\textsc{PolSysSat}(\mathbf{A})$ is harder than $\textsc{QuasiIdVal}(\mathbf{A})$ because an instance of (the negation of) $\textsc{QuasiIdVal}(\mathbf{A})$ can be reduced to solving a constant number of instances of $\textsc{PolSysSat}(\mathbf{A})$: given an instance

$$\forall \boldsymbol{x} \in A^n : (\bigwedge_{i \in \underline{k}} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})) \Rightarrow u(\boldsymbol{x}) = v(\boldsymbol{x}) \tag{2.1}$$

of $\textsc{QuasiIdVal}(\mathbf{A})$, we observe that (2.1) is not valid if and only if there are $a, b \in A$ with $a \neq b$ such that the system

$$(\bigwedge_{i \in \underline{k}} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})) \wedge u(\boldsymbol{x}) = a \wedge v(\boldsymbol{x}) = b \tag{2.2}$$

has a solution. We note that this reduction from (the negation of) $\textsc{QuasiIdVal}$ to $\textsc{PolSysSat}$ is not a many-one reduction, but just a truth table reduction: we solve $|A| \cdot (|A| - 1)$ polynomial systems in order to find a counterexample to the validity of a given quasi-identity.

When $\mathbf{A}$ has all constant functions in its term operations, which means that each polynomial function of $\mathbf{A}$ is a term function, and $|A| \geq 2$, then $\textsc{PolSysSat}(\mathbf{A})$ can be reduced to (the negation of) $\textsc{QuasiIdVal}(\mathbf{A})$ by observing that the system $\bigwedge_{i \in \underline{k}} p_i(\boldsymbol{x}) = q_i(\boldsymbol{x})$ has a solution if and only if

$$(\bigwedge_{i \in \underline{k}} p_i(\boldsymbol{x}) = q_i(\boldsymbol{x})) \Rightarrow y = z$$

is not valid, where y, z are distinct variables that do not appear among the x_i's. Without constants, this reduction shows that $\textsc{QuasiIdVal}(\mathbf{A})$ is harder than solving systems of *term* equations ($\textsc{TermSysSat}(\mathbf{A})$). If $\mathbf{A}$ is a group or a ring, every system of term equations is satisfied by $\boldsymbol{x} = (1, \dots, 1)$ (respectively $\boldsymbol{x} = (0, \dots, 0)$). This means that solving systems of term equations is trivial for groups and rings, and therefore for these algebras, the relation to $\textsc{TermSysSat}$ does not yield a meaningful lower bound on the complexity of $\textsc{QuasiIdVal}(\mathbf{A})$.

Such a bound can be obtained by comparing $\textsc{QuasiIdVal}(\mathbf{A})$ to the term equivalence problem $\textsc{TermEqv}(\mathbf{A})$. This is the problem that asks whether two given terms s, t induce the same function on $\mathbf{A}$. The quasi-identity validity problem is at least as hard as checking the validity of a single term equality because $\forall \boldsymbol{x} \in A^n : s(\boldsymbol{x}) = t(\boldsymbol{x})$ is valid if and only if the quasi-identity

$$\forall \boldsymbol{x} \in A^n, y \in A : y = y \Rightarrow s(\boldsymbol{x}) = t(\boldsymbol{x})$$

holds. For an algebra $\mathbf{A}$, let $\mathrm{Clo}(\mathbf{A})$ denote the set of its finitary term functions (this set has also been called the *clone* of $\mathbf{A}$). Strengthening the relation between $\textsc{TermEqv}$ and $\textsc{QuasiIdVal}$ given above, we have that for every algebra $\mathbf{B}$ of finite type with $\mathrm{Clo}(\mathbf{B}) \subseteq$

$\mathrm{Clo}(\mathbf{A})$, TermEqv($\mathbf{B}$) can be reduced to QuasiIdVal($\mathbf{A}$); we explain this reduction by an example. Let $\mathbf{A} = (A, \cdot)$ be a group, and let $\mathbf{B} = (B, f, g)$ with $f(x, y) := y \cdot (x \cdot y)$ and $g(x) := x \cdot x$. Suppose that we want to check whether the equality

$$f(g(x_1), f(x_2, x_3)) = g(x_3) \tag{2.3}$$

is valid in $\mathbf{B}$. Then by replacing f and g with their definition in terms of the operation $\cdot$, we obtain that (2.3) is valid in $\mathbf{B}$ if and only if the quasi-identity

$$x_1 = x_1 \Rightarrow (x_3 \cdot (x_2 \cdot x_3)) \cdot ((x_1 \cdot x_1) \cdot (x_3 \cdot (x_2 \cdot x_3))) = x_3 \cdot x_3 \tag{2.4}$$

holds in $\mathbf{A}$. However, in this example, we see that the quasi-identity in (2.4) is longer than the equality in (2.3) from which we started. In general, the reduction given above may produce a quasi-identity whose size is exponential in the length of the given equality, and therefore does not qualify as a polynomial time many-one reduction. This exponential increase in length is avoided if we introduce variables for all subterms that appear in (2.3). We observe that (2.3) is valid in $\mathbf{B}$ if and only if the quasi-identity

$$(z_1 = x_1 \cdot x_1 \wedge z_2 = x_3 \cdot (x_2 \cdot x_3) \wedge z_3 = z_2 \cdot (z_1 \cdot z_2) \wedge z_4 = x_3 \cdot x_3) \Rightarrow z_3 = z_4$$

holds in $\mathbf{A}$, and the size of the quasi-identity obtained in this way is bounded by a polynomial in the size of the input equality. Hence in this way, we obtain a polynomial time reduction of TermEqv($\mathbf{B}$) to QuasiIdVal($\mathbf{A}$). The problem TermEqv has been investigated for some classes of finite groups and rings (see e.g. [4, 9]). In [13], it is proved that for every non-nilpotent group $\mathbf{A}$, there is an algebra $\mathbf{B}$ with $\mathrm{Clo}(\mathbf{B}) = \mathrm{Clo}(\mathbf{A})$ such that TermEqv($\mathbf{B}$) is coNP-complete; this implies that every non-nilpotent group has a coNP-complete quasi-identity validity problem. We summarize these consequences of the literature in Table 1. In this table, we do not require that a ring has a unit element; a ring is *nilpotent* if there is a $k \in \mathbb{N}$ such that every product of at least k elements is 0, and a ring $\mathbf{R}$ with $r \cdot s = 0$ for all $r, s \in R$ is called a *zero ring*. Theorem 3.1 establishes that in both cases in which the complexity of QuasiIdVal is referred to as *open* in the above table, the answer is *coNP-complete*.

Table 1 Complexity of the studied problems as known before the present note.

$\mathbf{A}$	TermEqv($\mathbf{A}$)	QuasiIdVal($\mathbf{A}$)	PolSysSat($\mathbf{A}$)
module/abelian group/zero ring	P	P	P ([9, 15])
nonabelian nilpotent group	P ([9])	open	NPC ([9])
non-nilpotent solvable group	partially open	coNPC ([13])	NPC ([9])
non-solvable group	coNPC ([9])	coNPC	NPC ([9])
non-zero nilpotent ring	P ([4])	open	NPC ([15])
non-nilpotent ring	coNPC ([4])	coNPC	NPC ([15])

In [20], M. Volkov constructs a 10-element semigroup $\mathbf{Q}$ such that TermEqv($\mathbf{Q}$) is in P and QuasiIdVal($\mathbf{Q}$) is coNP-complete. Combining the results of the present note with [9, 4], we obtain that every finite nilpotent nonabelian group and every finite nilpotent nonzero ring, such as the quaternion group and the ring $2\mathbb{Z}_8$, have a tractable term equivalence and a coNP-complete quasi-identity validity problem.

3 Complexity for finite Mal'cev algebras

Groups and rings are all contained in the larger class of Mal'cev algebras. An algebra $\mathbf{A}$ is a *Mal'cev algebra* if it has a ternary term function $M \in \mathrm{Clo}(\mathbf{A})$ (called a *Mal'cev term*) such that $M(a,b,b) = M(b,b,a) = a$ for all $a, b \in A$ (cf. [16, 17]). For a group $M(x,y,z) = xy^{-1}z$, and for a ring $M(x,y,z) = x - y + z$ are examples of Mal'cev terms. In the present note, we show that for a finite Mal'cev algebra of finite type (i.e., having finitely many fundamental operations), QuasiIdVal is either in P or coNP-complete, and that the dividing line is the same as for PolSysSat. This dividing line can be expressed using a notion from universal algebra. Denoting the set of n-ary term functions of an algebra $\mathbf{A}$ by $\mathrm{Clo}_n(\mathbf{A})$, the algebra $\mathbf{A}$ is called *abelian* if for all $m \in \mathbb{N}$, for all $t \in \mathrm{Clo}_{1+m}(\mathbf{A})$ and for all $a, b \in A$ and $\boldsymbol{c}, \boldsymbol{d} \in A^m$ with $t(a, \boldsymbol{c}) = t(a, \boldsymbol{d})$, also $t(b, \boldsymbol{c}) = t(b, \boldsymbol{d})$ holds [17, Definition 4.146]. A group is abelian in this sense if and only if its operation is commutative, i.e., it is abelian in the sense of classic algebra, and a ring is abelian if and only if it is a zero ring. The main result of this note, proved in Section 5, is:

▶ **Theorem 3.1.** *Let $\mathbf{A}$ be a finite nonabelian Mal'cev algebra of finite type. Then QuasiIdVal$(\mathbf{A})$ is coNP-complete.*

If $\mathbf{A}$ is abelian, then it follows from [15] that PolSysSat$(\mathbf{A})$ is in P. Then as observed at the beginning of Section 2, the validity of a given quasi-identity Φ can be determined by checking the solvability of $|A| \cdot (|A| - 1)$ systems of polynomial equations. Hence we obtain:

▶ **Corollary 3.2.** *Let $\mathbf{A}$ be a finite Mal'cev algebra of finite type. Then QuasiIdVal$(\mathbf{A})$ is in P if $\mathbf{A}$ is abelian, and coNP-complete otherwise.*

4 Preliminaries on Mal'cev Algebras

Our proof of Theorem 3.1 requires the notions *congruence* and *commutator* from universal algebra, and we therefore introduce these briefly. A *congruence relation* of an algebra $\mathbf{A} = (A, F)$ is an equivalence relation α on A that satisfies the *compatibility condition*

$$(a_1, b_1) \in \alpha, \dots (a_n, b_n) \in \alpha \Longrightarrow (f(a_1, \dots, a_n), f(b_1, \dots, b_n)) \in \alpha$$

for each $n \in \mathbb{N}$, for each n-ary basic operation f of $\mathbf{A}$, and for each $\boldsymbol{a}, \boldsymbol{b} \in A^n$. For such an α and for an n-ary basic operation f of $\mathbf{A}$, the operation $f_\alpha(a_1/\alpha, \dots, a_n/\alpha) := f(a_1, \dots, a_n)/\alpha$ is well-defined; this allows to define a factor algebra $\mathbf{A}/\alpha$. For a group $\mathbf{G}$, each congruence relation α is of the form

$$\alpha = \{(g_1, g_2) \in G \times G \mid g_1^{-1} g_2 \in N_\alpha\},$$

where N_α is a normal subgroup of $\mathbf{G}$; similarly, every congruence of a ring $\mathbf{R}$ is of the form

$$\alpha = \{(r_1, r_2) \in R \times R \mid r_1 - r_2 \in I_\alpha\}$$

for some ideal I_α of $\mathbf{R}$. For $k \in \mathbb{N}$ and $\boldsymbol{a} = (a_1, \dots, a_k)$ and $\boldsymbol{b} = (b_1, \dots, b_k) \in A^k$, we write $\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b})$ for the smallest congruence relation on A containing $\{(a_1, b_1), \dots, (a_k, b_k)\}$ as a subset, and call $\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b})$ the *congruence* on $\mathbf{A}$ that is *generated by* $\{(a_1, b_1), \dots, (a_k, b_k)\}$. For the algebras that we consider, the generated congruence has a useful description using term operations. In fact, if $\mathbf{A}$ is a Mal'cev algebra, we have

$$\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}) = \{(t(\boldsymbol{a}, \boldsymbol{e}), t(\boldsymbol{b}, \boldsymbol{e})) \mid m \in \mathbb{N}, t \in \mathrm{Clo}_{k+m}(\mathbf{A}), \boldsymbol{e} \in A^m\};$$

it is easy to verify that the right hand side of this equation is a subuniverse of $\mathbf{A} \times \mathbf{A}$ containing $0_A = \{(a, a) \mid a \in A\}$; in a Mal'cev algebra, all subuniverses of $\mathbf{A} \times \mathbf{A}$ containing 0_A are congruence relations (cf., e.g., [11, Lemma 5.22]). We denote the set of congruence relations on $\mathbf{A}$ by $\mathrm{Con}(\mathbf{A})$.

The other concept from universal algebra that we use is the *commutator*. Here, one associates a congruence relation γ, denoted by $[\alpha, \beta]$, with every pair of congruences α, β from $\mathrm{Con}(\mathbf{A})$. The universal algebraic construction generalizes the *commutator subgroup* $[N, M]$ of two normal subgroups N, M of G, which is the subgroup generated by $\{n^{-1}m^{-1}nm \mid n \in N, m \in M\}$, and the *ideal product* IJ of two ideals I, J in rings, which is the ideal generated by $\{ij \mid i \in I, j \in J\} \cup \{ji \mid i \in I, j \in J\}$. For generalizing these concepts to arbitrary universal algebras, one starts by defining a relation $C(\alpha, \beta; \eta)$ between three congruences α, β, η that is designed to guarantee that $[\alpha, \beta] \leq \gamma$. Doing this formally, we say that for $\alpha, \beta, \eta \in \mathrm{Con}(\mathbf{A})$, the congruence α *centralizes* β *modulo* η if for all $m, n \in \mathbb{N}$, for all $\boldsymbol{a}, \boldsymbol{b} \in A^m$ and $\boldsymbol{c}, \boldsymbol{d} \in A^n$ with $\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}) \leq \alpha$ and $\Theta_{\mathbf{A}}(\boldsymbol{c}, \boldsymbol{d}) \leq \beta$ and for all $t \in \mathrm{Clo}_{m+n}(\mathbf{A})$ we have

$$(t(\boldsymbol{a}, \boldsymbol{c}), t(\boldsymbol{a}, \boldsymbol{d})) \in \eta \Rightarrow (t(\boldsymbol{b}, \boldsymbol{c}), t(\boldsymbol{b}, \boldsymbol{d})) \in \eta. \tag{4.1}$$

The *commutator* of α and β, denoted by $[\alpha, \beta]$, is then defined to be the smallest congruence relation η on $\mathbf{A}$ such that α centralizes β modulo η. What is given here is a reformulation of [8, Definition 3.2(2)]; other sources give slightly different, but equivalent, definitions of the commutator. For example, in [17, Definition 4.148], m is restricted to be equal to 1, which gives an equivalent condition (a rough explanation of this equivalence is that in the implication (4.1), we may change $\boldsymbol{a}$ to $\boldsymbol{b}$ by changing one component of $\boldsymbol{a}$ at a time, repeating this m times, see also [1, Proposition 2.1(2)]). Using the concept of commutators, we see that an algebra $\mathbf{A}$ is abelian if and only if $[1_A, 1_A] = 0_A$.

From this definition, it is not easy to determine the commutator $[\alpha, \beta]$ of two congruences, and therefore one seeks descriptions of $[\alpha, \beta]$ that allow us to compute its elements more directly. In Lemmas 4.2 and 4.3, we provide parametrizations of those commutators $[\alpha, \beta]$ where one of α and β is equal to $1_A = A \times A$. It is known that if $\mathbf{A}$ has a Mal'cev term, then $[\alpha, \beta] = [\beta, \alpha]$ ([17, Exercise 4.156(13)], a proof is written in [1, Lemma 2.5]). Hence $[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A] = [1_A, \Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b})]$.

▶ **Definition 4.1.** Let $\mathbf{A}$ be an algebra, let $k, m \in \mathbb{N}$, and let $t \in \mathrm{Clo}_{k+m}(\mathbf{A})$. Then $\boldsymbol{z} \in A^m$ is called a *right zero* of t if $t(\boldsymbol{x}, \boldsymbol{z}) = t(\boldsymbol{y}, \boldsymbol{z})$ for all $\boldsymbol{x}, \boldsymbol{y} \in A^k$.

▶ **Lemma 4.2.** *Let $\mathbf{A}$ be a Mal'cev algebra, let $k \in \mathbb{N}$, and let $\boldsymbol{a}, \boldsymbol{b} \in A^k$. Then $[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A]$ $= \{(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \mid m \in \mathbb{N}, \boldsymbol{w} \in A^m$, and $t \in \mathrm{Clo}_{k+m}(\mathbf{A})$ such that t has a right zero$\}$.*

Proof. We define

$$\Psi(\boldsymbol{a}, \boldsymbol{b}) := \{(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \mid m \in \mathbb{N}, \boldsymbol{w} \in A^m, t \in \mathrm{Clo}_{k+m}(\mathbf{A}), t \text{ has a right zero}\}.$$

For $\supseteq$, we observe that for a term t with a right zero $\boldsymbol{z}$, we have $t(\boldsymbol{a}, \boldsymbol{z}) = t(\boldsymbol{b}, \boldsymbol{z})$. Now using the term condition (4.1) (for $t'(\boldsymbol{x}, \boldsymbol{y}) := t(\boldsymbol{y}, \boldsymbol{x})$ and $\eta := [1_A, \Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b})]$) and observing that, obviously, $\Theta_{\mathbf{A}}(\boldsymbol{z}, \boldsymbol{w}) \leq 1_A$, we obtain $(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \in [1_A, \Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b})]$.

For $\subseteq$, we first show that $\Psi(\boldsymbol{a}, \boldsymbol{b})$ is a congruence of $\mathbf{A}$. To this end, we show that for all $n, p \in \mathbb{N}$, $s \in \mathrm{Clo}_{n+p}(\mathbf{A})$, $(c_1, d_1), \ldots, (c_n, d_n) \in \Psi(\boldsymbol{a}, \boldsymbol{b})$ and $\boldsymbol{e} \in A^p$, we have

$$(s(c_1, \ldots, c_n, \boldsymbol{e}), s(d_1, \ldots, d_n, \boldsymbol{e})) \in \Psi(\boldsymbol{a}, \boldsymbol{b}). \tag{4.2}$$

For each $i \in \underline{n}$, let t_i be a term function with right zero $\boldsymbol{z}^{(i)}$ and let $\boldsymbol{w}^{(i)}$ be a tuple from $\mathbf{A}$ such that $(c_i, d_i) = (t_i(\boldsymbol{a}, \boldsymbol{w}^{(i)}), t_i(\boldsymbol{b}, \boldsymbol{w}^{(i)}))$. Let

$$s'(\boldsymbol{x}, \boldsymbol{y}^{(1)}, \dots, \boldsymbol{y}^{(n)}, \boldsymbol{z}) := s(t_1(\boldsymbol{x}, \boldsymbol{y}^{(1)}), \dots, t_n(\boldsymbol{x}, \boldsymbol{y}^{(n)}), \boldsymbol{z}).$$

Since $\boldsymbol{z}^{(i)}$ is a right zero of t_i, $(\boldsymbol{z}^{(1)}, \dots, \boldsymbol{z}^{(n)}, \boldsymbol{e})$ is a right zero of s'. Hence

$$(s'(\boldsymbol{a}, \boldsymbol{w}^{(1)}, \dots, \boldsymbol{w}^{(n)}, \boldsymbol{e}), s'(\boldsymbol{b}, \boldsymbol{w}^{(1)}, \dots, \boldsymbol{w}^{(n)}, \boldsymbol{e})) \in \Psi(\boldsymbol{a}, \boldsymbol{b}),$$

and thus $(s(\boldsymbol{c}, \boldsymbol{e}), s(\boldsymbol{d}, \boldsymbol{e})) \in \Psi(\boldsymbol{a}, \boldsymbol{b})$, completing the proof of (4.2).

From (4.2), we see that $\Psi(\boldsymbol{a}, \boldsymbol{b})$ is a subuniverse of $\mathbf{A} \times \mathbf{A}$ that contains the diagonal 0_A as a subset. Since $\mathbf{A}$ is a Mal'cev algebra, this implies that $\Psi(\boldsymbol{a}, \boldsymbol{b})$ is a congruence of $\mathbf{A}$.

Next, we show that $\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b})$ centralizes 1_A modulo $\Psi(\boldsymbol{a}, \boldsymbol{b})$. To this end, let $f, g \in A$ with $(f, g) \in \Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b})$, let $m \in \mathbb{N}$, let $\boldsymbol{c}, \boldsymbol{d} \in A^m$, and let $t \in \mathrm{Clo}_{1+m}(\boldsymbol{A})$ be such that $(t(f, \boldsymbol{c}), t(f, \boldsymbol{d})) \in \Psi(\boldsymbol{a}, \boldsymbol{b})$. For proving the centralizing property, we need to show $(t(g, \boldsymbol{c}), t(g, \boldsymbol{d})) \in \Psi(\boldsymbol{a}, \boldsymbol{b})$. First, we observe that since $(f, g) \in \Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b})$, there are $n \in \mathbb{N}$, a term function $r \in \mathrm{Clo}_{k+n}(\mathbf{A})$ and $\boldsymbol{e} \in A^n$ such that $f = r(\boldsymbol{a}, \boldsymbol{e})$ and $g = r(\boldsymbol{b}, \boldsymbol{e})$. Denoting the Mal'cev term of $\mathbf{A}$ by $M(x, y, z)$, we define a term function $q \in \mathrm{Clo}_{k+(m+n+m+1)}(\mathbf{A})$ by

$$q(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v}, w) := M(t(r(\boldsymbol{x}, \boldsymbol{u}), \boldsymbol{y}), t(r(\boldsymbol{x}, \boldsymbol{u}), \boldsymbol{v}), t(w, \boldsymbol{v})).$$

Then $(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v}, w) := (\boldsymbol{c}, \boldsymbol{e}, \boldsymbol{c}, g)$ is a right zero of q because for all $\boldsymbol{x}, \boldsymbol{x'} \in A^k$, we have

$$\begin{aligned} q(\boldsymbol{x}, \boldsymbol{c}, \boldsymbol{e}, \boldsymbol{c}, g) &= M(t(r(\boldsymbol{x}, \boldsymbol{e}), \boldsymbol{c}), t(r(\boldsymbol{x}, \boldsymbol{e}), \boldsymbol{c}), t(g, \boldsymbol{c})) \\ &= t(g, \boldsymbol{c}) \\ &= M(t(r(\boldsymbol{x'}, \boldsymbol{e}), \boldsymbol{c}), t(r(\boldsymbol{x'}, \boldsymbol{e}), \boldsymbol{c}), t(g, \boldsymbol{c})) \\ &= q(\boldsymbol{x'}, \boldsymbol{c}, \boldsymbol{e}, \boldsymbol{c}, g). \end{aligned}$$

Hence

$$(q(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{c}, g), q(\boldsymbol{b}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{c}, g)) \in \Psi(\boldsymbol{a}, \boldsymbol{b}).$$

We have

$$\begin{aligned} q(\boldsymbol{a}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{c}, g) &= M(t(r(\boldsymbol{a}, \boldsymbol{e}), \boldsymbol{d}), t(r(\boldsymbol{a}, \boldsymbol{e}), \boldsymbol{c}), t(g, \boldsymbol{c})) \\ &= M(t(f, \boldsymbol{d}), t(f, \boldsymbol{c}), t(g, \boldsymbol{c})) \end{aligned}$$

and

$$\begin{aligned} q(\boldsymbol{b}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{c}, g)) &= M(t(r(\boldsymbol{b}, \boldsymbol{e}), \boldsymbol{d}), t(r(\boldsymbol{b}, \boldsymbol{e}), \boldsymbol{c}), t(g, \boldsymbol{c})) \\ &= M(t(g, \boldsymbol{d}), t(g, \boldsymbol{c}), t(g, \boldsymbol{c})) \\ &= t(g, \boldsymbol{d}). \end{aligned}$$

From the definition of $\Psi(\boldsymbol{a}, \boldsymbol{b})$, we therefore obtain

$$\big(M(t(f, \boldsymbol{d}), t(f, \boldsymbol{c}), t(g, \boldsymbol{c})),\ t(g, \boldsymbol{d})\big) \in \Psi(\boldsymbol{a}, \boldsymbol{b}).$$

Since $(t(f, \boldsymbol{c}), t(f, \boldsymbol{d})) \in \Psi(\boldsymbol{a}, \boldsymbol{b})$, we have that $M(t(f, \boldsymbol{d}), t(f, \boldsymbol{c}), t(g, \boldsymbol{c}))$ is congruent modulo $\Psi(\boldsymbol{a}, \boldsymbol{b})$ to $M(t(f, \boldsymbol{d}), t(f, \boldsymbol{d}), t(g, \boldsymbol{c})) = t(g, \boldsymbol{c})$. Thus we have $(t(g, \boldsymbol{c}), t(g, \boldsymbol{d})) \in \Psi(\boldsymbol{a}, \boldsymbol{b})$. Hence $\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b})$ centralizes 1_A modulo $\Psi(\boldsymbol{a}, \boldsymbol{b})$. Since $[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A]$ is defined as the intersection of all congruences ψ such that $\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b})$ centralizes 1_A modulo ψ, we obtain $[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A] \subseteq \Psi(\boldsymbol{a}, \boldsymbol{b})$. ◀

The next lemma tells that we can find one single term to parametrize commutators of the form $[\alpha, 1_A]$.

▶ **Lemma 4.3.** *Let* $\mathbf{A}$ *be a finite Mal'cev algebra and let* $k \in \mathbb{N}$*. Then there exist* $m \in \mathbb{N}$ *and* $t \in \operatorname{Clo}_{k+m}(\mathbf{A})$ *such that*

$$\begin{aligned}&t \text{ has a right zero, and for all } \boldsymbol{a}, \boldsymbol{b} \in A^k,\\ &\qquad\qquad \text{we have } [\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A] = \{(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \mid \boldsymbol{w} \in A^m\}.\end{aligned} \tag{4.3}$$

Proof. For each $m \in \mathbb{N}$ and each $t \in \operatorname{Clo}_{k+m}(\mathbf{A})$ with a right zero, let

$$\Psi(t, \boldsymbol{a}, \boldsymbol{b}) := \{(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \mid \boldsymbol{w} \in A^m\}.$$

We order the terms in

$$T := \{t \in \operatorname{Clo}_{k+m}(\mathbf{A}) \mid m \in \mathbb{N}, t \text{ has a right zero}\}$$

by $t_1 \leq t_2$ if $\Psi(t_1, \boldsymbol{a}, \boldsymbol{b}) \subseteq \Psi(t_2, \boldsymbol{a}, \boldsymbol{b})$ for all $\boldsymbol{a}, \boldsymbol{b} \in A^k$. The relation $\leq$ is a quasi-order. For the term function $\pi(x_1, \dots, x_k, y) := y$, we have $\Psi(\pi, \boldsymbol{a}, \boldsymbol{b}) = 0_A$. Since $\Psi(t, \boldsymbol{a}, \boldsymbol{b})$ can take only finitely many values, there is $t \in T$ such that $\pi \leq t$ and t is maximal in T with respect to $\leq$. Let $m \in \mathbb{N}$ be such that $t \in \operatorname{Clo}_{k+m}(\mathbf{A})$. Our claim is that this t satisfies (4.3). By Lemma 4.2, we have $[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A] \supseteq \{(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \mid \boldsymbol{w} \in A^m\}$ for all $\boldsymbol{a}, \boldsymbol{b} \in A^k$.

For proving the "$\subseteq$"-inclusion of (4.3), suppose that there are $\boldsymbol{c}, \boldsymbol{d} \in A^k$, $(f, g) \in [\Theta_{\mathbf{A}}(\boldsymbol{c}, \boldsymbol{d}), 1_A]$ and $(f, g) \notin \{(t(\boldsymbol{c}, \boldsymbol{w}), t(\boldsymbol{d}, \boldsymbol{w})) \mid \boldsymbol{w} \in A^m\}$. By Lemma 4.2, there are $n \in \mathbb{N}$, an $s \in \operatorname{Clo}_{k+n}(\mathbf{A})$ with a right zero and $\boldsymbol{e} \in A^n$ such that $(f, g) = (s(\boldsymbol{c}, \boldsymbol{e}), s(\boldsymbol{d}, \boldsymbol{e}))$. Denoting the Mal'cev term of $\mathbf{A}$ by $M(x, y, z)$, we define a term function $r \in \operatorname{Clo}_{k+(n+m+m)}(\mathbf{A})$ by

$$r(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) := M(s(\boldsymbol{u}, \boldsymbol{x}), t(\boldsymbol{u}, \boldsymbol{y}), t(\boldsymbol{u}, \boldsymbol{z})).$$

Let $\boldsymbol{h}, \boldsymbol{i}$ be the right zeros of s and t, respectively. Then $(\boldsymbol{h}, \boldsymbol{i}, \boldsymbol{i})$ is a right zero of r. We first show that $t \leq r$. To this end, let $\boldsymbol{a}, \boldsymbol{b} \in A^k$ and $\boldsymbol{w} \in A^m$. We want to show that $(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \in \Psi(r, \boldsymbol{a}, \boldsymbol{b})$. Since $\boldsymbol{h}$ is a right zero of s, we have $s(\boldsymbol{a}, \boldsymbol{h}) = s(\boldsymbol{b}, \boldsymbol{h})$, and thus $(s(\boldsymbol{a}, \boldsymbol{h}), s(\boldsymbol{b}, \boldsymbol{h})) \in 0_A = \Psi(\pi, \boldsymbol{a}, \boldsymbol{b})$. Since $\pi \leq t$, we therefore have $(s(\boldsymbol{a}, \boldsymbol{h}), s(\boldsymbol{b}, \boldsymbol{h})) \in \Psi(t, \boldsymbol{a}, \boldsymbol{b})$ and thus there is $\boldsymbol{v} \in A^m$ such that $(t(\boldsymbol{a}, \boldsymbol{v}), t(\boldsymbol{b}, \boldsymbol{v})) = (s(\boldsymbol{a}, \boldsymbol{h}), s(\boldsymbol{b}, \boldsymbol{h}))$. Hence

$$\begin{aligned}(r(\boldsymbol{a}, \boldsymbol{h}, \boldsymbol{v}, \boldsymbol{w}), r(\boldsymbol{b}, \boldsymbol{h}, \boldsymbol{v}, \boldsymbol{w})) &= \big(M(s(\boldsymbol{a}, \boldsymbol{h}), t(\boldsymbol{a}, \boldsymbol{v}), t(\boldsymbol{a}, \boldsymbol{w})), M(s(\boldsymbol{b}, \boldsymbol{h}), t(\boldsymbol{b}, \boldsymbol{v}), t(\boldsymbol{b}, \boldsymbol{w}))\big)\\ &= \big(M(s(\boldsymbol{a}, \boldsymbol{h}), s(\boldsymbol{a}, \boldsymbol{h}), t(\boldsymbol{a}, \boldsymbol{w})), M(s(\boldsymbol{b}, \boldsymbol{h}), s(\boldsymbol{b}, \boldsymbol{h}), t(\boldsymbol{b}, \boldsymbol{w}))\big)\\ &= (t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})),\end{aligned}$$

and thus $(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \in \Psi(r, \boldsymbol{a}, \boldsymbol{b})$. Hence $\Psi(t, \boldsymbol{a}, \boldsymbol{b}) \subseteq \Psi(r, \boldsymbol{a}, \boldsymbol{b})$ and thus $t \leq r$.

We will now establish that $r \not\leq t$. To this end, we observe that (f, g) is an element of $\Psi(r, \boldsymbol{c}, \boldsymbol{d})$ because

$$(s(\boldsymbol{c}, \boldsymbol{e}), s(\boldsymbol{d}, \boldsymbol{e})) = (r(\boldsymbol{c}, \boldsymbol{e}, \boldsymbol{i}, \boldsymbol{i}), r(\boldsymbol{d}, \boldsymbol{e}, \boldsymbol{i}, \boldsymbol{i})) \in \Psi(r, \boldsymbol{c}, \boldsymbol{d}).$$

Since $(f, g) \in \Psi(r, \boldsymbol{c}, \boldsymbol{d})$ and by assumption $(f, g) \notin \Psi(t, \boldsymbol{c}, \boldsymbol{d})$, we have $r \not\leq t$.

Now $t \leq r$ and $r \not\leq t$ contradict the maximality of t, which completes the proof that

$$[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A] \subseteq \{(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \mid \boldsymbol{w} \in A^m\}$$

for all $\boldsymbol{a}, \boldsymbol{b} \in A^k$. ◀

5 Graphs and the completeness proof

For establishing coNP-completeness, we use an NP-complete problem from graph theory. In our proof, we take the formal viewpoint to consider a *graph* as a relational structure $\mathbb{G} = (G, \rho)$ such that $\rho \subseteq G \times G$ is symmetric. We call the graph $\mathbb{G}$ *loopless* if there is no $v \in G$ with $(v, v) \in \rho$, and we say that $\mathbb{G}$ *contains a triangle* if there are $u, v, w \in G$ with $(u, v) \in \rho$, $(v, w) \in \rho$, $(u, w) \in \rho$. For a graph $\mathbb{H}$, the computational problem $\mathbb{H}$-Coloring studied in [10] asks whether for a finite input graph $\mathbb{G}$, there exists a homomorphism from $\mathbb{G}$ to $\mathbb{H}$. Theorem 1 of [10] states that if $\mathbb{H}$ is loopless and not bipartite, then $\mathbb{H}$-Coloring is NP-complete. Since a bipartite graph cannot contain a triangle, we obtain:

▶ **Corollary 5.1** ([10]). *Let $\mathbb{H}$ be a finite loopless graph that contains a triangle. Then $\mathbb{H}$-Coloring is NP-complete.*

Let $\mathbf{A}$ be a finite Mal'cev algebra and let $\mu \in \operatorname{Con}(\mathbf{A})$. The *difference graph* of $\mathbf{A}$ with respect to μ is the graph $\mathbb{H}_\mu = (H_\mu, \rho_\mu)$, where the set of vertices H_μ is equal to A^2. The set of edges ρ_μ is defined by

$$\rho_\mu = \{(\boldsymbol{a}, \boldsymbol{b}) \mid \boldsymbol{a}, \boldsymbol{b} \in A^2,\ \mu \leq [\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A]\}.$$

Intuitively, we draw an edge between two vectors $\boldsymbol{a}, \boldsymbol{b}$ from A^2 if $\boldsymbol{a}, \boldsymbol{b}$ are "sufficiently different". Here, "sufficiently different" means that the smallest congruence collapsing $\boldsymbol{a}$ and $\boldsymbol{b}$ is still large enough to have its commutator with 1_A above μ.

▶ **Lemma 5.2.** *Let $\mathbf{A}$ be a finite Mal'cev algebra, and let $\mu \in \operatorname{Con}(\mathbf{A})$ with $\mu > 0_A$. Then the difference graph $\mathbb{H}_\mu = (H_\mu, \rho_\mu)$ is loopless.*

Proof. Let $\boldsymbol{a} \in A^2$. We have to show that $(\boldsymbol{a}, \boldsymbol{a})$ is not an edge of $\mathbb{H}_\mu$. Suppose that $(\boldsymbol{a}, \boldsymbol{a}) \in \rho_\mu$. Then from the definition of ρ_μ, we obtain $\mu \leq [\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{a}), 1_A]$. Clearly, $\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{a}) = 0_A$. Furthermore, by [17, Lemma 4.149(i)], the commutator $[\alpha, \beta]$ is always contained in the intersection $\alpha \cap \beta$, and therefore $[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{a}), 1_A] = [0_A, 1_A] = 0_A$. Hence $\mu \leq 0_A$, contradicting the assumption $\mu > 0_A$. Thus $(\boldsymbol{a}, \boldsymbol{a})$ is not an edge of $\mathbb{H}_\mu$, and therefore $\mathbb{H}_\mu$ is loopless. ◀

▶ **Lemma 5.3.** *Let $\mathbf{A}$ be a finite nonabelian Mal'cev algebra. Then there is $\beta \in \operatorname{Con}(\mathbf{A})$ such that $\beta > 0_A$ and $\mathbb{H}_\beta = (H_\beta, \rho_\beta)$ has a triangle.*

Proof. Let ζ be the center of $\mathbf{A}$, i.e., the largest congruence with $[\zeta, 1_A] = 0_A$. We note that from [17, Lemma 4.149(ii)] it follows that such a largest congruence exists: in fact, ζ is the join of all congruences ζ' with $[\zeta', 1_A] = 0_A$. In particular, every congruence $\zeta' \in \operatorname{Con}(\mathbf{A})$ with $[\zeta', 1_A] = 0_A$ satisfies $\zeta' \leq \zeta$. Since $\mathbf{A}$ is nonabelian, $\zeta < 1_A$. Thus there are $a, b \in A$ such that $(a, b) \notin \zeta$. Let

$$\beta := [\Theta_{\mathbf{A}}(a, b), 1_A].$$

We first show that $\beta > 0_A$. Suppose $\beta = 0_A$. Then $[\Theta_{\mathbf{A}}(a, b), 1_A] = 0_A$, and therefore $\Theta_{\mathbf{A}}(a, b) \leq \zeta$. Then $(a, b) \in \zeta$, contradicting the choice of a and b. Hence $\beta > 0_A$. Next, we show that $\mathbb{H}_\beta$ has a triangle. To this end, we consider the vertices $\boldsymbol{u} = \binom{a}{a}$, $\boldsymbol{v} = \binom{a}{b}$, $\boldsymbol{w} = \binom{b}{b}$ of $\mathbb{H}_\beta$. For showing that $(\boldsymbol{u}, \boldsymbol{v})$ is an edge of $\mathbb{H}_\beta$, we observe that $(\boldsymbol{u}, \boldsymbol{v}) \in \rho_\beta$ if and only if $\beta \leq [\Theta_{\mathbf{A}}(\binom{a}{a}, \binom{a}{b}), 1_A]$. Now $\Theta_{\mathbf{A}}(\binom{a}{a}, \binom{a}{b})$ is the smallest congruence containing $\{(a, a), (a, b)\}$ and therefore $\Theta_{\mathbf{A}}(\binom{a}{a}, \binom{a}{b}) = \Theta_{\mathbf{A}}(a, b)$. Hence $\beta = [\Theta_{\mathbf{A}}(a, b), 1_A] = [\Theta_{\mathbf{A}}(\binom{a}{a}, \binom{a}{b}), 1_A]$, and therefore $(\boldsymbol{u}, \boldsymbol{v}) \in \rho_\beta$. Similarly, $(\boldsymbol{u}, \boldsymbol{w})$ and $(\boldsymbol{v}, \boldsymbol{w})$ are edges of $\mathbb{H}_\beta$. ◀

On a set $\mathcal{G}$ of graphs, we can define a quasi-order by $\mathbb{G} \preceq \mathbb{H}$ if there is a homomorphism from $\mathbb{G}$ to $\mathbb{H}$. We say that $\mathbb{G}$ is *maximal* in $\mathcal{G}$ with respect to $\preceq$ if for every $\mathbb{H} \in \mathcal{G}$ with $\mathbb{G} \preceq \mathbb{H}$, we also have $\mathbb{H} \preceq \mathbb{G}$. If $\mathcal{G}$ is finite and nonempty, it must contain at least one maximal element.

▶ **Lemma 5.4.** *Let $\mathbf{A}$ be a finite nonabelian Mal'cev algebra. For each $\gamma \in \mathrm{Con}(\mathbf{A})$, let $\mathbb{H}_\gamma$ be the difference graph of $\mathbf{A}$ with respect to γ. Let $\beta \in \mathrm{Con}(\mathbf{A})$ be such that $\beta > 0_A$ and $\mathbb{H}_\beta$ has a triangle. Let $\mu \in \mathrm{Con}(\mathbf{A})$ be such that $\mu > 0_A$ and $\mathbb{H}_\mu = (H_\mu, \rho_\mu)$ is maximal with respect to $\preceq$ in*

$$\{\mathbb{H}_\alpha \mid \alpha \in \mathrm{Con}(\mathbf{A}), \alpha > 0_A, \mathbb{H}_\beta \preceq \mathbb{H}_\alpha\}.$$

Let $m \in \mathbb{N}$ and let $t(x, y, z_1, \dots, z_m)$ be a term in the language of $\mathbf{A}$ such that its induced term function $t^{\mathbf{A}} \in \mathrm{Clo}_{2+m}(\mathbf{A})$ satisfies (4.3); such a term exists by Lemma 4.3. Let $\mathbb{G} = (G, \rho^{\mathbb{G}})$ be a graph. We assume that $G \cap H_\mu = \varnothing$. Let Φ be the quasi-identity

$$\Big(\bigwedge_{(u,v) \,\in\, \rho^{\mathbb{G}} \,\cup\, \rho_\mu} \big(a = t(x_u, y_u, \boldsymbol{z}_{(u,v)}) \ \wedge\ b = t(x_v, y_v, \boldsymbol{z}_{(u,v)})\big)\Big) \Rightarrow a = b$$

in the variables $\{x_u \mid u \in G \cup H_\mu\} \cup \{y_u \mid u \in G \cup H_\mu\} \cup \{(\boldsymbol{z}_{(u,v)})_i \mid i \in \underline{m}, (u, v) \in \rho^{\mathbb{G}} \cup \rho_\mu\} \cup \{a, b\}$. Then $\mathbb{G} \preceq \mathbb{H}_\mu$ if and only if Φ is not valid in $\mathbf{A}$.

Proof. For the "only if"-direction, let f be a homomorphism from $\mathbb{G}$ to $\mathbb{H}_\mu$. We set the variables in Φ in a way that contradicts the validity of Φ. First, we assign values to a and b such that $(a, b) \in \mu \setminus 0_A$. Next, we assign values to the variables x_u, y_u with $u \in G$. To this end, for each $u \in G$, we set $x_u, y_u \in A$ such that $\left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right) = f(u)$. Then for each pair $(u, v) \in \rho^{\mathbb{G}}$, the fact that f is a homomorphism yields $(f(u), f(v)) \in \rho_\mu$, and therefore we have $\mu \leq [\Theta(\left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right), \left(\begin{smallmatrix} x_v \\ y_v \end{smallmatrix}\right)), 1_A]$. Since $(a, b) \in \mu$, Lemma 4.3 allows us to find $\boldsymbol{z}_{(u,v)} \in A^m$ such that $t(x_u, y_u, \boldsymbol{z}_{(u,v)}) = a$ and $t(x_v, y_v, \boldsymbol{z}_{(u,v)}) = b$. In the next step, we assign values to the variables x_u, y_u with $u \in H_\mu$. To this end, we observe that each $u \in H_\mu$ is an element of A^2, and hence there are $x_u, y_u \in A$ such that $u = \left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right)$. Then for each $(u, v) \in \rho_\mu$, we have $(\left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right), \left(\begin{smallmatrix} x_v \\ y_v \end{smallmatrix}\right)) = (u, v) \in \rho_\mu$. Hence from the definition of ρ_μ, we obtain

$$\mu \leq [\Theta(\left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right), \left(\begin{smallmatrix} x_v \\ y_v \end{smallmatrix}\right)), 1_A].$$

Thus from Lemma 4.3 we obtain $\boldsymbol{z}_{(u,v)} \in A^m$ such that $t(x_u, y_u, \boldsymbol{z}_{(u,v)}) = a$ and $t(x_v, y_v, \boldsymbol{z}_{(u,v)}) = b$. Now this assignment of the variables confirms that Φ is not valid in $\mathbf{A}$.

For the "if"-direction, we assume that that the variables in Φ are assigned such that Φ does not hold and we construct a homomorphism $f : \mathbb{G} \to \mathbb{H}_\mu$. Let $\tau := \Theta_{\mathbf{A}}(a, b)$. Since $a \neq b$, we have $\tau > 0_A$. We first define a mapping g from $\mathbb{H}_\mu$ to $\mathbb{H}_\tau$ by

$$g(u) = \left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right) \text{ for all } u \in H_\mu.$$

Next, we prove that g is a homomorphism from $\mathbb{H}_\mu$ to $\mathbb{H}_\tau$. Since Φ does not hold, its precondition is fulfilled, and therefore for each $(u, v) \in \rho_\mu$, we have $a = t(x_u, y_u, \boldsymbol{z}_{(u,v)})$ and $b = t(x_v, y_v, \boldsymbol{z}_{(u,v)})$. Hence setting $\boldsymbol{a} := \left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right)$ and $\boldsymbol{b} := \left(\begin{smallmatrix} x_v \\ y_v \end{smallmatrix}\right)$ in Lemma 4.2, we obtain $(a, b) \in [\Theta(\left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right), \left(\begin{smallmatrix} x_v \\ y_v \end{smallmatrix}\right)), 1_A]$, and therefore $\tau \leq [\Theta(\left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right), \left(\begin{smallmatrix} x_v \\ y_v \end{smallmatrix}\right)), 1_A]$. From the definition of the difference graph $\mathbb{H}_\tau$, we now see that $(\left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right), \left(\begin{smallmatrix} x_v \\ y_v \end{smallmatrix}\right)) = (g(u), g(v))$ is an edge of $\mathbb{H}_\tau$. Hence g is a homomorphism from $\mathbb{H}_\mu$ to $\mathbb{H}_\tau$, which implies $\mathbb{H}_\mu \preceq \mathbb{H}_\tau$. Since $\mathbb{H}_\beta \preceq \mathbb{H}_\mu$, we have $\mathbb{H}_\beta \preceq \mathbb{H}_\tau$ and thus $\mathbb{H}_\tau \in \{\mathbb{H}_\alpha \mid \alpha \in \mathrm{Con}(\mathbf{A}), \alpha > 0_A, \mathbb{H}_\beta \preceq \mathbb{H}_\alpha\}$. By the maximality of $\mathbb{H}_\mu$ within this set, we therefore have $\mathbb{H}_\tau \preceq \mathbb{H}_\mu$, which yields a graph homomorphism $h : \mathbb{H}_\tau \to \mathbb{H}_\mu$.

Next, we define a homomorphism $j : \mathbb{G} \to \mathbb{H}_\tau$. For $u \in G$, we define $j(u) := \left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right)$. Using the same reasoning as for g, we show that j is a homomorphism: assume that $(u, v) \in \rho^{\mathbb{G}}$. Since Φ is not satisfied, we have $t(x_u, y_u, \boldsymbol{z}_{(u,v)}) = a$ and $t(x_v, y_v, \boldsymbol{z}_{(u,v)}) = b$. Hence $(a, b) \in [\Theta(\left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right), \left(\begin{smallmatrix} x_v \\ y_v \end{smallmatrix}\right)), 1_A]$, which implies that $\tau \leq [\Theta(\left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right), \left(\begin{smallmatrix} x_v \\ y_v \end{smallmatrix}\right)), 1_A]$. Thus $(j(u), j(v)) = (\left(\begin{smallmatrix} x_u \\ y_u \end{smallmatrix}\right), \left(\begin{smallmatrix} x_v \\ y_v \end{smallmatrix}\right))$ is an edge of $\mathbb{H}_\tau$, and therefore j is a homomorphism from $\mathbb{G}$ to $\mathbb{H}_\tau$.

Now $f := h \circ j$ is the required homomorphism from $\mathbb{G}$ to $\mathbb{H}_\mu$. ◀

We will now prove the main result.

Proof of Theorem 3.1. From Lemma 5.3, we obtain $\beta \in \mathrm{Con}(\mathbf{A})$ such that the difference graph $\mathbb{H}_\beta$ has a triangle. Let $\mathbb{H}_\mu$ be as in the assumptions of Lemma 5.4. Since $\mu > 0_A$, $\mathbb{H}_\mu$ is loopless. Now from $\mathbb{H}_\beta \preceq \mathbb{H}_\mu$, we obtain that $\mathbb{H}_\mu$ contains a triangle. Thus from Corollary 5.1, we obtain that $\mathbb{H}_\mu$-coloring is NP-complete. By Lemma 5.4, the existence of a $\mathbb{H}_\mu$-coloring of a given graph $\mathbb{G}$ can be determined by checking the validity of a quasi-identity Φ that can be computed in time polynomial in the size of $\mathbb{G}$. This implies that $\textsc{QuasiIdVal}(\mathbf{A})$ is coNP-complete. ◀

Proof of Corollary 3.2. Given Theorem 3.1, we only have to verify that $\textsc{QuasiIdVal}(\mathbf{A})$ is in P when $\mathbf{A}$ is an abelian finite Mal'cev algebra of finite type. In Section 2, we observed that $\textsc{QuasiIdVal}(\mathbf{A})$ can be reduced to $\textsc{PolSysSat}(\mathbf{A})$ using a truth table reduction. In fact, a counterexample to the validity of a quasi-identity can be found as the solution of one of $|A| \cdot (|A| - 1)$ many polynomial systems: one passes from a quasi-identity such as (2.1) to the systems given in (2.2). From this reduction, we see that it suffices to show that $\textsc{PolSysSat}(\mathbf{A})$ is in P when $\mathbf{A}$ is an abelian finite Mal'cev algebra of finite type. Since a Mal'cev algebra generates a congruence modular variety, it follows from [15, Corollary 3.14] (which is proved using [7, Theorem 33]) that in this case $\textsc{PolSysSat}(\mathbf{A})$ can be solved in polynomial time. ◀

As special cases, we obtain:

▶ **Corollary 5.5.** *For a finite group* $\mathbf{G}$, $\textsc{QuasiIdVal}(\mathbf{G})$ *is in P if* $\mathbf{G}$ *is abelian, and coNP-complete otherwise. For a finite ring* $\mathbf{R}$ *(not necessarily commutative and not necessarily with unit),* $\textsc{QuasiIdVal}(\mathbf{R})$ *is in P if* $\mathbf{R}$ *is a zero ring, i.e.,* $R \cdot R = 0$*, and coNP-complete otherwise.*

References

1 Erhard Aichinger. The polynomial functions of certain algebras that are simple modulo their center. In *Contributions to general algebra. 17*, pages 9–24. Heyn, Klagenfurt, 2006.

2 Erhard Aichinger, Nebojša Mudrinski, and Jakub Opršal. Complexity of term representations of finitary functions. *Internat. J. Algebra Comput.*, 28(6):1101–1118, 2018. doi:10.1142/S0218196718500480.

3 Andrei A. Bulatov. A dichotomy theorem for nonuniform CSPs. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 319–330, 2017. doi:10.1109/FOCS.2017.37.

4 Stanley Burris and John Lawrence. The equivalence problem for finite rings. *J. Symbolic Comput.*, 15(1):67–71, 1993. doi:10.1142/S021819671100625X.

5 Stanley Burris and Hanamantagouda P. Sankappanavar. *A course in universal algebra.* Springer New York Heidelberg Berlin, 1981.

6 Pete L. Clark. The Combinatorial Nullstellensätze revisited. *Electron. J. Combin.*, 21(4):Paper 4.15, 17, 2014. doi:10.37236/4359.

7 Tomás Feder and Moshe Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: a study through Datalog and group theory. *SIAM J. Comput.*, 28(1):57–104 (electronic), 1999. doi:10.1137/S0097539794266766.

8 Ralph Freese and Ralph N. McKenzie. *Commutator Theory for Congruence Modular varieties*, volume 125 of *London Math. Soc. Lecture Note Ser.* Cambridge University Press, 1987.

9 Mikael Goldmann and Alexander Russell. The complexity of solving equations over finite groups. *Inform. and Comput.*, 178(1):253–262, 2002. doi:10.1006/inco.2002.3173.

10 Pavol Hell and Jaroslav Nešetřil. On the complexity of H-coloring. *J. Combin. Theory Ser. B*, 48(1):92–110, 1990. doi:10.1016/0095-8956(90)90132-J.

11 David Hobby and Ralph N. McKenzie. *The structure of finite algebras*, volume 76 of *Contemporary Mathematics*. American Mathematical Society, Providence, RI, 1988. doi:10.1090/conm/076.

12 Gábor Horváth. The complexity of the equivalence and equation solvability problems over nilpotent rings and groups. *Algebra universalis*, 66(4):391–403, 2011. doi:10.1007/s00012-011-0163-y.

13 Gábor Horváth and Csaba Szabó. The extended equivalence and equation solvability problems for groups. *Discrete Mathematics & Theoretical Computer Science*, Vol. 13 no. 4, 2011. doi:10.46298/dmtcs.536.

14 Paweł M. Idziak, Piotr Kawałek, and Jacek Krzaczkowski. Intermediate problems in modular circuits satisfiability. In *Proceedings of the 35th Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '20, pages 578–590. Association for Computing Machinery, 2020. doi:10.1145/3373718.3394780.

15 Benoit Larose and László Zádori. Taylor terms, constraint satisfaction and the complexity of polynomial equations over finite algebras. *Internat. J. Algebra Comput.*, 16(3):563–581, 2006. doi:10.1142/S0218196706003116.

16 Anatoly I. Mal'cev. On the general theory of algebraic systems. *Mat. Sb. N.S.*, 35(77):3–20, 1954.

17 Ralph N. McKenzie, George F. McNulty, and Walter F. Taylor. *Algebras, lattices, varieties, Volume I.* Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, California, 1987.

18 Michael Sipser. *Introduction to the Theory of Computation.* Cengage Learning, Boston, MA, USA, 3rd edition, 2012. URL: https://books.google.at/books?id=1aMKAAAAQBAJ.

19 Guy Terjanian. Sur les corps finis. *C. R. Acad. Sci. Paris Sér. A-B*, 262:A167–A169, 1966.

20 Mikhail V. Volkov. Checking quasi-identities in a finite semigroup may be computationally hard. *Studia Logica*, 78(1-2):349–356, 2004. doi:10.1007/s11225-005-0356-5.

21 Armin Weiß. Hardness of equations over finite solvable groups under the exponential time hypothesis. In *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 102:1–102:19, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2020.102.

22 László Zádori. Solvability of systems of polynomial equations over finite algebras. *Internat. J. Algebra Comput.*, 17(4):821–835, 2007. doi:10.1142/S0218196707003809.

23 Dmitriy Zhuk. A proof of the CSP dichotomy conjecture. *Journal of the ACM*, 67(5):1–78, 2020. doi:10.1145/3402029.