JOHANNES KEPLER
UNIVERSITY LINZ

Author
**Simon Grünbacher**

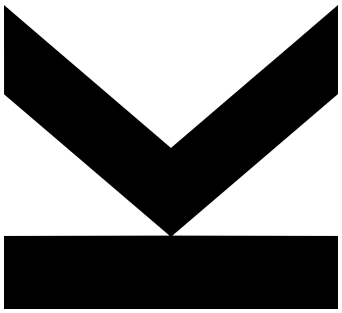
Submission
**Institute for Integrated Circuits**

First Supervisor
Assoz. Univ.-Prof. DI Dr.
**Richard Küng**

Second Supervisor
Assoz. Univ.-Prof. DI Dr.
**Erhard Aichinger**

April 2024

# The complexity of quasi-identities over finite algebras with a Mal'cev term



**JOHANNES KEPLER**
**UNIVERSITY LINZ**
Altenbergerstraße 69
4040 Linz, Austria
www.jku.at
DVR 0093696

# Kurzfassung

Eine Quasi-Identität ist eine Formel der Form

$$(\bigwedge_{i=1}^{k} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})) \Rightarrow u(\boldsymbol{x}) = v(\boldsymbol{x}),$$

wobei $(s_i)_{i=1}^k, (t_i)_{i=1}^k, u, v$ Terme über einer Algebra $\mathbf{A}$ in den Variablen $x_1, \ldots, x_n$ sind. Das Problem QuasiIdVal($\boldsymbol{A}$) stellt die Frage, ob eine gegebene Quasi-Identität für alle Werte $\boldsymbol{x} \in A^n$ gilt. Ziel dieser Arbeit ist es, die Komplexität von QuasiIdVal($\boldsymbol{A}$) für verschiedene endliche Algebren $\boldsymbol{A}$ festzustellen. Insbesondere wird ein Dichotomietheorem für endliche Ringe und Gruppen bewiesen. Dieses Theorem besagt, dass QuasiIdVal unter bestimmten Bedingungen in P ist und ansonsten coNP-vollständig.

# Abstract

A quasi-identity is a formula of the form

$$(\bigwedge_{i=1}^{k} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})) \Rightarrow u(\boldsymbol{x}) = v(\boldsymbol{x}),$$

where $(s_i)_{i=1}^k, (t_i)_{i=1}^k, u, v$ are expressions built from the variables $x_1, \ldots, x_n$ and the fundamental operations of an algebra $\mathbf{A}$. The problem QuasiIdVal($\mathbf{A}$) asks whether a given quasi-identity holds for all values $\boldsymbol{x} \in A^n$. In this thesis, we seek to determine the computational complexity of QuasiIdVal($\boldsymbol{A}$) for various finite algebras $\boldsymbol{A}$. In particular, we find a P/coNPC dichotomy for finite rings and groups.

# Contents

# 1 Introduction

A *quasi-identity* is an implication between a system of equations, the *precondition*, and a single equation, the *conclusion.* We want to find out how hard it is to check whether such a quasi-identity holds over some fixed algebra. An *algebra* $\mathbf{A}$ is a pair $(A, F)$, where $A$ is a nonempty set and $F$ is a set of *fundamental operations* on $A$ of finite arity. For example, the ring $(\mathbb{Z}, +, -, 0, \cdot)$ is an algebra with $A = \mathbb{Z}$ and fundamental operations $F = \{+, -, 0, \cdot\}$. For an algebra $\mathbf{A}$, a *term* is an expression involving fundamental operations and variables. For example, $x + (y \cdot z)$ is a term in the algebra $(\mathbb{Z}, +, -, 0, \cdot)$. On the other hand, $x + 1$ is not a term in $(\mathbb{Z}, +, -, 0, \cdot)$, because 1 is neither a fundamental operation nor a variable. We call such an expression involving constants, variables and fundamental operations a *polynomial*. We are interested in the computational complexity of the following problem:

**Definition 1.1.** Let $\mathbf{A}$ be an algebra. The problem QuasiIdVal($\mathbf{A}$) is the following decision problem:
**Given:** A quasi-identity $\Phi$ of the form

$$(\bigwedge_{i=1}^{k} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})) \Rightarrow u(\boldsymbol{x}) = v(\boldsymbol{x}),$$

where $s_1, \ldots, s_k, t_1, \ldots, t_k, u, v$ are terms in $x_1, \ldots, x_n$ over $\mathbf{A}$.
**Asked:** Does $\Phi$ hold for all $\boldsymbol{x} \in A^n$?

Note that we only allow terms to appear in our quasi-identities. One could also consider polynomial quasi-identities. In chapter 2, we will show how the complexity of this more general problem follows from known results.

For example, the formula $\Phi_1$ defined by

$$\forall x, y, z : (x + y = z \wedge x + z = y) \Rightarrow x = z + z$$

is a quasi-identity with precondition $x+y=z \wedge x+z=y$ and conclusion $x=z+z$. Our goal is to find out whether $\phi_1$ holds in the algebra $(\mathbb{Z}_3,+)$. This can be done systematically using linear algebra, which shows that quasi-identities in $(\mathbb{Z}_3,+)$ can be verified in polynomial time. As a second example, consider the formula $\Phi_2$ over $(\mathbb{Z}_3,+,\cdot)$ defined by

$$\forall x,y,z:(x\cdot y=z\wedge x+z=y)\Rightarrow x=z+z.$$

We cannot use linear algebra anymore to find out whether $\Phi_2$ holds in $(\mathbb{Z}_3,+,\cdot)$. However, the problem QuasiIdVal$(\mathbb{Z}_3,+,\cdot)$ of verifying quasi-identities over $(\mathbb{Z}_3,+,\cdot)$ is still in the problem class coNP, since we can disprove a quasi-identity by providing a counterexample. In fact, QuasiIdVal$(\mathbb{Z}_3,+,-,0,\cdot)$ is coNP-complete: We can provide a simple reduction from 3-Colorability by noting that the graph $(V,E)$ is 3-colorable if and only if the quasi-identity

$$0=0\Rightarrow\prod_{(u,v)\in E}(x_u-x_v)=0$$

does not hold.

In this thesis, we try to find out what distinguishes these two cases. In particular, we want to know the complexity for finite rings and groups. In *universal algebra,* rings and groups belong to the larger class of *Mal'cev algebras.* Our main result is a P/coNPC dichotomy theorem for Mal'cev algebras, which was originally published in [1]. In an earlier version of the result, this dichotomy was only proven for subdirectly irreducible expanded groups. The author then proved a full dichotomy in the case of finite groups. Erhard Aichinger then proved Lemma 5.10, which shows that a certain property of groups also holds for finite Mal'cev algebras. This allowed Aichinger to generalize the proof from groups to Mal'cev algebras, yielding Theorem 6.4. To the author's knowledge, there is only one other publication which explicitly considers this question: In [13], M. Volkov constructs a 10-element semigroup $\mathbf{Q}$ such that QuasiIdVal$(\mathbf{Q})$ is coNP-complete, but checking the validity of a single term equation (TermEqv$(\mathbf{Q})$) is solvable in polynomial time. We will see in later chapters that this is also the case for all nonabelian nilpotent groups.

# 2 Relationship to systems satisfiability

In this chapter, we relate the complexity of QuasiIdVal to the complexity of the polynomial satisfiability problem. Note that a polynomial over an algebra $\mathbf{A} = (A, F)$ is just a term over the extended algebra $\mathbf{A}^+ = (A, F \cup C_A)$, where $C_A$ contains a constant operation $c_a : A^0 \to A$ of arity 0 for each $a \in A$. In other words, a polynomial is just an expression involving function symbols from $F$, variables and constants from $A$.

**Definition 2.1.** Let $\mathbf{A} = (A, F)$ be a finite algebra. The *polynomial satisfiability problem* over $\mathbf{A}$, denoted by PolSysSat($\mathbf{A}$), is the following decision problem:
**Given:** Polynomials $p_1, q_1, \ldots, p_k, q_k$ over $\mathbf{A}$ in the variables $x_1, \ldots, x_n$;
**Asked:** Does the formula

$$\exists \boldsymbol{x} \in A^n : \bigwedge_{i=1}^{k} p_i(\boldsymbol{x}) = q_i(\boldsymbol{x})$$

hold?

Clearly PolSysSat($\mathbf{A}$) $\in$ NP for all finite $\mathbf{A}$ of finite type, because we can prove satisfiability by providing a solution. On the other hand, the problem is not NP-complete for every algebra. For example, a polynomial system over the abelian group $(\mathbb{Z}_3, +)$ is just an affine linear system. Therefore, PolSysSat$(\mathbb{Z}_3, +, -, 0) \in$ P. On the other hand, one can see that PolSysSat$(\mathbb{Z}_3, +, -, 0, \cdot)$ is NP-complete by observing that the graph $G = (V, E)$ is 3-colorable if and only if the system $\bigwedge_{(u,v)\in E}(x_u - x_v) \cdot (x_u - x_v) = 1$ has a solution.

We can solve QuasiIdVal($\mathbf{A}$) by solving a finite number of instances of PolSysSat($\mathbf{A}$) :

**Lemma 2.2.** *Let $\mathbf{A} = (A, F)$ be a finite algebra. Then there is a truth-table reduction from QuasiIdVal($\mathbf{A}$) to the complement of PolSysSat($\mathbf{A}$).*

*Proof.* Let $\Phi$ be an instance of QuasiIdVal($\mathbf{A}$), defined as

$$(\bigwedge_{i=1}^{k} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})) \Rightarrow u(\boldsymbol{x}) = v(\boldsymbol{x}).$$

Then the quasi-identity $\Phi$ is valid if and only if the polynomial system

$$(\bigwedge_{i=1}^{k} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})) \wedge u(\boldsymbol{x}) = a \wedge v(\boldsymbol{x}) = b$$

is unsatisfiable for all constants $a, b \in A$ with $a \neq b$. □

It is not clear whether such a direct reduction is possible in the other direction. This is because PolSysSat allows expressions with constants, while QuasiIdVal does not. Let PolQuasiIdVal($\mathbf{A}$) be the problem of deciding the validity of *polynomial* quasi-identities. We then get the following:

**Lemma 2.3.** *Let* $\mathbf{A}$ *be a finite algebra with* $|A| > 1$. *Then* PolSysSat($\mathbf{A}$) *can be reduced to the complement of* PolQuasiIdVal($\mathbf{A}$).

*Proof.* The polynomial system

$$\bigwedge_{i=1}^{k} p_i(\boldsymbol{x}) = q_i(\boldsymbol{x})$$

has a solution if and only if the polynomial quasi-identity

$$(\bigwedge_{i=1}^{k} p_i(\boldsymbol{x}) = q_i(\boldsymbol{x})) \Rightarrow y = z$$

is not valid. □

Likewise, the problem of solving a system of *term* equations (TermSysSat($\mathbf{A}$)) can be reduced to the complement of QuasiIdVal($\mathbf{A}$). However, every system of term equations over a group has the solution $\boldsymbol{x} = (1, \ldots, 1)$. Similarly, every system of term equations over a ring $\mathbf{R} = (R, +, -, \cdot)$ has the solution $(0, \ldots, 0)$. Therefore, this reduction does not yield coNP-completeness results for rings and groups. We mention in passing that for finite algebras, the complexity of TermSysSat was recently classified in [11].

We briefly comment on the tools used in [10] to determine the complexity of systems satisfiability. The key observation is that PolSysSat($\mathbf{A}$) can be reduced to a constraint satisfaction problem. For example, the system $x + (x \cdot z) = 0 \wedge x + y = z$ is satisfiable in $(\mathbb{Z}_3, +, \cdot, 0)$ if and only if the constraint satisfaction problem

$$\rho.(x, z, t_1) \wedge \rho_+(x, t_1, t_2) \wedge \rho_0(t_2) \wedge \rho_+(x, y, z)$$

is satisfiable, where $\rho_f(\boldsymbol{x}, y) : \Longleftrightarrow \ f(\boldsymbol{x}) = y$ denotes the relation defined by the graph of the fundamental operation. For a finite set of relations $\mathcal{R} \subset \bigcup_{n\in\mathbb{N}} \mathcal{P}(D^n)$, over a finite *domain* $D$, the problem CSP($\mathcal{R}$) is the question whether there exists an assignment of values in $D$ to the variables that satisfy specific relations taken from the set $\mathcal{R}$. For example, $k-$Colorability can be viewed as CSP($\mathcal{R}$), where $\mathcal{R} = \{\{(i, j) \mid i, j \in \{1, \ldots, k\}, i \neq j\}\}$ and $D = \{1, \ldots, k\}$. Our example above now provides an intuitive argument why the problems PolSysSat$(\mathbb{Z}_3, +, \cdot, 0)$ and CSP$(\{\rho_+, \rho., \rho_0\})$ are equivalent. After decades-long research, Zhuk [14] and Bulatov [2] independently proved the CSP-Dichotomy Theorem. This theorem states that for every finite set of relations $\mathcal{R}$ over a finite domain $D$, the problem CSP($\mathcal{R}$) is solvable in polynomial time if a certain condition holds for $\mathcal{R}$, and NP-complete otherwise. This completely determines the complexity of PolSysSat($\mathbf{A}$) and therefore PolQuasiIdVal($\mathbf{A}$). However, it is not obvious how to translate what this condition means for groups and rings. In [10], the authors use the part of the dichotomy theorem that was known at the time to prove the following:

**Theorem 2.4.** *Let $\mathbf{G} = (G, \cdot, (\cdot)^{-1}, 1)$ be a finite group and let $\mathbf{R} = (R, +, -, 0, \cdot)$ be a finite ring.*

- *If $a \cdot b = b \cdot a$ for all $a, b \in G$, then PolSysSat($\mathbf{G}$) $\in P$; otherwise PolSysSat($\mathbf{G}$) is NP-complete.*
- *If $a \cdot b = 0$ for all $a, b \in R$, then PolSysSat($\mathbf{R}$) $\in P$; otherwise PolSysSat($\mathbf{R}$) is NP-complete.*

Together with Lemma 2.2 and Lemma 2.3 we get

**Corollary 2.5.** *Let $\mathbf{G} = (G, \cdot, (\cdot)^{-1}, 1)$ be a finite group and let $\mathbf{R} = (R, +, -, \cdot)$ be a finite ring.*

- *If* $a \cdot b = b \cdot a$ *for all* $a, b \in G$*, then* PolQuasiIdVal($\mathbf{G}$) *and* QuasiIdVal($\mathbf{G}$) *are in P; otherwise* PolQuasiIdVal($\mathbf{G}$) *is NP-complete.*
- *If* $a \cdot b = 0$ *for all* $a, b \in R$*, then* PolQuasiIdVal($\mathbf{R}$) *and* QuasiIdVal($\mathbf{R}$) *are in P; otherwise* PolQuasiIdVal($\mathbf{R}$) *is NP-complete.*

This determines the complexity of PolQuasiIdVal, but leaves the complexity of QuasiIdVal open for nonabelian groups and nonzero rings.

# 3 Relationship to the equivalence problem

In this chapter, we cover existing results for the complexity of the term equivalence problem. We relate the results to the complexity of the quasi-identity validity problem.

**Definition 3.1.** Let $\mathbf{A}$ be an algebra. The problem TermEqv($\mathbf{A}$) is the following problem:
**Given:** Terms $s, t$ in $n$ variables over the language of $\mathbf{A}$
**Asked:** Does $s(\boldsymbol{x}) = t(\boldsymbol{x})$ hold for all $\boldsymbol{x} \in A^n$?

If $\mathbf{A}$ is finite, then TermEqv($\mathbf{A}$) is in coNP: If there is an $\boldsymbol{x} \in A^n$ with $s(\boldsymbol{x}) \neq t(\boldsymbol{x})$, then we can verify this counterexample in polynomial time by evaluating $s$ and $t$ at $\boldsymbol{x}$. The term equivalence $s(\boldsymbol{x}) = t(\boldsymbol{x})$ is valid if and only if the quasi-identity $x_1 = x_1 \Rightarrow s(\boldsymbol{x}) = t(\boldsymbol{x})$ is valid. Thus TermEqv can be reduced to QuasiIdVal. Since TermEqv is known to be hard in many cases, this allows us to prove hardness of quasi-identity validity for some groups and rings. To get a feeling for the complexity of TermEqv, we study two examples:

**Example 1.** TermEqv($\mathbb{Z}_3, +$): Let $s$ and $t$ be two terms over $(\mathbb{Z}_3, +)$. The functions induced by $s$ and $t$ are linear functions from $\mathbb{Z}_3^n$ to $\mathbb{Z}_3$. Thus, they are equal if and only if they are equal for all elements of a basis of $\mathbb{Z}_3^n$. This can be checked in polynomial time by evaluating $s$ and $t$ on the standard basis $e_1, \ldots, e_n \in \mathbb{Z}_3^n$. Therefore TermEqv($\mathbb{Z}_3, +$) $\in$ P.

**Example 2.** TermEqv($\mathbb{Z}_3, +, -, 0, \cdot$): This problem is considerably harder. We can prove coNP-completeness by providing a reduction from the complement of 3-Colorability. Given a graph $G = (V, E)$, consider the equation $\Phi$:

$$0 \approx \prod_{(u,v) \in E} (x_u - x_v)$$

If $G$ is not 3−colorable, then every assignment $(x_v)_{v\in V} \in \mathbb{Z}_3$ has an edge $(u,v) \in E$ with $x_u = x_v$, and hence $\prod_{(u,v)\in E}(x_u - x_v) = 0$. If $\Phi$ is not valid, then there exists a counterexample $(x_v)_{v\in V}$ with $\prod_{(u,v)\in E}(x_u - x_v) \neq 0$ and thus $x_u \neq x_v$ for all $(u,v) \in E$. Thus this counterexample provides a valid coloring of $G$.

For finite rings, the complexity of the term equivalence problem was fully classified in [3].

**Theorem 3.2** ([3]). *Let $\mathbf{R} = (R, +, -, 0, \cdot)$ be a finite ring.*

- *If there is a $k \in \mathbb{N}$ such that $x_1 \dots x_k = 0$ for all $x_1, \dots, x_k \in R$, then $\textsc{TermEqv}(\mathbf{R}) \in P$.*
- *Otherwise $\textsc{TermEqv}(\mathbf{R})$ is coNP-complete.*

We call a finite ring *nilpotent* if there is a $k \in \mathbb{N}$ such that $x_1 \dots x_k = 0$ for all $x_1, \dots, x_k \in R$. For example, the ring $R_k := \{(a_{ij})_{i,j} \in \mathbb{Z}^{(k+1)\times(k+1)} \mid \forall i \leq j : a_{ij} = 0\}$ is nilpotent for all $k \in \mathbb{N}$. On the other hand, every nontrivial ring that contains a unit element 1 is non-nilpotent.

We now shift our attention to term equivalence for groups. We start by introducing some basic notation: For a group $\mathbf{G} = (G, \cdot, (\cdot)^{-1}, 1)$, we call a subgroup $N \subseteq G$ *normal* if $g^{-1}ng \in N$ for all $n \in N$ and $g \in G$. Given two normal subgroups $M$ and $N$, their *commutator* $[M, N]$ is defined as the normal subgroup generated by the set $\{x^{-1}y^{-1}xy \mid x \in N, y \in M\} \subseteq G$. We can use the commutator to define a sequence $G^{(k)}$ of normal subgroups defined by $G^{(0)} := G$ and $G^{(k+1)} = [G^{(k)}, G]$. This sequence is called the *lower central series*. We now call $\mathbf{G}$ *nilpotent* if there exists a $k \in \mathbb{N}$ such that $G^{(k)} = \{1\}$. This is analogous to the definition of rings, because a ring is nilpotent if the series $R, R \cdot R, \dots$ terminates with $\{0\}$. In [6], the authors show that $\textsc{TermEqv}(\mathbf{G}) \in$ P for finite nilpotent groups $\mathbf{G}$. In the case of groups, non-nilpotence does not imply hardness of TermEqv as it does in the case of rings. For example, consider the alternating group $A_4$, which consists of even permutations on the set $\{1,2,3,4\}$. $A_4$ is not nilpotent, as $[A_4, A_4] = V =: \{(), (1\ 2), (3\ 4), (1\ 2)(3\ 4)\}$ and $[A_4, V] = V$. In [9], the authors find that $\textsc{TermEqv}(A_4) \in$ P. Somewhat surprisingly, they also prove that $\textsc{TermEqv}((A_4, \cdot, (\cdot)^{-1}, 1, [\cdot,\cdot]))$ is coNP-complete, where $[x, y] := x^{-1}y^{-1}xy$. This means that adding the commutator term as a fundamental operation changes the complexity. The reason is that the property $1 \in \{x, y\} \Rightarrow [x, y] = 1$ of the commutator can be used to encode the conjunction of constraints, allowing for a reduction

from colorability. This reduction cannot be carried out when the commutator is not a fundamental operation, because expanding a nested commutator increases the expression size exponentially. For example, $[[x,y],z] = [x,y]^{-1}z^{-1}[x,y]z = (x^{-1}y^{-1}xy)^{-1}z^{-1}x^{-1}y^{-1}xyz$. In the case of groups, one can more generally prove that for every non-nilpotent group $\mathbf{G}$, there is a term function $t$ such that $\textsc{TermEqv}(G, \cdot, (\cdot)^{-1}, 1, t)$ is coNP-complete (see [8]). This shows that the complexity of $\textsc{TermEqv}(\mathbf{A})$ depends not only on the term functions of the algebra $\mathbf{A}$, but also on the set of fundamental operations. This does not happen for quasi-identity validity. For example, given the quasi-identity $\Phi_1$

$$x_1 = x_2 \Rightarrow [[x_1, x_2], x_3] = 1$$

over $(A_4, \cdot, (\cdot)^{-1}, 1, [\cdot,\cdot])$, we can form an equivalent quasi-identity $\Phi_2$

$$x_1 = x_2 \wedge y_1 = x_1^{-1}x_2^{-1}x_1x_2 \wedge y_2 = y_1^{-1}x_3^{-1}y_1x_3 \Rightarrow y_2 = 1$$

over $(A_4, \cdot, (\cdot)^{-1}, 1)$. As the following lemma shows, we can always introduce local variables in the precondition to avoid expanding deeply nested terms:

**Lemma 3.3.** *Let $\mathbf{A}_1 = (A, F_1)$ and $\mathbf{A}_2 = (A, F_2)$ be two algebras over the same universe such that $F_1$ and $F_2$ are finite. If $\mathbf{A}_1$ and $\mathbf{A}_2$ have the same term functions, then $\textsc{QuasiIdVal}(\mathbf{A}_1)$ is equivalent to $\textsc{QuasiIdVal}(\mathbf{A}_2)$.*

*Proof.* By symmetry, it suffices to show that $\textsc{QuasiIdVal}(\mathbf{A}_1)$ can be reduced to $\textsc{QuasiIdVal}(\mathbf{A}_2)$. For every fundamental operation $f \in F_1$ of $\mathbf{A}_1$ choose a term $t_f$ over $\mathbf{A}_2$ which induces the function $f$. Now let $\Phi_1$ be the quasi-identity

$$(\bigwedge_{i=1}^{k} s_i(\boldsymbol{x}) = t_i(\boldsymbol{x})) \Rightarrow u(\boldsymbol{x}) = v(\boldsymbol{x})$$

over $\mathbf{A}_1$. For a term $w$, let $\sigma(w)$ be the set of subexpressions appearing in $\Phi$. We define $\sigma$ recursively as $\sigma(x) = \{x\}$ for a variable $x$ and $\sigma(f(w_1, \ldots, w_n)) = \{f(w_1, \ldots, w_n\} \cup \bigcup_{i=1}^{n} \sigma(w_i)$. Let $S = \bigcup\{\sigma(w) \mid w \in \{s_1, t_1, \ldots, s_k, t_k, u, v\}\}$ be the set of subexpressions appearing in $\Phi$, and let $X \subseteq S$ be the set of variables. Clearly $|S|$ is bounded by the number of function symbols and variables appearing in $\Phi$. Now let $(y_s)_{s \in S}$ be a sequence of

variables distinct from those appearing in $\Phi$. Consider the following quasi-identity $\Phi_2$ over $\mathbf{A}_2$ :

$$\Big(\bigwedge_{f(g_1,\ldots,g_l)\in S\setminus X} y_{f(g_1,\ldots,g_l)} = t_f(y_{g_1},\ldots,y_{g_l})\Big) \wedge \Big(\bigwedge_{x\in X} y_x = x\Big) \Rightarrow y_u = y_v$$

The size of $\Phi_2$ is now bounded by the size of $S$ times the maximal size of $t_f$. We show that $\Phi_2$ is valid if and only if $\Phi_1$ is valid: Let $(y_s)_{s\in S}$ be a counterexample to the validity of $\Phi_2$. Since the precondition holds, it follows that for $i \in \underline{k}$, we have $s_i((y_x)_{x\in X}) = y_{s_i} = y_{t_i} = t_i((y_x)_{x\in X})$. Since the conclusion of $\Phi_2$ fails, we have $u((y_x)_{x\in X}) = y_u \neq y_v = v((y_x)_{x\in X})$. Thus $(y_x)_{x\in X}$ is a counterexample to the validity of $\Phi_1$. By the same argument, we can turn a counterexample $\boldsymbol{x}$ to the validity of $\Phi_1$ into a counterexample to the validity of $\Phi_2$ by choosing $y_s := s(\boldsymbol{x})$ for $s \in S$. □

The results from our reductions from the term equivalence problem and to the system satisfiability problem are summarized in Table 3.1.

**Table 3.1:** Complexity of the studied problems as known before [1].

| $\mathbf{A}$ | TermEqv($\mathbf{A}$) | QuasiIdVal($\mathbf{A}$) | PolSysSat($\mathbf{A}$) |
|---|---|---|---|
| module/abelian group/zero ring | P | P | P ([6, 10]) |
| nonabelian nilpotent group | P ([6]) | open | NPC ([6]) |
| non-nilpotent solvable group | partially open | coNPC ([8]) | NPC ([6]) |
| non-solvable group | coNPC ([6]) | coNPC | NPC ([6]) |
| non-zero nilpotent ring | P ([3]) | open | NPC ([10]) |
| non-nilpotent ring | coNPC ([3]) | coNPC | NPC ([10]) |

# 4 A hardness proof for commutative rings

In this chapter, we prove coNP-completeness of QuasiIdVal($\mathbf{R}$) in the case where $\mathbf{R} = (R, +, -, 0, \cdot)$ is a commutative ring which has $a, b \in R$ with $ab \neq 0$. For such a ring, we call a subset $I \subseteq R$ an *ideal* if for all $a, b \in I, r \in R$ we have $\{0, a+b, a-b, r \cdot a\} \subseteq I$. Note that we do not require $R$ to have a multiplicative unit. For example, $3\mathbb{Z}_{27} = (\{0, 3, 6, \ldots, 24\}, +, -, 0, \cdot)$ is a ring and $I = \{0, 9, 18\}$ is an ideal of $3\mathbb{Z}_{27}$. The intersection of a collection of ideals is once again an ideal. This allows us to define *the ideal generated by* $A \subseteq R$ as

$$\langle A \rangle := \bigcap \{I \subseteq R \mid A \subseteq I, I \text{ is an ideal of } \mathbf{R}\}$$

It is easy to see that $\langle A \rangle = \{a_1 + \cdots + a_n + r_1 a_{n+1} + \cdots + r_m a_{n+m} \mid n, m \in \mathbb{N}_0, a_1, \ldots, a_{n+m} \in A, r_1, \ldots, r_m \in R\}$, since the expression on the right-hand side is a part of the intersection and must be contained in every ideal that contains $A$. Given an ideal $I$ of $\mathbf{R}$ and $r \in \mathbf{R}$, we define the class of $r$ modulo $I$ as

$$r/I := \{r + i \mid i \in I\}.$$

For ideals $I, J \subseteq R$, their *product* is defined as

$$I \cdot J = \{i_1 \cdot j_1 + \cdots + i_k \cdot j_k \mid k \in \mathbb{N}, i_1, \ldots, i_k \in I, j_1, \ldots, j_k \in J\}.$$

It is straightforward to verify that $I \cdot J$ is again an ideal of $\mathbf{R}$.

**Lemma 4.1.** *Let* $\mathbf{R}$ *be a finite ring, let* $a, b \in R$ *and let* $I = \langle a, b \rangle$ *be an ideal of* $\mathbf{R}$*. Then*

$$R \cdot I = \{r \cdot a + s \cdot b \mid r, s \in R\}$$

*Proof.* ($\supseteq$) : Clearly $a, b \in I$, therefore $r \cdot a + s \cdot b \in R \cdot I$ for all $r, s \in R$. ($\subseteq$) : Let $x \in R \cdot I$. We first consider the case where $x = yz$ for some $y \in R$ and $z \in I$. By the characterization

of $I$ discussed above, we are in one of the following cases: (1): $z = a + ra + sb$ for some $r, s \in R$. In this case $x = y(a + ra + sb) = (y + yr)a + ysb$. (2): $z = b + ra + sb$ for some $r, s \in R$. In this case $x = y(b + ra + sb) = yra + (y + ys)b$. (3): $z = a + b + ra + sb$ for some $r, s \in R$. In this case $x = y(a + ra + sb) = (y + yr)a + (y + ys)b$. (4): $z = ra + sb$ for some $r, s \in R$. In this case $x = y(ra + sb) = yra + ysb$.

In all those cases $x$ is of the required form. Now consider the case where $x = y_1z_1 + \cdots + y_kz_k$ for some $y_1, \ldots, y_k \in R, z_1, \ldots, z_k \in I$. Using the reasoning above, we choose $r_1, \ldots, r_k, s_1, \ldots, s_k \in R$ such that $y_iz_i = r_ia + s_ib$. Then $x = (r_1 + \cdots + r_k)a + (s_1 + \cdots + s_k)b$, as required.

□

For an ideal $I$, we define its *annihilator*

$$A(I) := \{r \in \mathbb{R} \mid \forall i \in I : r \cdot i = 0\},$$

which is again an ideal of $\mathbf{R}$. Note that all products in $\mathbf{R}$ vanish precisely if $A(R) = R$.

The first instinct for proving coNP-completeness of QuasiIdVal($\mathbf{R}$) for nonzero rings might be a variation of the construction given in the introduction. In other words, we might reduce from colorability and use a quasi-identity like $0 = 0 \Rightarrow \prod_{(u,v)\in E}(x_u - v_v) = 0$ to express non-colorability. The problem with this ansatz is that long products may vanish. For example, in the ring $3\mathbb{Z}_{27}$, all products of length 3 are equal to zero. The above quasi-identity would therefore be valid for all graphs $(V, E)$ with more than three edges. Considering the case of $3\mathbb{Z}_{27}$ more closely, we might come up with the following quasi-identity:

$$\Big(\bigwedge_{(u,v)\in E} y_{u,v}(x_u - x_v) = z\Big) \Rightarrow z = 0$$

Indeed, this quasi-identity is valid if and only if the graph $(V, E)$ is 3-colorable. Clearly, a counterexample to the validity of this quasi-identity yields a proper graph coloring. Also, whenever we have a proper graph coloring $c : V \to \{0, 3, 6\}$, choosing $x_v := c(v), y_{u,v} := (c(u) - c(v))$ and $z = 9$ yields a counterexample to the quasi-identity. There are two problems that block the generalization of this proof to arbitrary commutative rings:

- This proof requires the property that if $x_u - x_v \notin A(R)$, then $R \cdot \langle x_u - x_v \rangle$ contains the unique nonzero minimal ideal $\{0, 9, 18\}$. This is a property that is quite specific to $3\mathbb{Z}_{27}$.
- We clearly cannot use this reduction from Colorability if the ring is $2\mathbb{Z}_8$, because 2-Colorability is too easy.

The second problem can be avoided by considering pairs of variables. To solve the first problem, we need a way to control the value of $z$ in counterexamples - so far we only know that $z \neq 0$. We do this by introducing suitable equations in the precondition. This will ensure that the constraint satisfaction problem defined by the relation $\rho(x_u, x_v) :\iff (R \cdot \langle x_u - x_v \rangle \ni z)$ is hard for any $z$ that might appear in a counterexample. It is now convenient to pick the constraint satisfaction problem to reduce from depending on the ring $\mathbf{R}$.

We are therefore interested in the special case of the CSP-Dichotomy Theorem where the only constraint is a symmetric binary relation $\rho \subseteq D \times D$ over a finite domain $D$. In this case $\mathrm{CSP}(\rho)$ is equivalent to the question whether there is a graph homomorphism $h : G \to H$ from a given graph $G = (V, E)$ into the fixed undirected graph $H = (D, \rho)$. In this case we write $G \preceq H$. If $D = \{1, \ldots, k\}$ and $\rho = \{(i, j) \in D \times D \mid i \neq j\}$, then $\mathrm{CSP}(\rho)$ is equivalent to $k-$Colorability. The complexity of $\mathrm{CSP}(\rho)$ was classified by Hell and Nešetřil:

**Theorem 4.2** ([7], Theorem 1)**.** *Let $D$ be a finite set and let $\rho \subseteq D \times D$ be symmetric such that $(a, a) \notin \rho$ for all $a \in D$. Then the problem CSP($\rho$) is NP-complete if the graph $(D, \rho)$ is non-bipartite.*

We are now ready to prove the main theorem of this chapter:

**Theorem 4.3.** *Let $\mathbf{R} = (R, +, -, 0, \cdot)$ be a finite commutative ring. If there exist $a, b \in R$ with $ab \neq 0$, then QuasiIdVal($\mathbf{R}$) is coNP−complete.*

*Proof.* Now choose $a, b \in R$ with $ab \neq 0$. Then $a \notin A(R)$ and therefore $a/A(R) \neq 0/A(R)$.

For every ideal $I \neq 0$ of $\mathbf{R}$, let $\rho_I \subset W \times W$ be the binary relation over $W := (\mathbf{R}/A(R))^2$ defined as

$$\rho_I := \{(\boldsymbol{x}, \boldsymbol{y}) \mid R \cdot \langle x_1 - y_1, x_2 - y_2 \rangle \supseteq I\}.$$

It is easy to see that the graph $H_I := (W, \rho_I)$ is symmetric and contains no loops. For $J := R \cdot \langle a \rangle$, we see that $\rho_J$ contains the edges $(\binom{a}{0}, \binom{0}{0}), (\binom{a}{0}, \binom{0}{a})$ and $(\binom{0}{a}, \binom{0}{0})$. Let

$$S := \{H_I \mid I \text{ is an ideal of } \mathbf{R}, I \neq 0, H_J \preceq H_I\}.$$

Now $S$ is a finite nonempty set and $\preceq$ defines a quasi-order on $S$. Therefore $S$ has an element $H_M$ which is maximal in the sense that $H_M \preceq H_N$ implies $H_N \preceq H_M$ for all $H_N \in S$. The graph $H_M$ is now symmetric and loopless. Because $H_M$ is loopless and $H_J$ contains a triangle, we see that $H_M$ must also contain a triangle. Therefore $\mathrm{CSP}(\rho_M)$ is NP$-$complete. We perform a reduction from (the complement of) $\mathrm{CSP}(\rho_M)$ to QuasiIdVal($\mathbf{R}$). Let $G = (V, E)$ be an undirected graph. We assume without loss that $V$ and $W$ are disjoint. Now consider the quasi-identity $\Phi$ defined by

$$\Big(\bigwedge_{(u,v)\in\rho_M\cup E} s_{(u,v)}(x_u - x_v) + t_{(u,v)}(y_u - y_v) = z\Big) \Rightarrow z = 0.$$

We prove that $\Phi$ is valid if and only if $G \not\preceq H_M$ : ($\Rightarrow$) : Assume we have a graph homomorphism $h : V \to W$ from $G$ to $H_M$. We show that there is a counterexample to $\Phi$. Choose $z$ to be a nonzero element of $M \subseteq R$. For $u \in V$ choose $(x_u, y_u) =: h(u)$. For $u \in W$ choose $(x_u, y_u) =: u$. We are left to define $s_{(u,v)}$ and $t_{(u,v)}$ for $(u, v) \in \rho_M \cup E$. If $(u, v) \in \rho_M$, then $R \cdot \langle x_u - x_v, y_u - y_v \rangle \supseteq M \ni z$. Using Lemma 4.1, we choose $s_{(u,v)}, t_{(u,v)} \in R$ with $s_{(u,v)}(x_u - x_v) + t_{(u,v)}(y_u - y_v) = z$. If $(u, v) \in E$, then $(h(u), h(v)) \in W$ and thus $R \cdot \langle x_u - x_v, y_u - y_v \rangle \supseteq M \ni z$, which again allows us to choose $s_{(u,v)}, t_{(u,v)} \in R$ appropriately.

($\Leftarrow$) : Assume that the variables $z, (x_u)_{u\in V\cup W}, (s_{(u,v)})_{(u,v)\in\rho_M\cup E}, (t_{(u,v)})_{(u,v)\in\rho_M\cup E}$ are set in such a way that $\Phi$ does not hold. Let $N := \langle z \rangle$ and let $g : W \to W$ be the function defined by $g(u) = (x_u, y_u)$. We show that $g$ is a homomorphism from $H_M$ to $H_N$. Let $(u, v) \in \rho_M$. Since $\Phi$ does not hold, its precondition must be satisfied, and thus $s_{(u,v)}(x_u - x_v) + t_{(u,v)}(y_u - y_v) = z$. Therefore $(g(u), g(v)) \in \rho_N$. By the same argument, $i : V \to W$ defined by $i(u) = (x_u, y_u)$ is a graph homomorphism from $G$ to $H_N$. Since $z \in N$, we have $N \neq 0$ and thus $N \in S$. We have shown $H_M \preceq H_N$. By transitivity of $\preceq$ and $z \neq 0$, we get $H_N \in S$. By maximality of $H_M$ in $S$, this implies $H_N \preceq H_M$, which gives us a graph homomorphism $j$ from $H_N$ to $H_M$. Now $j \circ i$ is the required graph homomorphism from $G$ to $H$. □

We now seek to generalize this proof to other finite algebras. Several concepts seem specific to rings:

- A ring has ideals.
- We can form the product of two ideals.
- If $R \cdot R = \{0\}$, then we can find a polynomial-time algorithm for PolSysSat($\mathbf{R}$) and thus QuasiIdVal($\mathbf{R}$)
- There is a term $t(x_1, x_2, y_1, \ldots, y_k)$ over the language of $\mathbf{R}$ such that

$$R \cdot \langle x_1, x_2 \rangle = \{t(x_1, x_2, y_1, \ldots, y_k) \mid y_1, \ldots, y_k \in R\}.$$

In the following chapter, we will see how the concept of ideals and ideal products can be generalized.

# 5 Universal algebra

In this chapter, we give a brief overview of universal algebra. One of the aims of universal algebra is to study how theorems about rings and groups generalize to other structures.

**Definition 5.1.** An *algebra* is a pair $\mathbf{A} = (A, F)$, where $F$ is a set of higher order functions $f : A^{a(f)} \to A$ on the nonempty set $A$.

If $F = \{f_1, \ldots, f_k\}$, we also write $(A, f_1, \ldots, f_k)$. For example, every group $\mathbf{G} = (G, \cdot, (\cdot)^{-1}, 1)$ and every ring $\mathbf{R} = (R, +, -, 0, \cdot)$ is an algebra. In this case, we interpret the group constant 1 as an operation of arity 0. Next we generalize the concept of the ideal of a ring. The reason why ideals are interesting is that $+, -, \cdot$ are all well defined on the classes of the equivalence relation $x \equiv_I y :\iff x - y \in I$. In general, we call such equivalence relations congruences:

**Definition 5.2.** Let $\mathbf{A} = (A, F)$ be an algebra. An equivalence relation $\alpha \subseteq A \times A$ is called a *congruence* of $\mathbf{A}$ if for all fundamental operations $f : A^k \to A$ and all $(a_1, b_1), \ldots, (a_k, b_k) \in \alpha$, we have $(f(a_1, \ldots, a_k), f(b_1, \ldots, b_k)) \in \alpha$. We write $\mathrm{Con}(\mathbf{A})$ for the set of all congruences on $\mathbf{A}$.

Given a congruence $\alpha$ and $a \in A$, we can define the *class* of $x$ as

$$a/\alpha := \{b \in A \mid (a, b) \in \alpha\}.$$

We write $A/\alpha := \{a/\alpha \mid a \in A\}$. For a $k$-ary fundamental operation $f$ of $\mathbf{A}$ and a congruence $\alpha$, the operation $f_\alpha : (A/\alpha)^k \to A/\alpha$ with

$$f_\alpha(a_1/\alpha, \ldots, a_k/\alpha) := f(a_1, \ldots, a_k)/\alpha$$

is well defined. This allows us to define the *quotient algebra* $\mathbf{A}/\alpha := (A/\alpha, \{f_\alpha \mid f \in F\})$.

It is easy to see that the intersection of congruences is again a congruence. Given a relation $R \subseteq A^2$, we therefore define the *congruence generated by* $R$ as

$$\Theta_{\mathbf{A}}(R) := \bigcap \{\theta \in \mathrm{Con}(\mathbf{A}) \mid R \subseteq \theta\}.$$

For $\boldsymbol{a}, \boldsymbol{b} \in A^n$, we write $\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}) := \Theta_{\mathbf{A}}(\{(a_1, b_1), \ldots, (a_n, b_n)\})$.

For groups, congruences are determined by their 1-classes, which are normal subgroups. A subgroup $N$ of a group $G$ is *normal* if $g^{-1}xg \in N$ for all $n \in N$ and $g \in G$.

**Lemma 5.3.** *Let* $\mathbf{G} = (G, \cdot, (\cdot)^{-1}, 1)$ *be a group and let* $\alpha$ *be a congruence of* $\mathbf{G}$*. Then* $\alpha = \{(a, b) \mid ab^{-1} \in N\}$ *for a normal subgroup* $N$ *of* $G$*.*

*Proof.* Choose $N := 1/\alpha \subseteq G$. Clearly $1 \in N$ and for $a, b \in N$, we have $a/\alpha = b/\alpha = 1/\alpha$ and thus

$$a \cdot b \in (a \cdot b)/\alpha = (a/\alpha) \cdot_\alpha (b/\alpha) = (1/\alpha) \cdot_\alpha (1/\alpha) = 1/\alpha = N$$

Similarly $a^{-1} \in N$, therefore $N$ is a subgroup of $G$. Now let $x \in N$ and $g \in G$. We then have

$$gxg^{-1} \in (gxg^{-1})/\alpha = (g/\alpha) \cdot_\alpha (x/\alpha) \cdot_\alpha (g/\alpha)^{-1} = (g/\alpha) \cdot_\alpha (g/\alpha)^{-1} = 1\alpha = N.$$

Therefore $N$ is a normal subgroup. □

**Lemma 5.4.** *Let* $\mathbf{G}$ *be a group and let* $N$ *be a normal subgroup of* $G$*. Then* $\alpha := \{(a, b) \in G \times G \mid ab^{-1} \in N\}$ *is a congruence of* $\mathbf{G}$*.*

*Proof.* Reflexivity: For all $a \in G$, we have $aa^{-1} = 1 \in N$. Symmetry: For all $ab^{-1} \in N$, we have $ba^{-1} = (ab^{-1})^{-1} \in N$. Transitivity: For all $ab^{-1}, bc^{-1} \in N$ we have $ac^{-1} = ab^{-1}bc^{-1} \in N$. We show that $\alpha$ preserves the fundamental operations $\cdot, (\cdot)^{-1}$ and 1. To this end, let $a_1b_1^{-1}, a_2b_2^{-1} \in N$. For $(\cdot)$, we have

$$\begin{aligned}(a_1 \cdot a_2)(b_1 \cdot b_2)^{-1} &= a_1a_2b_2^{-1}b_1^{-1} \\ &= (a_1a_1^{-1})(a_1a_2b_2^{-1}b_1^{-1})(a_1a_1^{-1}) \\ &= a_1(\underbrace{(a_2b_2^{-1})}_{\in N}\underbrace{(b_1^{-1}a_1)}_{\in N})a_1^{-1} \in N.\end{aligned}$$

For $(\cdot)^{-1}$ we have $(a_1^{-1})(b_1^{-1})^{-1} = a_1^{-1}b_1 = a_1^{-1}(b_1a_1^{-1})a_1 = a_1^{-1}(a_1b_1^{-1})^{-1}a_1 \in N$. For 1, we clearly have $1 \in N$. □

For rings, congruences are determined by their 0−classes, which are ideals:

**Lemma 5.5.** *Let* $\mathbf{R} = (R, +, -, 0, \cdot)$ *be a ring and let* $I \subseteq R$ *be a two-sided ideal of* $\mathbf{R}$. *Then* $\alpha := \{(a, b) \in R \times R \mid a - b \in I\}$ *is a congruence of* $\mathbf{R}$.

*Proof.* Reflexivity: For all $a \in R$, we have $a-a = 0 \in I$. Symmetry: For all $a-b \in I$ we have $b-a = -(a-b) \in I$. Transitivity: For $a-b, b-c \in I$ we have $a-c = (a-b)+(b-c) \in I$. To see that $\alpha$ is a congruence, let $a_1 - b_1, a_2 - b_2 \in I$. We prove that all fundamental operations preserve $\alpha$. Addition: $(a_1+a_2)-(b_1+b_2) = (a_1-b_1)+(a_2-b_2) \in I$. Subtraction: $(a_1 - a_2) - (b_1 - b_2) = (a_1 - b_1) - (a_2 - b_2) \in I$. Zero: $0 \in I$. Multiplication:

$$(a_1 \cdot a_2) - (b_1 \cdot b_2) = a_1(a_2 - b_2) + (a_1 - b_1)b_2 \in I.$$

□

Next, we need to generalize the concept of the ideal product to the setting of congruences. If $I$ and $J$ are ideals of a ring $\mathbf{R}$ and if $a - b \in I, c - d \in J$, then $ac - bc - ad + bd = (a - b)(c - d) \in IJ$. More generally $t(a, c) - t(b, c) - t(a, d) + t(b, d) \in IJ$ for all term functions $t$. In particular, $t(a, c) - t(a, d) \in IJ$ implies $t(b, c) - t(b, d) \in IJ$.

This last property turns out to be useful for defining the commutator:

**Definition 5.6** ([5], Definition 3.2)**.** Let $\mathbf{A}$ be an algebra and let $\alpha, \beta, \eta$ be congruences of $\mathbf{A}$. We say that $\alpha$ *centralizes* $\beta$ *modulo* $\eta$ if for all $m, n \in \mathbb{N}$, $(a_1, b_1), \ldots, (a_m, b_m) \in \alpha$ and $(c_1, d_1), \ldots, (c_n, d_n) \in \beta$ and all term functions $t$ in $m + n$ arguments,

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

In this case, we also write $C(\alpha, \beta, \eta)$.

**Lemma 5.7.** *Let* $\mathbf{A}$ *be an algebra, let* $\alpha, \beta$ *be congruences on* $\mathbf{A}$ *and let*

$$\mu := \bigcap \{\eta \in \mathrm{Con}(\mathbf{A}) \mid C(\alpha, \beta, \eta)\}.$$

*Then $C(\alpha, \beta, \mu)$ holds.*

*Proof.* Let $m, n \in \mathbb{N}$ and let $(a_1, b_1), \dots, (a_m, b_m) \in \alpha, (c_1, d_1), \dots, (c_n, d_n) \in \beta$. Let $t$ be a term in $m + n$ arguments. Assume $(t(\boldsymbol{b}, \boldsymbol{c}), t(\boldsymbol{b}, \boldsymbol{d})) \in \mu$. Let $\eta$ be a congruence with $C(\alpha, \beta, \eta)$. Then $(t(\boldsymbol{a}, \boldsymbol{c}), t(\boldsymbol{a}, \boldsymbol{d})) \in \eta$, and therefore $(t(\boldsymbol{b}, \boldsymbol{c}), t(\boldsymbol{b}, \boldsymbol{d})) \in \eta$. Hence $(t(\boldsymbol{b}, \boldsymbol{c}), t(\boldsymbol{b}, \boldsymbol{d})) \in \mu$. □

The congruence $[\alpha, \beta] := \mu$ is called the *commutator* of $\alpha$ and $\beta$ (see e.g. [5, Definition 3.2])

**Lemma 5.8.** *Let* $\mathbf{R}$ *be a commutative ring and let* $I, J$ *be ideals of* $\mathbf{R}$.
*Then* $[\equiv_I, \equiv_J] = \equiv_{IJ}$ .

*Proof.* ($\subseteq$) : We show that $\equiv_I$ centralizes $\equiv_J$ modulo $\equiv_{IJ}$: Let $m, n \in \mathbb{N}$, let $a_1 - b_1, \dots, a_m - b_m \in I, c_1 - d_1, \dots, c_n - d_n \in J$. Let $t$ be a term function in $m + n$ arguments. Because $\mathbf{R}$ is a commutative ring, $t(\boldsymbol{x}, \boldsymbol{y}) = \sum_{(\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{N}_0^m \times \mathbb{N}_0^n} c(w) \boldsymbol{x}^{\boldsymbol{u}} \boldsymbol{y}^{\boldsymbol{v}}$ for some coefficient function $c : \mathbb{N}_0^m \times \mathbb{N}_0^n \to \{0, 1, -1, 1 + 1, -(1 + 1), \dots\} \subseteq R$ of finite support.

Because $\equiv_I$ and $\equiv_J$ are congruences, we get $\boldsymbol{a}^{\boldsymbol{u}} \equiv_I \boldsymbol{b}^{\boldsymbol{u}}$ and $\boldsymbol{c}^{\boldsymbol{v}} \equiv_I \boldsymbol{d}^{\boldsymbol{v}}$ for all $(\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{N}_0^m \times \mathbb{N}_0^n$. Therefore

$$t(\boldsymbol{a}, \boldsymbol{c}) - t(\boldsymbol{b}, \boldsymbol{c}) - t(\boldsymbol{a}, \boldsymbol{d}) + t(\boldsymbol{b}, \boldsymbol{d}) = \sum_{(\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{N}_0^m \times \mathbb{N}_0^n} c(w)(\boldsymbol{a}^{\boldsymbol{u}} - \boldsymbol{b}^{\boldsymbol{u}})(\boldsymbol{c}^{\boldsymbol{v}} - \boldsymbol{d}^{v}) \in IJ$$

and thus $t(\boldsymbol{a}, \boldsymbol{c}) - t(\boldsymbol{a}, \boldsymbol{d}) \in IJ$ implies $t(\boldsymbol{b}, \boldsymbol{c}) - t(\boldsymbol{b}, \boldsymbol{d}) \in IJ$.

($\supseteq$) : Let $\eta$ be a congruence such that $\equiv_I$ centralizes $\equiv_J$ modulo $\eta$. We show that $IJ \subseteq 0/\eta$ : Let $i \in I$ and $j \in J$. Choose $t(x, y) = x \cdot y$. Clearly $(t(i, 0), t(0, 0)) \in \eta$ and thus $(t(i, j), t(0, j)) \in \eta$, hence $ij \in 0/\eta$. □

For non-commutative rings, we get $[\equiv_I, \equiv_J] = \equiv_{\langle (IJ) \cup (JI) \rangle}$ . In the case of groups, the commutator of the congruences induced by the normal subgroups $M$ and $N$ is the congruence induced by the commutator $[M, N] := \langle \{x^{-1} y^{-1} x y \mid x \in M, y \in N\} \rangle$. These observations are explored in Chapter 1 of [5].

**Definition 5.9.** An algebra $\mathbf{A}$ is called *abelian* if the congruences $1_A := A \times A$ and $0_A := \{(a, a) \mid a \in A\}$ safisfy $[1_A, 1_A] = 0_A$.

For abelian algebras, the centralizing condition becomes $t(\boldsymbol{a}, \boldsymbol{c}) = t(\boldsymbol{b}, \boldsymbol{c}) \Rightarrow t(\boldsymbol{a}, \boldsymbol{d}) = t(\boldsymbol{b}, \boldsymbol{d})$ for all $\boldsymbol{a}, \boldsymbol{b} \in A^m, \boldsymbol{c}, \boldsymbol{d} \in A^n$. In other words, if the term function $t$ is constant along one edge of a square, it is constant along the parallel edge as well. This property can be observed for affine functions in a vector space.

If $\mathbf{R} = (R, +, -, 0, \cdot)$ is a ring, then $\mathbf{R}$ is abelian if and only if $a \cdot b = 0$ for all $a, b \in R$. Such rings are sometimes called *zero rings.* If $\mathbf{G}$ is a group, then $\mathbf{G}$ is abelian if and only if $ab = ba$ for all $a, b \in G$.

In our effort to generalize Theorem 4.3, we need one last ingredient: We need to find a term function $t$ in $2 + m$ arguments such that for all $\boldsymbol{a}, \boldsymbol{b} \in A^2$, we have

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

In the search of such a term function, it is helpful to require the existence of a ternary term function $m$ with the property that $m(x, x, y) = y = m(y, x, x)$ for all $x, y \in A$. We call algebras with such a term function *Mal'cev algebras.* Every ring is a Mal'cev algebra because $m(x, y, z) := x - y + z$ satisfies the condition. Similarly, every group is a Mal'cev algebra with $m(x, y, z) := xy^{-1}z$. Under this additional requirement, it is now possible to prove the following:

**Lemma 5.10.** *Let $\mathbf{A}$ be a finite Mal'cev algebra and let $k \in \mathbb{N}$. Then there exist $m \in \mathbb{N}$ and a term $t$ of arity $k + m$ such that for all $\boldsymbol{a}, \boldsymbol{b} \in A^k$, we have*

$$[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A] = \{(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \mid \boldsymbol{w} \in A^m\}. \tag{5.1}$$

A full proof can be found in [1]. We give a proof for only the special case of groups.
*Proof for groups.* Assume that $\mathbf{A} = (A, \cdot, (\cdot)^{-1}, 1)$ is a finite group. Then $1/[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A] = [\langle a_1 b_1^{-1}, \ldots, a_k b_k^{-1} \rangle, A] =: H(\boldsymbol{a}, \boldsymbol{b})$ for all $\boldsymbol{a}, \boldsymbol{b} \in A^k$. For $r \in \mathbb{N}, l : \underline{r} \to \mathbb{N}, j : \bigcup_{i \in \underline{r}} \underline{l(i)} \times \{i\} \to \underline{k}$ let $t_{r,l,j}$ be the term function

$$t_{r,l,j}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{x}, \boldsymbol{y}) := \prod_{i=1}^{r} [\prod_{m=1}^{l(i)} x_{m,i} a_{j(m,i)} b_{j(m,i)}^{-1} x_{m,i}^{-1}, y_i].$$

Let $T$ be the set of all such term functions. Note that $T$ is closed under multiplication and for every $t \in T$ of arity $k+k+u+v$, the set $I(t, \boldsymbol{a}, \boldsymbol{b}) := \{t_{r,l,j}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{x} \in A^u, \boldsymbol{y} \in A^v\}$ has the property that $\bigcup_{t \in T} I(t, \boldsymbol{a}, \boldsymbol{b}) = H(\boldsymbol{a}, \boldsymbol{b})$ and $\{1\} \subseteq I(t, \boldsymbol{a}, \boldsymbol{b}) \subseteq H(\boldsymbol{a}, \boldsymbol{b})$ for all $t \in T$ and $\boldsymbol{a}, \boldsymbol{b} \in A^k$. We write $s \preceq t$ if $I(s, \boldsymbol{a}, \boldsymbol{b}) \subseteq I(s, \boldsymbol{a}, \boldsymbol{b})$ for all $\boldsymbol{a}, \boldsymbol{b} \in A^k$. We now choose $s \in T$ such that $I(s)$ is maximal with respect to $\preceq$ . Such an $s$ must exist as there are only finitely many functions which map pairs $\boldsymbol{a}, \boldsymbol{b} \in A^k$ to subsets of $A$. Seeking a contradiction, assume that there are $\boldsymbol{a}, \boldsymbol{b} \in A^k$ and $x \in H(\boldsymbol{a}, \boldsymbol{b})$ such that $x \notin I(s, \boldsymbol{a}, \boldsymbol{b})$. Choose $t \in T$ with $x \in I(t, \boldsymbol{a}, \boldsymbol{b})$. We now have $s \cdot t \in T$ and $I(s \cdot t, \boldsymbol{c}, \boldsymbol{d}) \supseteq I(s, \boldsymbol{c}, \boldsymbol{d}) \cup I(t, \boldsymbol{c}, \boldsymbol{d}) \supsetneq I(s, \boldsymbol{c}, \boldsymbol{d})$ for all $\boldsymbol{c}, \boldsymbol{d} \in A^k$, contradicting maximality of $s$. We can use this term $s$ to induce the commutator as

$$[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A] = \{(y \cdot s(\boldsymbol{a}, \boldsymbol{w}), y) \mid \boldsymbol{w} \in A^m, y \in A\}.$$

□

In the case of general Mal'cev algebras, the proof follows similar steps:

1. Identify a set $T = \bigcup_{m \geq 0} T_m$ of term functions of arity $k + k + m$ such that their images satisfy $[\Theta_{\mathbf{A}}(\boldsymbol{a}, \boldsymbol{b}), 1_A] = \bigcup_{m \geq 0} \bigcup_{t \in T_m} \{(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \mid \boldsymbol{w} \in A^m\}$.
2. Use finiteness of the commutator to choose a term $s$ such that the function $(\boldsymbol{a}, \boldsymbol{b}) \mapsto \{(t(\boldsymbol{a}, \boldsymbol{w}), t(\boldsymbol{b}, \boldsymbol{w})) \mid \boldsymbol{w} \in A^m\}$ is maximal with respect to pointwise set inclusion.
3. Prove that this term $s$ actually generates the entire commutator for all $\boldsymbol{a}, \boldsymbol{b} \in A^k$.

# 6 Graphs and the completeness proof

We are now in a position to prove the general version of the main theorem. A version of this chapter first appeared in [1]. For establishing coNP-completeness, we again use the result of Hell and Nešetřil.

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

**Lemma 6.1.** *Let $\mathbf{A}$ be a finite Mal'cev algebra, and let $\mu \in \mathrm{Con}(\mathbf{A})$ 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 $\rho_\mu$, 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 [12, 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 6.2.** *Let $\mathbf{A}$ be a finite nonabelian Mal'cev algebra. Then there is $\beta \in \mathrm{Con}(\mathbf{A})$ such that $\beta > 0_A$ and $\mathbb{H}_\beta = (H, \rho_\beta)$ has a triangle.*

*Proof.* Let $\zeta$ be the center of $\mathbf{A}$, i.e., the largest congruence with $[\zeta, 1_A] = 0_A$. We note that from [12, Lemma 4.149(ii)] it follows that such a largest congruence exists: in fact, $\zeta$ is the join of all congruences $\zeta'$ with $[\zeta', 1_A] = 0_A$. In particular, every congruence $\zeta' \in \mathrm{Con}(\mathbf{A})$ 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 6.3.** *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 $\gamma$. 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 $t(x, y, z_1, \ldots, z_m)$ 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* (5.1)*; such a term exists by Lemma 5.10. Let $\mathbb{G} = (G, \rho^{\mathbb{G}})$ be a graph. We assume that $G \cap H_\mu = \varnothing$. Let $\Phi$ 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* $\Phi$ *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 $\Phi$ in a way that contradicts the validity of $\Phi$. 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 $\binom{x_u}{y_u} = f(u)$. 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 $\mu \leq [\Theta(\binom{x_u}{y_u}, \binom{x_v}{y_v}), 1_A]$. Since $(a,b) \in \mu$, Lemma 5.10 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 $u = \binom{x_u}{y_u}$. Then for each $(u,v) \in \rho_\mu$, we have $(\binom{x_u}{y_u}, \binom{x_v}{y_v}) = (u,v) \in \rho_\mu$. Hence from the definition of $\rho_\mu$, we obtain

$$\mu \leq [\Theta(\tbinom{x_u}{y_u}, \tbinom{x_v}{y_v}), 1_A].$$

Thus from Lemma 5.10 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 $\Phi$ is not valid in $\mathbf{A}$.

For the "if"-direction, we assume that that the variables in $\Phi$ are assigned such that $\Phi$ 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

$$g(u) = \tbinom{x_u}{y_u} \text{ for all } u \in H_\mu.$$

Next, we prove that $g$ is a homomorphism from $\mathbb{H}_\mu$ to $\mathbb{H}_\tau$. Since $\Phi$ 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 $\boldsymbol{a} := \binom{x_u}{y_u}$ and $\boldsymbol{b} := \binom{x_v}{y_v}$ in Lemma 5.10, we obtain $(a,b) \in [\Theta(\binom{x_u}{y_u}, \binom{x_v}{y_v}), 1_A]$, and therefore $\tau \leq [\Theta(\binom{x_u}{y_u}, \binom{x_v}{y_v}), 1_A]$. From the definition of the difference graph $\mathbb{H}_\tau$, we now see that $(\binom{x_u}{y_u}, \binom{x_v}{y_v}) = (g(u), g(v))$ 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 $j(u) := \binom{x_u}{y_u}$. Using the same reasoning as for $g$, we show that $j$ is a homomorphism: assume that $(u, v) \in \rho^{\mathbb{G}}$. Since $\Phi$ 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 $(a, b) \in [\Theta(\binom{x_u}{y_u}, \binom{x_v}{y_v}), 1_A]$, which implies that $\tau \leq [\Theta(\binom{x_u}{y_u}, \binom{x_v}{y_v}), 1_A]$. Thus $(j(u), j(v)) = (\binom{x_u}{y_u}, \binom{x_v}{y_v})$ 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.

**Theorem 6.4.** *Let* $\mathbf{A}$ *be a finite nonabelian Mal'cev algebra of finite type. Then* QuasiIdVal($\mathbf{A}$) *is coNP-complete.*

*Proof.* From Lemma 6.2, 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 6.3. 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 Theorem 4.2, we obtain that $\mathbb{H}_\mu$-coloring is NP-complete. By Lemma 6.3, 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 $\Phi$ that can be computed in time polynomial in the size of $\mathbb{G}$. This implies that QuasiIdVal($\mathbf{A}$) is coNP-complete. □

We therefore obtain a full dichotomy:

**Corollary 6.5.** *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.*

*Proof.* Given Theorem 6.4, we only have to verify that QuasiIdVal($\mathbf{A}$) is in $P$ when $\mathbf{A}$ is an abelian finite Mal'cev algebra of finite type. In chapter 2, we observed that QuasiIdVal($\mathbf{A}$) can be reduced to 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. From this reduction, we see that it is sufficient to show that 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 [10,

Corollary 3.14] (which is proved using [4, Theorem 33]) that in this case PolSysSat($\mathbf{A}$) can be solved in polynomial time. □

As special cases, we obtain:

**Corollary 6.6.** *For a finite group* $\mathbf{G}$*,* 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),* QuasiIdVal($\mathbf{R}$) *is in P if* $\mathbf{R}$ *is a zero ring, i.e.,* $R \cdot R = 0$*, and coNP-complete otherwise.*

# 7 Hardness of systems satisfiability revisited

In this chapter, we use the techniques developed in the proof of Theorem 6.4 to find a new proof for NP-completeness of PolSysSat($\mathbf{A}$) for nonabelian Mal'cev algebras, which was first proven in [10].

**Lemma 7.1.** *Let $\mathbf{A} = (A, F)$ be a finite nonabelian algebra with Mal'cev term. Then PolSysSat($\mathbf{A}$) is NP-complete.*

*Proof.* Let $\zeta$ be the center of $\mathbf{A}$. Choose $a, b \in A$ such that $(a, b) \notin \zeta$. Choose $(u, v) \in [\Theta(a, b), \mathbf{1}_\mathbf{A}]$ with $u \neq v$ and let $\mu := \Theta(u, v)$. Let $\mathbb{H}_\mu = (H, \rho_\mu)$ be the difference graph as defined in the previous section. Consider the vectors $\boldsymbol{r} = \binom{a}{a}$, $\boldsymbol{s} = \binom{a}{b}$, $\boldsymbol{t} = \binom{b}{a}$. Note that $[\Theta(\boldsymbol{r}, \boldsymbol{s}), \mathbf{1}_\mathbf{A}] = [\Theta(\boldsymbol{s}, \boldsymbol{t}), \mathbf{1}_\mathbf{A}] = [\Theta(\boldsymbol{r}, \boldsymbol{t}), \mathbf{1}_\mathbf{A}] = [\Theta(a, b), \mathbf{1}_\mathbf{A}] \geq \mu$. Therefore $(\boldsymbol{r}, \boldsymbol{s}), (\boldsymbol{s}, \boldsymbol{t}), (\boldsymbol{t}, \boldsymbol{r}) \in \rho_\mu$ and hence the $\mathbb{H}_\mu$ is non-bipartite and loop-free. Therefore $\mathbb{H}_\mu$-Colorability is again NP-complete. We will provide a reduction from $\mathbb{H}_\mu$-Colorability to PolSysSat($\mathbf{A}$).

To this end, let $G = (V, E)$ be a graph. We will construct a system of polynomial equations which is satisfiable if and only if there exists a homomorphism from $G$ to $\mathbb{H}_\mu$. Using Lemma 5.10, choose $m \in \mathbb{N}$ and a term $t$ in $2 + m$ variables such that for all $\boldsymbol{x}, \boldsymbol{y} \in A^2$,

$$[\Theta(\boldsymbol{x}, \boldsymbol{y}), \mathbf{1}_\mathbf{A}] = \{(t(\boldsymbol{x}, \boldsymbol{z}), t(\boldsymbol{y}, \boldsymbol{z})) \mid \boldsymbol{z} \in A^m\}.$$

Consider the polynomial system $\Phi$:

$$\bigwedge_{(i,j) \in E} t(x_i, y_i, z_{(i,j)}) = u \wedge t(x_j, y_j, z_{(i,j)}) = v$$

We need to show that $\Phi$ is satisfiable if and only if there exists a graph $h$ homomorphism from $G$ to $\mathbb{H}_\mu$. ($\Rightarrow$) : Assume that $\Phi$ is satisfiable. Choose a satisfying assignment $(x_i)_{i\in V}, (z_{(i,j)})_{(i,j)\in E}$. Consider the function $h : V \to H, v \mapsto \binom{x_v}{y_v}$. We need to show that $h$ is a graph homomorphism. To this end, let $(i,j) \in E$. We know that $(u,v) = (t(x_i,y_i,z_{(i,j)}), t(x_j,y_i,z_{(i,j)}))$, therefore $(u,v) \in [\Theta(\binom{x_i}{y_i}, \binom{x_j}{y_j}), \mathbf{1_A}]$, thus $\mu \leq [\Theta(\binom{x_i}{y_i}, \binom{x_j}{y_j}), \mathbf{1_A}]$, hence $(\binom{x_i}{y_i}, \binom{x_j}{y_j}) \in \rho_\mu$. ($\Leftarrow$) : Let $h$ be a graph homomorphism from $G$ to $\mathbb{H}_\mu$. We need to find a satisfying assignment of $\Phi$. To this end, let $\binom{x_i}{y_i} := h(i)$ for $i \in V$. We then know that for $(i,j) \in E$, $(x_i,x_j) \in \mathbb{H}_\mu$, therefore $\Theta(u,v) \leq [\Theta(\binom{x_i}{y_i}, \binom{x_j}{y_j}), \mathbf{1_A}]$, thus $(u,v) \in \{(t(x_i,y_i,\boldsymbol{z}), t(x_j,y_j,\boldsymbol{z})) \mid \boldsymbol{z} \in A^m\}$. Hence we can find a $\boldsymbol{z}_{(i,j)} \in A^m$ such that $t(x_i,y_i,\boldsymbol{z}_{(i,j)}) = u$ and $t(x_j,y_j,\boldsymbol{z}_{(i,j)}) = v$. □

# Bibliography


[1] Erhard Aichinger and Simon Grünbacher. "The Complexity of Checking Quasi-Identities over Finite Algebras with a Mal'cev Term". In: *40th International Symposium on Theoretical Aspects of Computer Science (STACS 2023)*. Ed. by Petra Berenbrink et al. Vol. 254. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 4:1–4:12. ISBN: 978-3-95977-266-2. DOI: 10.4230/LIPIcs.STACS.2023.4 (cit. on pp. 2, 10, 20, 22).

[2] Andrei A. Bulatov. "A Dichotomy Theorem for Nonuniform CSPs". In: *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*. 2017, pp. 319–330. DOI: 10.1109/FOCS.2017.37 (cit. on p. 5).

[3] Stanley Burris and John Lawrence. "The equivalence problem for finite rings". In: *J. Symbolic Comput.* 15.1 (1993), pp. 67–71. ISSN: 0747-7171. DOI: 10.1142/S021819671100625X (cit. on pp. 8, 10).

[4] Tomás Feder and Moshe Y. Vardi. "The computational structure of monotone monadic SNP and constraint satisfaction: a study through Datalog and group theory". In: *SIAM J. Comput.* 28.1 (1999), 57–104 (electronic). ISSN: 0097-5397. DOI: 10.1137/S0097539794266766 (cit. on p. 26).

[5] Ralph Freese and Ralph N. McKenzie. *Commutator Theory for Congruence Modular varieties*. Vol. 125. London Math. Soc. Lecture Note Ser. Cambridge University Press, 1987 (cit. on pp. 18, 19).

[6] Mikael Goldmann and Alexander Russell. "The Complexity of Solving Equations over Finite Groups". In: *Information and Computation* 178.1 (2002), pp. 253–262. ISSN: 0890-5401. DOI: https://doi.org/10.1006/inco.2002.3173 (cit. on pp. 8, 10).

[7] Pavol Hell and Jaroslav Nešetřil. "On the complexity of $H$-coloring". In: *J. Combin. Theory Ser. B* 48.1 (1990), pp. 92–110. ISSN: 0095-8956. DOI: 10.1016/0095-8956(90)90132-J (cit. on p. 13).

[8] Gábor Horváth and Csaba Szabo. "The extended equivalence and equation solvability problems for groups". In: *Discrete Mathematics & Theoretical Computer Science* 13 (Nov. 2011), pp. 23–32. DOI: 10.46298/dmtcs.536 (cit. on pp. 9, 10).

[9] Gábor Horváth and Csaba Szabó. "Equivalence and equation solvability problems for the alternating group A4". In: *Journal of Pure and Applied Algebra* 216.10 (2012), pp. 2170–2176. ISSN: 0022-4049. DOI: https://doi.org/10.1016/j.jpaa.2012.02.007 (cit. on p. 8).

[10] Benoit Larose and László Zádori. "Taylor Terms, Constraint Satisfaction and the Complexity of Polynomial Equations over Finite Algebras". In: *International Journal of Algebra and Computation* 16.03 (June 2006), pp. 563–581. DOI: 10.1142/s0218196706003116 (cit. on pp. 5, 10, 25, 27).

[11] Peter Mayr. "On the Complexity Dichotomy for the Satisfiability of Systems of Term Equations over Finite Algebras". In: *48th International Symposium on Mathematical Foundations of Computer Science (MFCS 2023)*. Ed. by Jérôme Leroux, Sylvain Lombardy, and David Peleg. Vol. 272. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 66:1–66:12. ISBN: 978-3-95977-292-1. DOI: 10.4230/LIPIcs.MFCS.2023.66 (cit. on p. 4).

[12] Ralph N. McKenzie, George F. McNulty, and Walter F. Taylor. *Algebras, lattices, varieties, Volume I*. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, California, 1987 (cit. on pp. 22, 23).

[13] Mikhail V. Volkov. "Checking quasi-identities in a finite semigroup may be computationally hard". In: *Studia Logica* 78.1-2 (2004), pp. 349–356. ISSN: 0039-3215. DOI: 10.1007/s11225-005-0356-5 (cit. on p. 2).

[14] Dmitriy Zhuk. "A Proof of the CSP Dichotomy Conjecture". In: *Journal of the ACM* 67.5 (2020), pp. 1–78. DOI: 10.1145/3402029 (cit. on p. 5).