Satisfiability of Circuits and Equations over Finite Malcev Algebras

Paweł M. Idziak
Department of Theoretical Computer Science, Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland

Piotr Kawałek
Department of Theoretical Computer Science, Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland

Jacek Krzaczkowski
Department of Computer Science, Faculty of Mathematics, Physics and Computer Science, Maria Curie-Sklodowska University, Lublin, Poland

Abstract

We show that the satisfiability of circuits over finite Malcev algebra $\mathbf{A}$ is NP-complete or $\mathbf{A}$ is nilpotent. This strengthens the result from our earlier paper [18] where nilpotency has been enforced, however with the use of a stronger assumption that no homomorphic image of $\mathbf{A}$ has NP-complete circuits satisfiability. Our methods are moreover strong enough to extend our result of [14] from groups to Malcev algebras. Namely we show that tractability of checking if an equation over such an algebra $\mathbf{A}$ has a solution enforces its nice structure: $\mathbf{A}$ must have a nilpotent congruence ν such that also the quotient algebra $\mathbf{A}/\nu$ is nilpotent. Otherwise, if $\mathbf{A}$ has no such congruence ν then the Exponential Time Hypothesis yields a quasipolynomial lower bound. Both our results contain important steps towards a full characterization of finite algebras with tractable circuit satisfiability as well as equation satisfiability.

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness; Theory of computation → Complexity classes

Keywords and phrases Circuit satisfiability, solving equations, Exponential Time Hypothesis

Digital Object Identifier 10.4230/LIPIcs.STACS.2022.37

Funding The project is partially supported by Polish NCN Grant # 2014/14/A/ST6/00138.
Piotr Kawałek: This research is partially supported by the Priority Research Area SciMat under the program Excellence Initiative – Research University at the Jagiellonian University in Kraków.

1 Introduction

The problem of deciding whether an equation over an algebraic structure has a solution has got quite deep attention both in mathematics and computer science. Let us only mention two crucial examples:

- 10^{th} Hilbert's problem for Diophantine equations, i.e. equations over the ring of integers – shown to be undecidable by Matiyasevich [24],
- the problem SAT of satisfiability of Boolean formulas – shown to be NP-complete by Cook [4].

In this paper we consider equations of the form $\mathbf{t}(x_1, \dots, x_n) = \mathbf{s}(x_1, \dots, x_n)$, where $\mathbf{t}$ and $\mathbf{s}$ are polynomials over a fixed finite algebra $\mathbf{A}$ (i.e. a finite set A with finitely many operations), i.e. terms with some of their variables already evaluated by elements of $\mathbf{A}$. We are interested in the complexity of the problem POLSAT($\mathbf{A}$), i.e. the problem of deciding whether an equation (of two polynomials) over $\mathbf{A}$ has a solution in $\mathbf{A}$.

Recently, this problem for equations over finite groups and other finite algebraic structures (like e.g. rings [10], semigroups [2, 21] or lattices [25]) attracted many researchers. For groups the story began with the paper [7] of Goldmann and Russell where NP-completeness of PolSat has been shown for nonsolvable groups and a polynomial time algorithm has been created for nilpotent groups. However the gap between solvable and nilpotent groups remained unfilled. More recently several examples of solvable but non-nilpotent groups with PolSat tractable in polynomial time have been provided [12, 13, 11, 5]. Among such groups there is the symmetric group $\mathbf{S}_3$ and the alternating group $\mathbf{A}_4$. Those two examples are of a special interest, as after endowing them by additional operations definable in terms of group multiplication the PolSat problem becomes NP-complete for such extensions. For the group $\mathbf{S}_3$ this phenomenon has been first described in [8], while for $\mathbf{A}_4$ in [13]. The NP-hardness for $\mathbf{A}_4$ has been obtained by extending this group by the binary commutator operation $[x, y] = x^{-1}y^{-1}xy$, which is heavily used in group theory.

The existence of such examples made the expectation of characterizing finite algebras with PolSat solvable in polynomial time rather hopeless. For this reason our paper [18] modified PolSat to make it independent of the choice of the basic operations in the algebra. This has been done by interpreting the size of a term (or a polynomial) not as the size of the tree that represent this term but as the size of a circuit computing it. One can easily see that the tree (built up with group multiplication only) computing the n-ary term $[\dots[[x_1, x_2], x_3]\dots x_n]$ has exponential size, while there is a circuit of linear size computing this term. This small change allows us to expand a finite algebra $\mathbf{A}$ by (finitely many) operations definable by polynomials of $\mathbf{A}$ without actually changing the complexity. Thus by a circuit satisfiability over an algebra $\mathbf{A}$ we mean the following problem

CSat($\mathbf{A}$): given a circuit over $\mathbf{A}$ with two output gates $\mathbf{g}_1, \mathbf{g}_2$, is there an assignment of values to the input gates $\overline{x} = (x_1, \dots, x_n)$ that gives the same output on $\mathbf{g}_1, \mathbf{g}_2$, i.e. $\mathbf{g}_1(\overline{x}) = \mathbf{g}_2(\overline{x})$.

Note here that the characterizations given in [10, 25] show that the problems PolSat and CSat are tractable for the same rings and lattices: namely the only tractable rings are the nilpotent rings and the only tractable lattices are the distributive lattices.

The paper [18] shows that replacing polynomials by circuits representing those polynomials allows us to attack the complexity of CSat in a more general setting than just for particular algebraic structures like groups, rings or lattices. The setting considered there includes all of the above structures and many more, i.e. algebras from the so-called congruence modular varieties. Roughly speaking, most of the structures in classical abstract algebra (except semigroups) are included.

In this paper we improve the result of [18] and generalize the result of [14] from groups to more general algebraic structures. First, in both of those two improvements we restrict ourselves to the so-called Malcev algebras, i.e. algebras having a ternary term $\mathbf{d}(x, y, z)$ that satisfies $\mathbf{d}(x, x, y) = y = \mathbf{d}(y, x, x)$. Note that groups and in fact all algebras that are extensions of groups (like rings or Boolean algebras) are Malcev, as the term $\mathbf{d}(x, y, z) = xy^{-1}z$ does the job. Also many generalizations of groups, like quasigroups or loops, are Malcev. Next we refer to the monograph [6] where a commutator theory is developed in a way that generalizes the commutator $[H, K]$ of normal subgroups H, K in the group theory and the ideal multiplication $I \cdot J$ in the ring theory. In general setting we use congruences instead of normal subgroups or ideals. The book [6] shows how for two congruences α, β of an algebra $\mathbf{A}$ define their commutator $[\alpha, \beta]$ and can serve as a reference source. With the help of the commutator one can define notions of abelianess, solvability and nilpotency for arbitrary algebras. First, for a congruence θ and $i = 1, 2, \dots$ we put

$$\begin{array}{rclcrcl} \theta^{(0)} & = & \theta & & \theta^{[0]} & = & \theta \\ \theta^{(i+1)} & = & [\theta, \theta^{(i)}] & & \theta^{[i+1]} & = & [\theta^{[i]}, \theta^{[i]}]. \end{array}$$

Now, a congruence θ of $\mathbf{A}$ is called *k-nilpotent* [or *k-solvable*] if $\theta^{(k)} = 0_{\mathbf{A}}$ [$\theta^{[k]} = 0_{\mathbf{A}}$] and the algebra $\mathbf{A}$ is *nilpotent* [*solvable*] if 1_A is k-nilpotent [k-solvable] for some finite k. In particular θ [or $\mathbf{A}$] is Abelian if $[\theta, \theta] = 0_{\mathbf{A}}$ [or $[1_{\mathbf{A}}, 1_{\mathbf{A}}] = 0_{\mathbf{A}}$]. In a similar way we can define what it means for a congruence θ to be Abelian, nilpotent or solvable over a smaller congruence α, by simply saying that an appropriate commutator power of θ is contained in α. Finally we define $\theta^{(\omega)} = \bigcap_{i=0}^{\infty} \theta^{(i)}$, and note that since in a finite algebra the descending chain $\theta^{(0)} \geqslant \theta^{(1)} \geqslant \theta^{(2)} \geqslant \dots$ stabilizes the congruence $\theta^{(\omega)}$ is in fact one of the $\theta^{(i)}$'s.

In our study of solvable Malcev algebra a series of congruences $0 = \nu_0 \leqslant \nu_1 \leqslant \dots \leqslant \nu_{h-1} \leqslant \nu_h = 1_{\mathbf{A}}$ in which ν_i is the largest nilpotent congruence over ν_{i-1} will play a crucial role. This series is called the nilpotent series (or sometimes the Fitting series in the group theory), and we will use this teminology as well. We define the nilpotent (or Fitting) rank $\mathsf{nr}(\alpha)$ of a congruence α to be the smallest integer k for which there is a sequence of congruences $0 = \alpha_0 \leqslant \alpha_1 \leqslant \dots \leqslant \alpha_{k-1} \leqslant \alpha_k = \alpha$ where each α_i is nilpotent over α_{i-1}. Note here that ν_k is the largest congruence with nilpotent rank k, so that we have $\mathsf{nr}(\alpha) \leqslant k$ iff $\alpha \leqslant \nu_k$. By the same token any solvable congruence in a Malcev algebra has finite nilpotent rank.

For a finite Malcev algebra $\mathbf{A}$ and a covering pair $\alpha \prec \beta$ of congruences (i.e. without any congruence between them) there are tools to describe the behaviour of $\mathbf{A}$ locally, depending on whether β is Abelian over α. In the case it is not, tame congruence theory (as described in [9]) tells us that $\mathbf{A}$ has a unary idempotent polynomial $\mathbf{e}$ (i.e. $\mathbf{e}(\mathbf{e}(x)) = \mathbf{e}(x)$ for all $x \in A$) with a two element range $\{0, 1\} = \mathbf{e}(A)$ so that the induced algebra $\mathbf{A}|_{\{0,1\}}$ (i.e. the set $\{0, 1\}$ with all the polynomials of $\mathbf{A}$ that preserve that set) has Boolean operations $\wedge, \vee, \neg$ definable by polynomials. As one can expect the presence of a Boolean behaviour results in NP-hardness of CSat($\mathbf{A}$) (see the proof of Corollary 1.3 for details). If there is no local Boolean behaviour in $\mathbf{A}$, i.e. $\mathbf{A}$ behaves locally in the Abelian fashion, then $\mathbf{A}$ is solvable. Thus our goal is to understand solvable algebras with tractable PolSat or even CSat.

In [18] it was shown that if a finite Malcev algebra $\mathbf{A}$ is not nilpotent then $\mathbf{A}$ has a nonnilpotent quotient $\mathbf{A}'$ with NP-complete CSat($\mathbf{A}'$). Here we significantly improve that result to be read as follows.

▶ **Theorem 1.1.** *If a finite Malcev algebra $\mathbf{A}$ is not nilpotent then CSat($\mathbf{A}$) is NP-complete.*

Unfortunately [18] does not provide a proof that nilpotency is already strong enough to force tractability. In fact [16] describes examples of nilpotent Malcev algebras with CSat outside P under the assumption of Exponential Time Hypothesis. On the other hand [18] (and independently [22]) provides an argument that supernilpotent Malcev algebras have tractable CSat. Due to [20], for algebras with finitely many basic operations this stronger condition of supernilpotency simply means that an algebra is nilpotent and decomposes into a direct product of algebras of prime power order. (Note here that due to Sylow's results every nilpotent group is already supernilpotent; the same is true for rings). Obviously the examples from [16] are not supernilpotent. In fact they are rather far from being supernilpotent. The very same paper [16] isolates a concept of supernilpotent rank, similar to the nilpotent rank, with the help of supernilpotent congruences instead of nilpotent ones. Thus we say that the supernilpotent rank of a congruence α is at most k (and write $\mathsf{sr}(\alpha) \leqslant k$) if there is a sequence

of congruences $0 = \alpha_0 \leqslant \alpha_1 \leqslant \ldots \leqslant \alpha_{k-1} \leqslant \alpha_k = \alpha$ in which each α_i is supernilpotent over α_{i-1}. (Note here that in [16] a congruence α with $\mathsf{sr}(\alpha) \leqslant k$ has been called k-step supernilpotent.) Analogously as in the case of the nilpotent series $(\nu_k)_k$ we can define the supernilpotent series $(\sigma_k)_k$ in which σ_k is the largest congruence of supernilpotent rank k (all that is needed to do that is to observe that the join of two supernilpotent congruences is supernilpotent). We also have $\mathsf{sr}(\alpha) \leqslant k$ iff $\alpha \leqslant \sigma_k$. Moreover, since every supernilpotent congruence is nilpotent, $\sigma_k \leqslant \nu_k$ and $\mathsf{nr}(\alpha) \leqslant \mathsf{sr}(\alpha)$. Note that the concepts of nilpotency and supernilpotency coincide in groups so that $\mathsf{nr}(\alpha) = \mathsf{sr}(\alpha)$ for every congruence α of a finite group. This however is not the case in general, as for nilpotent but not supernilpotent algebra $\mathbf{A}$ we have $\mathsf{nr}(\mathbf{A}) = 1 < \mathsf{sr}(\mathbf{A})$.

We have already noticed that the examples from [16] are not supernilpotent. Actually their supernilpotent rank is at least 3. Very recently Kompatscher [23] applied the technique from [16] to show that (under ETH) no finite nilpotent algebra with supernilpotent rank at least 3 has tractable CSat (but note here that Kompatscher use the term Fitting rank, for what we call here supernilpotent rank). However even $\mathsf{sr}(\mathbf{A}) \leqslant 2$ does not suffice to have tractable CSat (or ETH fails). Appropriate examples are created in [17].

The sequence of papers [16, 26, 14] explores the same idea to show that for a solvable group $\mathbf{G}$ with $\mathsf{nr}(\mathbf{G}) \geqslant 3$ the problem PolSat($\mathbf{G}$) is not tractable (if ETH holds). Actually [14] combines some premature results from [16] and [26]. Here we leave the group realm and use tame congruence theory (instead of well understood group techniques) to bound the nilpotent rank of solvable algebras with tractable PolSat.

▶ **Theorem 1.2.** *Let $\mathbf{A}$ be a finite solvable Malcev algebra with nilpotent rank $h \geqslant 3$. Then checking if an equation of length ℓ over $\mathbf{A}$ has a solution needs at least $2^{\Omega(\log^{h-1}\ell)}$ steps, or the Exponential Time Hypothesis fails.*

Combining Theorem 1.2 with the possibility of eliminating nonabelian (and therefore Boolean) local behaviour we will also get the following Corollary.

▶ **Corollary 1.3.** *Let $\mathbf{A}$ be a finite Malcev algebra. If PolSat($\mathbf{A}$) $\in P$ then $\mathbf{A}$ is solvable and $\mathsf{nr}(\mathbf{A}) \leqslant 2$, or ETH fails.*

We conclude the description of our results by noting that (under the Exponential Time Hypothesis) nilpotent rank 2 does not put PolSat into P, as some dihedral groups described in [17] show.

2 Proof of the Theorems

To prove the results formulated in the Introduction we need some preparation stated in two Lemmas below. Their proofs are postponed to Section 3 as they make some (or sometimes even quite heavy) use of the theory of commutator in congruence modular varieties (or the modular commutator theory, for short) and the tame congruence theory described in [6] and [9], respectively. This section however does not require the knowledge of these theories. All we need to state our Lemmas is the concept of a join irreducible congruence i.e. a congruence θ that cannot be represented as a join $\theta_1 \vee \theta_2$ for $\theta_1, \theta_2 < \theta$. Such a congruence has a unique subcover, which will be denoted by θ_-. Moreover θ has to be principal, i.e. it is generated by a single pair, say (a, b). This last fact is to be denoted by $\theta = \Theta(a, b)$. For these and all other basic algebraic concepts and notation we refer the reader to [3].

▶ **Lemma 2.1.** *Let* $\mathbf{A}$ *be a finite solvable Malcev algebra. Then for every join irreducible congruence* $\delta = \Theta(e, a)$ *with* $\mathsf{nr}(\delta) > \mathsf{nr}(\delta_-) > 0$ *there is another join irreducible congruence* $\delta^\star = \Theta(e^\star, a^\star)$ *with* $\mathsf{nr}(\delta^\star) = \mathsf{nr}(\delta) - 1$ *and* $\mathsf{nr}(\delta^\star) > \mathsf{nr}\left(\delta^\star_-\right)$*, and for each integer* n *there is an* n*-ary polynomial* $\mathbf{and}_n(x_1, \ldots, x_n)$ *such that for* $x_1, \ldots, x_n \in \{e, a\}$ *we have*

$$\mathbf{and}_n(x_1, \ldots, x_n) = \begin{cases} a^\star, & \text{if } x_1 = \ldots = x_n = a, \\ e^\star, & \text{otherwise.} \end{cases}$$

The polynomial $\mathbf{and}_n$ *can be constructed in a time bounded by* $2^{O(n)}$ *while the circuit that computes* $\mathbf{and}_n$ *can be constructed in a linear time* $O(n)$*.*

▶ **Lemma 2.2.** *In a finite solvable Malcev algebra* $\mathbf{A}$ *with nilpotent rank* $h \geqslant 2$ *there are:*

- *a join irreducible congruence* $\delta^{h-1} = \Theta(e_{h-1}, a_{h-1})$ *with* $h - 1 = \mathsf{nr}\left(\delta^{h-1}\right) > \mathsf{nr}\left(\delta^{h-1}_-\right) = h - 2$*,*
- *a partition of* A *into two nonempty disjoint subsets* $A = A_\perp \cup A_\top$*,*

such that for any 3*-CNF-formula* Φ *with* m *clauses there exists a* $3m$*-ary polynomial* $\mathbf{sat}_\Phi$ *of* $\mathbf{A}$ *with range contained in* $\{e_{h-1}, a_{h-1}\}$ *and such that for* $z_1^1, z_2^1, z_3^1, \ldots, z_1^m, z_2^m, z_3^m \in \{\top, \perp\}$ *and* $x_1^1, x_2^1, x_3^1, \ldots, x_1^m, x_2^m, x_3^m \in A$ *with* $x_i^j \in A_{z_i^j}$ *we have*

$$\Phi(z_1^1, z_2^1, z_3^1, \ldots, z_1^m, z_2^m, z_3^m) = \top \quad \textit{iff} \quad \mathbf{sat}_\Phi(x_1^1, x_2^1, x_3^1, \ldots, x_1^m, x_2^m, x_3^m) = a_{h-1}.$$

The polynomial $\mathbf{sat}_\Phi$ *can be constructed (from the formula* Φ*) in time bounded by* $2^{O(m)}$ *while the circuit that computes* $\mathbf{sat}_\Phi$ *can be constructed in linear time* $O(m)$*.*

Now we are ready to prove our results stated in the Introduction.

Proof of Theorem 1.2. We are going to translate a 3-CNF formula Φ with m clauses into an equation of length $2^{O(m^{1/(h-1)})}$ (and with the very same time needed for this translation) such that the formula Φ is satisfiable iff the corresponding equation has a solution. This, according to ETH (together with the Sparsification Lemma) shows that the time needed to check if an equation of length ℓ has a solution is at least $2^{\Omega(\log^{h-1}\ell)}$. For if not, a PolSat($\mathbf{A}$) algorithm working in $2^{o(\log^{h-1}\ell)}$ time would solve 3-CNF-SAT in $2^{o(m)}$, contrary to ETH.

Without loss of generality we assume that $m = k^{h-1}$. We will produce a $3m$-ary polynomial $\Phi\mathbf{Sat}$ represented by a tree having:

- exactly h levels,
- $3m$ leaves, all of them on the level h,
- $m/k = k^{h-2}$ nodes on level $h-1$, each of which labeled by $3k$-ary polynomial of the form **sat** provided by Lemma 2.2,
- $m/k^{h-l} = k^{l-1}$ nodes at the l-th level (for $l = h-2, ..., 1$), each of which labeled by k-ary polynomial of the form $\mathbf{and}_k$ supplied by Lemma 2.1.

To do that we start with a 3-CNF formula Φ with $m = k^{h-1}$ clauses and group them into m/k groups each of which containing exactly k clauses. Thus Φ can be represented as $\bigwedge_{j=1}^{m/k} \Phi_j$, with Φ_j being a conjunction of k clauses in the j-th group. Now, with the help of Lemma 2.2 we produce:

- the partition $A = A_\top \cup A_\perp$,
- the join irreducible congruence δ^{h-1} with $\mathsf{nr}\left(\delta^{h-1}\right) = h - 1 = 1 + \mathsf{nr}\left(\delta^{h-1}_-\right)$,
- the elements e_{h-1}, a_{h-1} with $\delta^{h-1} = \Theta(e_{h-1}, a_{h-1})$,
- and for each Φ_j a corresponding $3k$-ary polynomial $\mathbf{sat}_{\Phi_j}$ with the property described by the Lemma 2.2.

Next we go down with $l = h-1, ..., 2$ to use Lemma 2.1 and for $\delta^l = \Theta(e_l, a_l)$ satisfying $\mathsf{nr}\left(\delta^l\right) = l$ we produce

- the congruence $\delta^{l-1} = \Theta(e_{l-1}, a_{l-1})$ with $\mathsf{nr}\left(\delta^{l-1}\right) = l-1$ by putting $\delta^{l-1} = \left(\delta^{l}\right)^*$, $e_{l-1} = e_l^*$ and $a_{l-1} = a_l^*$,
- the k-ary polynomial $\mathbf{and}_k^{l-1}$ with the range $\{e_{l-1}, a_{l-1}\}$, so that

$$\mathbf{and}_k^{l-1}(x_1, \dots, x_k) = \begin{cases} a_{l-1}, & \text{if } x_1 = \dots = x_k = a_l, \\ e_{l-1}, & \text{otherwise,} \end{cases}$$

whenever $x_1, \dots, x_k \in \{e_l, a_l\}$.

To get the polynomial $\Phi\mathbf{Sat}$ we first compute $\mathbf{sat}_{\Phi_1}(\overline{x}), \dots, \mathbf{sat}_{\Phi_{m/k}}(\overline{x})$. Note that in fact in each of the $\mathbf{sat}_{\Phi_j}(\overline{x})$'s at most $3k$ variables (from $\overline{x}$) may occur, as the z_i^j's in different clauses do not have to be different. Next, to pass from level $l = h-1, \dots, 2$ to $l-1$ we group k^{l-1} values into k^{l-2} groups, each of which having k elements and apply the k-ary polynomial $\mathbf{and}_k^{l-1}$ to each of these groups to get k^{l-2} values on level $l-1$. Note that, due the properties of the ranges of the $\mathbf{sat}_{\Phi_j}$'s and of the $\mathbf{and}_k^l$'s, the only values that may occur at the l-th level (with $l = h-1, \dots, 1$) are e_l and a_l. Moreover the value a_l occurs only if the conjunction of k^{h-l} clauses that were used to compute this value is properly evaluated (to be satisfied). In particular arriving at level 1 we get one of the values e_1 or a_1 so that the resulting $3m$-ary polynomial $\Phi\mathbf{Sat}(x_1^1, x_2^1, x_3^1, \dots, x_1^m, x_2^m, x_3^m)$ satisfies

$$\Phi(z_1^1, z_2^1, z_3^1, \dots, z_1^m, z_2^m, z_3^m) = \top \quad \text{iff} \quad \Phi\mathbf{Sat}(x_1^1, x_2^1, x_3^1, \dots, x_1^m, x_2^m, x_3^m) = a_1,$$

whenever $z_1^1, z_2^1, z_3^1, \dots, z_1^m, z_2^m, z_3^m \in \{\top, \bot\}$ and $x_1^1, x_2^1, x_3^1, \dots, x_1^m, x_2^m, x_3^m \in A$ are such that $x_i^j \in A_{z_i^j}$. This means that the equation $\Phi\mathbf{Sat}(x_1^1, x_2^1, x_3^1, \dots, x_1^m, x_2^m, x_3^m) = a_1$ has a solution iff the formula Φ is satisfiable.

To conclude the proof we observe that Lemmas 2.1 and 2.2 guarantee that the polynomials of the form $\mathbf{sat}_{\Phi_j}$ and $\mathbf{and}_k^l$ have their size bounded by $2^{O(k)}$ and in fact they can be constructed in the very same amount of steps. Thus composing them to get $\Phi\mathbf{Sat}$ we need roughly $\left(2^{O(k)}\right)^h = 2^{O(m^{1/(h-1)})}$ steps, as promised. ◀

Proof of Corollary 1.3. As we have already mentioned in the Introduction a finite Malcev algebra admits only two kinds of a local behaviour. One of them is Boolean (or type **3** in the sense of tame congruence theory [9]). Modifying our argument used in Section 5 of [18] we show that the presence of type **3** leads to NP-completeness. Indeed, in this case the algebra has:

- two elements, say $0, 1$,
- an idempotent unary polynomial $\mathbf{e}_{01}(x)$ with range $\{0, 1\}$,
- two binary polynomials $\wedge, \vee$ and a unary polynomial $\neg$ that act on the set $\{0, 1\}$ like Boolean operation, i.e. meet, join and negation, respectively.

Let c be a constant bounding the sizes of all these four polynomials.

The presence of these polynomials allows us to translate each 3-CNF-SAT instance $\Phi \equiv \bigwedge_{i=1}^m \ell_1^i \vee \ell_2^i \vee \ell_3^i$, where $\ell_j^i \in \left\{z_i^j, \neg z_i^j\right\}$, into the equation of the algebra $\mathbf{A}$

$$\bigwedge_{j=1}^{m} \delta_1^j \mathbf{e}_{01}(x_1^j) \vee \delta_2^j \mathbf{e}_{01}(x_2^j) \vee \delta_3^j \mathbf{e}_{01}(x_3^j) = 1, \tag{1}$$

where

$$\delta_j^i \mathbf{e}_{01}(x_i^j) = \begin{cases} \mathbf{e}_{01}(x_i^j), & \text{if the literal } \ell_j^i \text{ is the variable, i.e., } \ell_j^i = z_i^j, \\ \neg \mathbf{e}_{01}(x_i^j), & \text{if } \ell_j^i \text{ is the negated variable, i.e., } \ell_j^i = \neg z_i^j. \end{cases}$$

It should be obvious that the formula Φ is satisfiable if and only if the equation (1) has a solution. However we need to take care of the size of this equation. But this can be secured by representing the m-ary conjunction in (1) in a balanced form, i.e. by a complete binary tree. This ensures us that the size of our equation is bounded by $O(c^{\log m})$, i.e. by a polynomial in m. This shows that in the presence of local Boolean behaviour in $\mathbf{A}$ the problem PolSat($\mathbf{A}$) is NP-complete, so that in view of ETH it cannot be in P.

Now we may assume that there is no local Boolean behaviour in $\mathbf{A}$, or in other words that the algebra $\mathbf{A}$ is solvable. In this case we simply refer to Theorem 1.2 to conclude that PolSat($\mathbf{A}$) $\in$ P implies $\mathsf{nr}(\mathbf{A}) \leqslant 2$, as claimed. ◀

By using the circuits constructed in Lemma 2.1 we can easily derive Theorem 1.1.

Proof of Theorem 1.1. The first part of the proof of Corollary 1.3 shows that the presence of Boolean local behaviour leads to NP-completeness of PolSat($\mathbf{A}$) and therefore also of CSat($\mathbf{A}$). Thus we may assume that $\mathbf{A}$ is solvable. In this case Lemma 2.2 supplies us with an element $a_{h-1} \in A$ so that each 3-CNF formula Φ can be turned, in a polynomial time, into a corresponding circuit $\mathbf{sat}_\Phi$ such that the equation $\mathbf{sat}_\Phi(\overline{x}) = a_{h-1}$ has a solution iff Φ is satisfiable.

This reduction obviously shows NP-completeness of CSat($\mathbf{A}$). ◀

3 Proofs of the Lemmas

For the proofs of Lemmas 2.1 and 2.2 we need an auxiliary Lemma. It uses the concept of the centralizer $(\alpha : \beta)$, that is the largest congruence θ satisfying $[\theta, \beta] \leqslant \alpha$.

▶ **Lemma 3.1.** *Let $\mathbf{A}$ be a finite solvable Malcev algebra and $\alpha \prec \beta$ be a covering pair of its congruences such that $\mathsf{nr}(\beta) > \mathsf{nr}(\alpha) > 0$. Then there is a join irreducible congruence γ with $\mathsf{nr}(\gamma) = \mathsf{nr}(\alpha)$ and $\alpha \leqslant (\gamma_- : \gamma)$ but $\beta \not\leqslant (\gamma_- : \gamma)$.*

Moreover for any pair $(e', a') \in \gamma - \gamma_-$ and $(e, a) \notin (\gamma_- : \gamma)$ there is a binary polynomial $\mathbf{s}_{ea}(x, y)$ of $\mathbf{A}$, satisfying

$$\begin{aligned} &\mathbf{s}_{ea}(e', y) = e', \qquad \text{for all } y \in A, \\ &\mathbf{s}_{ea}(a', e) \overset{\gamma_-}{\equiv} e', \\ &\mathbf{s}_{ea}(a', a) = a'. \end{aligned} \tag{2}$$

Proof. Let $k = \mathsf{nr}(\beta) > \mathsf{nr}(\alpha) = k - 1$, i.e. $\beta \not\leqslant \nu_{k-1} \geqslant \alpha$. Since $\mathsf{nr}\left(\beta^{(\omega)}\right) = \mathsf{nr}(\beta) - 1$ we know that $\beta^{(\omega)} \not\leqslant \nu_{k-2}$ so there is a congruence φ such that $\beta^{(\omega)} \cap \nu_{k-2} \leqslant \varphi \prec \beta^{(\omega)}$. Put $\rho_0 = \beta^{(\omega)}$ and $\rho_{i+1} = [\rho_i, \alpha]$ and observe that $\rho_i \leqslant \alpha^{(i)}$ whenever $i \geqslant 1$. Since $\alpha^{(\omega)} = \alpha^{(j)}$ holds for some j, we have $\rho_j \leqslant \alpha^{(j)} = \alpha^{(\omega)} \leqslant \beta^{(\omega)} \cap \nu_{k-2} \leqslant \varphi$. As $\rho_0 = \beta^{(\omega)} \not\leqslant \varphi$ then the minimal integer ℓ for which $\rho_\ell \leqslant \varphi$ is at least 1. Thus $\rho_{\ell-1} \vee \varphi = \beta^{(\omega)}$ so that in fact $\rho_1 = \left[\beta^{(\omega)}, \alpha\right] = [\rho_{\ell-1} \vee \varphi, \alpha] = [\rho_{\ell-1}, \alpha] \vee [\varphi, \alpha] = \rho_\ell \vee [\varphi, \alpha] \leqslant \varphi$. This shows $\alpha \leqslant \left(\varphi : \beta^{(\omega)}\right)$. Obviously $\beta \not\leqslant \left(\varphi : \beta^{(\omega)}\right)$ as $\left[\beta^{(\omega)}, \beta\right] = \beta^{(\omega)} \not\leqslant \varphi$. Now we pick a minimal congruence γ below $\beta^{(\omega)}$ but not below φ. Obviously γ is join irreducible and the covering pair $\varphi \prec \beta^{(\omega)}$ transposes down to $\gamma_- \prec \gamma$ where γ_- is the unique subcover of γ. Consequently $(\gamma_- : \gamma) = \left(\varphi : \beta^{(\omega)}\right)$, i. e. γ has the properties described in the Lemma. To calculate $\mathsf{nr}(\gamma)$ note first that $\gamma \leqslant \beta^{(\omega)} \leqslant \nu_{k-1}$ so that $\mathsf{nr}(\gamma) \leqslant k - 1$. But $\mathsf{nr}(\gamma) \leqslant k - 2$ would give $\gamma \leqslant \nu_{k-2} \cap \beta^{(\omega)} \leqslant \varphi$, contrary to our choice of γ.

For the second part of the Lemma we fix any pair $(e', a') \in \gamma - \gamma_-$. Due to the join irreducibility of γ we know that $\gamma = \Theta(e', a')$. Now if $(e, a) \notin (\gamma_- : \gamma)$ then $[\Theta(e, a), \Theta(e', a')] \not\leqslant \gamma_-$ so that Exercise 6.6 in [6] supplies us with a binary polynomial $\mathbf{s}(x, y)$ of $\mathbf{A}$ such that

$$\begin{aligned}&\mathbf{s}(e', e) \overset{\gamma_-}{\equiv} \mathbf{s}(a', e),\\ &\mathbf{s}(e', a) \overset{\gamma_-}{\not\equiv} \mathbf{s}(a', a).\end{aligned}$$

The very last line gives $\Theta(\mathbf{s}(e', a), \mathbf{s}(a', a)) = \gamma \ni (e', a')$ and therefore there is a unary polynomial $\mathbf{p}$ of $\mathbf{A}$ that takes the pair $(\mathbf{s}(e', a), \mathbf{s}(a', a))$ to (e', a'). Using Malcev polynomial $\mathbf{d}$ we modify $\mathbf{s}(x, y)$ to a new polynomial $\mathbf{s}_{ea}(x, y) = \mathbf{d}(\mathbf{ps}(x, y), \mathbf{ps}(e', y), e')$ for which it should be easy to check that

$$\begin{array}{lllllll} \mathbf{s}_{ea}(e', y) &=& \mathbf{d}(\mathbf{ps}(e', y), \mathbf{ps}(e', y), e') &=& e', & & \\ \mathbf{s}_{ea}(a', e) &=& \mathbf{d}(\mathbf{ps}(a', e), \mathbf{ps}(e', e), e') &\overset{\gamma_-}{\equiv}& \mathbf{d}(\mathbf{ps}(e', e), \mathbf{ps}(e', e), e') &=& e', \\ \mathbf{s}_{ea}(a', a) &=& \mathbf{d}(\mathbf{ps}(a', a), \mathbf{ps}(e', a), e') &=& \mathbf{d}(a', e', e') &=& a', \end{array}$$

so that all three conditions of (2) hold. ◀

With the help of Lemma 3.1 we are ready to prove Lemmas 2.1 and 2.2.

Proof of Lemma 2.1. We start our proof by referring to Lemma 3.1 with $(\alpha, \beta) = (\delta_-, \delta)$ to get a join irreducible congruence γ with $\delta \not\leqslant (\gamma_- : \gamma) \geqslant \delta_-$. We fix a (γ_-, γ)-minimal set V of $\mathbf{A}$ and a pair $(e', a') \in \gamma|_V - \gamma_-$. On the other hand $\delta = \Theta(e, a)$ ensures us that $(e, a) \notin (\gamma_- : \gamma)$ so that Lemma 3.1 supplies us with a polynomial $\mathbf{s}_{ea}(x, y)$ satisfying (2). Our first goal is to consecutively modify this polynomial $\mathbf{s}_{ea}$ to force the middle line in the display (2) to be the real equality instead of $\overset{\gamma_-}{\equiv}$. First we replace $\mathbf{s}_{ea}(x, y)$ with $\mathbf{e}_V\mathbf{s}_{ea}(\mathbf{e}_V(x), y)$, where $\mathbf{e}_V$ is a unary idempotent polynomial of $\mathbf{A}$ with range V. This new $\mathbf{s}_{ea}(x, y)$ not only satisfies (2) but also has its range contained in V and for any fixed y we have $\mathbf{s}_{ea}(A, y) = \mathbf{s}_{ea}(V, y)$.

Now, note that the first two lines of (2) tell us that the unary polynomial $\mathbf{s}^0(x) = \mathbf{s}_{ea}(x, e)$ does not permute V and consequently $\mathbf{s}^0(A) = \mathbf{s}^0(V) \subsetneq V$. Note also that for all $\varphi \prec \psi \leqslant \gamma$ we have $\mathbf{s}^0(\psi) \subseteq \varphi$, as otherwise the range of $\mathbf{s}^0$ would contain an (φ, ψ)-minimal set properly contained in V. This however (in view of Lemma 4.30 of [9]) cannot happen as V is a minimal set of type **2**. Now if a maximal chain of congruences strictly below γ has exactly l congruences then by replacing the polynomial $\mathbf{s}_{ea}(x, y)$ with $\mathbf{s}_{ea}(\dots\mathbf{s}_{ea}(\mathbf{s}_{ea}(x, y), y)\dots, y)$, where the iteration in the variable x is done l times, we keep the first and the third line of (2) to be true, while the middle one can be replaced by the equality. Thus we end up with a new polynomial $\mathbf{s}_{ea}(x, y)$ satisfying

$$\begin{aligned}&\mathbf{s}_{ea}(e', y) = e', \qquad \text{for all } y \in A,\\ &\mathbf{s}_{ea}(a', e) = e',\\ &\mathbf{s}_{ea}(a', a) = a'.\end{aligned} \tag{3}$$

Now the n-ary polynomial $\mathbf{and}^0_n(x_1, \dots, x_n) = \mathbf{s}_{ea}(\dots\mathbf{s}_{ea}(\mathbf{s}_{ea}(a', x_1), x_2)\dots, x_n)$ satisfies

$$\mathbf{and}^0_n(x_1, \dots, x_n) = \begin{cases} a', & \text{if } x_1 = \dots = x_n = a,\\ e', & \text{otherwise,}\end{cases} \tag{4}$$

whenever $x_1, \dots, x_n \in \{e, a\}$.

To conclude our argument note that $\gamma \not\leqslant \nu_{\mathsf{nr}(\delta_-)-1}$ and then simply pick a minimal congruence $\delta^\star$ below γ but not below $\nu_{\mathsf{nr}(\delta_-)-1}$. Obviously $\delta^\star$ is join irreducible (with the unique subcover $\delta^\star_-$) so that it is principal, say $\delta^\star = \Theta(e^\star, a^\star)$. Also, by minimality we have $\mathsf{nr}(\delta^\star) = \mathsf{nr}(\delta_-)$ and $\mathsf{nr}(\delta^\star_-) = \mathsf{nr}(\delta_-) - 1$.

Finally $(e^\star, a^\star) \in \delta^\star \subseteq \gamma = \Theta(e', a')$, so that there is a unary polynomial $\mathbf{p}$ of $\mathbf{A}$ that maps e' onto $e^\star$ and a' onto $a^\star$. By (4) it should be clear that the polynomial $\mathbf{and}_n(x_1, \dots, x_n) = \mathbf{p}(\mathbf{and}^0_n(x_1, \dots, x_n))$ does the job described in the Lemma.

To bound the length of the polynomial $\mathbf{and}_n$ and the size of the circuits that compute $\mathbf{and}_n$, first note that the polynomial $\mathbf{s}_{ea}$ can be realized by a circuit of a constant size (independent of n). Thus the entire circuit computing $\mathbf{and}^0_n$, and therefore of $\mathbf{and}_n$, can be constructed in a linear time $O(n)$. On the other hand unwinding this circuit to the tree (corresponding to the polynomial $\mathbf{and}_n$) enlarges the size exponentially and requires at most exponential time $2^{O(n)}$. ◀

The proof of Lemma 2.2 is slightly more involved.

Proof of Lemma 2.2. We start with observing that $\nu_{h-1} < 1_{\mathbf{A}}$ so that we can pick a cover $\beta \succ \nu_{h-1}$. This together with α set to ν_{h-1} allows us to use Lemma 3.1 to produce a join irreducible congruence γ with $\nu_{h-1} \leqslant (\gamma_- : \gamma) \not\geqslant \beta$. In particular we know that $(\gamma_- : \gamma) \neq 1_{\mathbf{A}}$.

By Lemma 3.1 we know that $\mathsf{nr}(\gamma) = \mathsf{nr}(\nu_{h-1}) = h - 1$ so that $\gamma \not\leqslant \nu_{h-2}$ and we can pick δ^{h-1} to be a minimal congruence below γ but not below ν_{h-2}. Obviously δ^{h-1} is join irreducible so that it is a principal congruence, say $\delta^{h-1} = \Theta(e_{h-1}, a_{h-1})$. By the very same token we know that γ is principal, but here we need to choose its generating pair more carefully. First we fix a (γ_-, γ)-minimal set V and then a pair $(e', a') \in \gamma|_V - \gamma_-$ to have $\gamma = \Theta(e', a')$. The (γ_-, γ)-trace of V containing both e' and a' is denoted by N. We know that the induced algebra $(\mathbf{A}|_N)/\gamma_-$ is polynomially equivalent to a (one dimensional) vector space and we may assume that e'/γ_- is its zero element with respect to the vectors addition $+$ which has to be a polynomial of $\mathbf{A}$. Since $\Theta(e_{h-1}, a_{h-1}) = \delta^{h-1} \leqslant \gamma = \Theta(e', a')$ we can pick a unary polynomial $\mathbf{p}$ of $\mathbf{A}$ that maps e' to e_{h-1} and a' to a_{h-1}. This polynomial is going to be used at the end of the proof.

Now we put $\tau = (\gamma_- : \gamma)$ and choose a transversal $\{d_0, d_1, \dots, d_r\}$ of A/τ. If $i \neq j$ then $(d_i, d_j) \notin (\gamma_- : \gamma)$ and Lemma 3.1 gives us a binary polynomial $\mathbf{s}_{ij} = \mathbf{s}_{d_i, d_j}$ satisfying

$$\begin{aligned} &\mathbf{s}_{ij}(e', y) = e', && \text{for all } y \in A,\\ &\mathbf{s}_{ij}(a', d_i) \overset{\gamma_-}{\equiv} e',\\ &\mathbf{s}_{ij}(a', d_j) = a'. \end{aligned} \tag{5}$$

As in the proof of Lemma 2.1 we replace $\mathbf{s}_{ij}(x, y)$ by $\mathbf{e}_V \mathbf{s}_{ij}(\mathbf{e}_V(x), y)$, where $\mathbf{e}_V$ is the unary idempotent polynomial of $\mathbf{A}$ with the range V. Obviously the properties in the display (5) hold for this new $\mathbf{s}_{ij}$, but this new polynomial has the range contained in V and for any fixed $y \in A$ the mapping $V \ni v \longmapsto \mathbf{s}_{ij}(v, y) \in V$ is either a permutation of V or collapses $\gamma|_V$ to γ_-, i.e. it is constant modulo γ_- on $\gamma|_V$-classes. Thus, iterating $\mathbf{s}_{ij}(v, y)$ in the first variable a sufficient number of times, we can modify $\mathbf{s}_{ij}$ to additionally have that (for each fixed $y \in A$) the new polynomial $\mathbf{s}_{ij}(v, y)$ is either the identity map on V or it is constant modulo γ_- on $\gamma|_V$-classes. Actually, in the second case, i.e. if $\mathbf{s}_{ij}(v, y)$ collapses $\gamma|_V$ to γ_-, it collapses the entire trace N to $\mathbf{s}_{ij}(e', y)/\gamma_- = e'/\gamma_-$. Summing up, we produced polynomials $\mathbf{s}_{ij}$ satisfying

$$\begin{aligned} \mathbf{s}_{ij}(e', y) &= e', && \text{for each } y \in A,\\ \mathbf{s}_{ij}(v, d_i) &\overset{\gamma_-}{\equiv} e', && \text{for each } v \in N,\\ \mathbf{s}_{ij}(v, d_j) &= v, && \text{for each } v \in V. \end{aligned}$$

Now, using the fact that $[\gamma,\tau] \leqslant \gamma_-$ we can keep the above equalities modulo γ_- by varying the second variable modulo τ:

$$\begin{array}{rcll} \mathbf{s}_{ij}(e',y) & = & e', & \text{for each } \ y \in A, \\ \mathbf{s}_{ij}(v,y) & \overset{\gamma_-}{\equiv} & e', & \text{for each } \ v \in N \text{ and } \ y \in d_i/\tau, \\ \mathbf{s}_{ij}(v,y) & \overset{\gamma_-}{\equiv} & v, & \text{for each } \ v \in V \text{ and } \ y \in d_j/\tau. \end{array} \tag{6}$$

Now for each $j = 0,\dots,r$ define

$$\mathbf{s}_j(x,y) = \mathbf{s}_{i_1 j}(\dots \mathbf{s}_{i_{r-1} j}(\mathbf{s}_{i_r j}(x,y),y)\dots,y),$$

where $\{j,i_1,\dots,i_r\} = \{0,1,\dots,r\}$. Obviously $\mathbf{s}_j$ has the range contained in V and

$$\begin{array}{rcll} \mathbf{s}_j(e',y) & = & e', & \text{for each } \ y \in A, \\ \mathbf{s}_j(v,y) & \overset{\gamma_-}{\equiv} & e', & \text{for each } \ v \in N \text{ and } \ y \in A - d_j/\tau, \\ \mathbf{s}_j(v,y) & \overset{\gamma_-}{\equiv} & v, & \text{for each } \ v \in V \text{ and } \ y \in d_j/\tau. \end{array} \tag{7}$$

Indeed, the first and the last item follow directly from the definition of $\mathbf{s}_j$. To see the middle one, note that for $v \in N$ and $y \in d_{i_\ell}/\tau$, defining $v' = \mathbf{s}_{i_{\ell+1}}(\dots \mathbf{s}_{i_{r-1} j}(\mathbf{s}_{i_r j}(v,y),y)\dots,y)$ we have

$$v' \overset{\gamma}{\equiv} \mathbf{s}_{i_{\ell+1} j}(\dots \mathbf{s}_{i_{r-1} j}(\mathbf{s}_{i_r j}(e',y),y)\dots,y) = e',$$

i.e. $v' \in N$ so that $\mathbf{s}_{i_\ell j}(v',y) \overset{\gamma_-}{\equiv} e'$, and consequently

$$\mathbf{s}_j(v,y) = \mathbf{s}_{i_1 j}(\dots \mathbf{s}_{i_{\ell-1} j}(\mathbf{s}_{i_\ell j}(v',y),y)\dots,y) \overset{\gamma_-}{\equiv} \mathbf{s}_{i_1 j}(\dots \mathbf{s}_{i_{\ell-1} j}(e',y)\dots,y) = e'.$$

Recall that $(\mathbf{A}|_N)/\gamma_-$ is polynomially equivalent to a vector space in which e'/γ_- serves as a zero element, while the addition is defined by $x + y = \mathbf{d}(x,e',y)$ and $x - y = \mathbf{d}(x,y,e')$. Obviously this addition does not behave so nice outside the trace N and before factoring out by γ_- but since for $v \in N$ (and arbitrary $y \in A$) the elements $\mathbf{s}_j(v,y)$ are in $\mathbf{A}|_N$, it makes sense to sum them up and define

$$\mathbf{s}_\top(x,y) = \sum_{j=1}^{r} \mathbf{s}_j(x,y)$$

(by associating the "summands" to the left) to observe that

$$\begin{array}{rcll} \mathbf{s}_\top(e',y) & = & e', & \text{for each } \ y \in A, \\ \mathbf{s}_\top(v,y) & \overset{\gamma_-}{\equiv} & e', & \text{for each } \ v \in N \text{ and } \ y \in d_0/\tau, \\ \mathbf{s}_\top(v,y) & \overset{\gamma_-}{\equiv} & v, & \text{for each } \ v \in N \text{ and } \ y \in A - d_0/\tau. \end{array} \tag{8}$$

To see (8) first note that in the sum defining $\mathbf{s}_\top$ for $v \in N$, at most one summand lies outside of e'/γ_-, namely $\mathbf{s}_j(v,y)$ with $j \neq 0$ for which $y \in d_j/\tau$. This obviously gives the last two lines in (8) as well as $\mathbf{s}_\top(e',y) \overset{\gamma_-}{\equiv} e'$. To replace this by the equality note that due to the first line in (7) all summands in $\mathbf{s}_\top(e',y)$ are equal to e' so that "summing" them up with the help of Malcev polynomial $\mathbf{d}$ returns e'.

Now we put $\mathbf{s}_\bot(x,y) = \mathbf{s}_0(x,y)$ and observe that for any fixed $v \in N$ both $\mathbf{s}_\top(v,y)$ and $\mathbf{s}_\bot(v,y)$ have their ranges contained in $e'/\gamma_- \cup v/\gamma_-$. Moreover $\mathbf{s}_\top$ and $\mathbf{s}_\bot$ are complementary in the sense that $\mathbf{s}_\top(v,y) + \mathbf{s}_\bot(v,y) = v$. In fact they switch their values (between v and e') depending on whether y is in $A_\bot = d_0/\tau$ or in its complement $A_\top = A - A_\bot = \bigcup_{i=1}^{r} d_i/\tau$.

Now, for $\bar{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3) \in \{\top, \bot\}^3$ we define the polynomial

$$\begin{aligned}\breve{\mathbf{s}}_{\bar{\varepsilon}}(v, y_1, y_2, y_3) = {} & \mathbf{s}_{\varepsilon_3}(\mathbf{s}_{\varepsilon_2}(\mathbf{s}_{\varepsilon_1}(v, y_1), y_2), y_3) + \mathbf{s}_{\varepsilon_1}(v, y_1) + \mathbf{s}_{\varepsilon_2}(v, y_2) + \mathbf{s}_{\varepsilon_3}(v, y_3) \\ & - \mathbf{s}_{\varepsilon_2}(\mathbf{s}_{\varepsilon_1}(v, y_1), y_2) - \mathbf{s}_{\varepsilon_3}(\mathbf{s}_{\varepsilon_1}(v, y_1), y_3) - \mathbf{s}_{\varepsilon_2}(\mathbf{s}_{\varepsilon_3}(v, y_2), y_3),\end{aligned}$$

Using (7) and (8) one can calculate that

$$\begin{array}{lcll} \breve{\mathbf{s}}_{\bar{\varepsilon}}(e', y_1, y_2, y_3) & = & e', & \\ \breve{\mathbf{s}}_{\bar{\varepsilon}}(v, y_1, y_2, y_3) & \overset{\gamma_-}{\equiv} & e', & \text{if } y_i \notin A_{\varepsilon_i} \text{ for } i \in \{1,2,3\}, \\ \breve{\mathbf{s}}_{\bar{\varepsilon}}(v, y_1, y_2, y_3) & \overset{\gamma_-}{\equiv} & v, & \text{if } y_i \in A_{\varepsilon_i} \text{ for some } i, \end{array} \tag{9}$$

due to the fact that, modulo γ_-, each of the seven summands in $\breve{\mathbf{s}}_{\bar{\varepsilon}}(v, y_1, y_2, y_3)$ is either e', or v, or $-v$.
Arguing like in the proof of Lemma 2.1, by going down from γ_- to 0 through a maximal chain of congruences (and iterating $\breve{\mathbf{s}}_{\bar{\varepsilon}}$ in the first variable), we improve (9) to get

$$\begin{array}{lcll} \breve{\mathbf{s}}_{\bar{\varepsilon}}(e', y_1, y_2, y_3) & = & e', & \\ \breve{\mathbf{s}}_{\bar{\varepsilon}}(v, y_1, y_2, y_3) & = & e', & \text{if } y_i \notin A_{\varepsilon_i} \text{ for every } i \in \{1,2,3\}, \\ \breve{\mathbf{s}}_{\bar{\varepsilon}}(v, y_1, y_2, y_3) & = & v, & \text{if } y_i \in A_{\varepsilon_i} \text{ for some } i. \end{array} \tag{10}$$

With the help of the polynomials $\breve{\mathbf{s}}_{\bar{\varepsilon}}$ (which have been invented to code clauses) we can create the polynomial $\mathbf{sat}_\Phi$ for a given 3-CNF formula $\Phi \equiv \bigwedge_{j=1}^m \ell_1^j \vee \ell_2^j \vee \ell_3^j$, where $\ell_i^j \in \left\{z_i^j, \neg z_i^j\right\}$, by first putting

$$\varepsilon_i^j = \begin{cases} \top, & \text{if the literal } \ell_i^j \text{ is the variable, i.e., } \ell_i^j = z_i^j, \\ \bot, & \text{if the literal } \ell_i^j \text{ is the negated variable, i.e., } \ell_i^j = \neg z_i^j, \end{cases}$$

then $\bar{\varepsilon}^j = (\varepsilon_1^j, \varepsilon_2^j, \varepsilon_3^j)$ and finally

$$\mathbf{sat}_\Phi(x_1^1, x_2^1, x_3^1, \dots, x_1^m, x_2^m, x_3^m) = \mathbf{p}(\breve{\mathbf{s}}_{\bar{\varepsilon}^m}(\dots \breve{\mathbf{s}}_{\bar{\varepsilon}^1}(a', x_1^1, x_2^1, x_3^1), \dots x_1^m, x_2^m, x_3^m)).$$

By (10) it should be clear that for any evaluation of the z_i^j's in $\{\top, \bot\}$ and x_i^j's in A so that $x_i^j \in A_{z_i^j}$ we have

$$\Phi(z_1^1, z_2^1, z_3^1, \dots, z_1^m, z_2^m, z_3^m) = \top \quad \text{iff} \quad \mathbf{sat}_\Phi(x_1^1, x_2^1, x_3^1, \dots, x_1^m, x_2^m, x_3^m) = a_{h-1}.$$

as required by the Lemma.

The complexity arguments simply repeat those from the proof of Lemma 2.1 to bound the time needed to construct the polynomial $\mathbf{sat}_\Phi$ by $2^{O(m)}$, while to construct the circuit computing $\mathbf{sat}_\Phi$ by $O(m)$. ◀

4 Conclusions and Open Problems

Since Theorem 1.1 that improves the result from [18], one can hope for a classification of finite algebras from congruence modular varieties that have tractable CSAT. Indeed, by [18, Corollary 6.5], such an algebra has to decompose into a direct product $\mathbf{S} \times \mathbf{L}$ of a solvable algebra $\mathbf{S}$ and an algebra $\mathbf{L}$ that behaves pretty similarly to a lattice (at least locally). Now, our Theorem 1.1 forces $\mathbf{S}$ to be nilpotent without actually assuming (like it has been done in [18]) that CSAT is tractable for all quotients of $\mathbf{S}$. The paper [18] also enforces that the algebra $\mathbf{L}$ has to behave not only like a lattice, but in fact like a distributive lattice, provided CSAT is tractable for all quotients of $\mathbf{L}$. The natural problem is to eliminate this strong assumption about quotients also from the $\mathbf{L}$ side, i.e. for algebras from congruence distributive varieties. Thus, we are left with the following

▶ **Question 1.** *Does tractability of* CSat *for a finite algebra* $\mathbf{L}$ *from a congruence distributive variety implies tractability of* CSat *for all quotients of* $\mathbf{L}$*?*

As we have already mentioned in the Introduction the classification of finite algebras with tractable CSat is not fully done on the solvable side. Now, by Theorem 1.1, we know that such an algebra $\mathbf{S}$ has to be nilpotent. Moreover, due to the work of Kompatscher [23], we know that in fact $\mathsf{sr}(\mathbf{A}) \leqslant 2$. Unfortunately this bound does not suffice to have tractable CSat($\mathbf{S}$), as it has been shown by some examples described in [17]. On the other hand one cannot hope to strengthen this bound and force $\mathbf{S}$ to be supernilpotent (in which case [18] gives a polynomial time algorithm for CSat). This in turn has been witnessed by our examples provided in [15]. Thus we are left with the following

▶ **Problem 2.** *Characterize finite nilpotent Malcev algebras of supernilpotent rank at most* 2 *with tractable* CSat*.*

A very similar, and somehow connected, problem CEqv($\mathbf{A}$) of circuit equivalence has also been considered in [18]. This time we ask if two circuits over $\mathbf{A}$ compute the same function. In this case, if a finite algebra $\mathbf{A}$ is taken from a congruence modular variety then [18] shows that tractability of CEqv($\mathbf{A}$) implies that $\mathbf{A}$ is solvable. Note that there is no lattice-like part here, as the co-NP-completeness of CEqv even over the 2-element lattice eliminates this part. Arguing like in the proof of Theorem 1.1 we can force nilpotency of algebras with tractable CEqv. Again, Kompatscher [23] forces such algebras to have supernilpotent rank bounded by 2. Suprisingly, here we do not have any single example of a nilpotent algebra $\mathbf{A}$ with $\mathsf{sr}(\mathbf{A}) = 2$ and intractable CEqv($\mathbf{A}$). In fact among the examples in [17] of algebras with $\mathsf{sr}(\mathbf{A}) = 2$ but intractable CSat($\mathbf{A}$) there are 2-nilpotent algebras. However, [19] shows that for 2-nilpotent algebras the problem CEqv is tractable. Therefore the answer to the next Problem differs from the answer to Problem 2.

▶ **Problem 3.** *Characterize finite nilpotent Malcev algebras of supernilpotent rank at most* 2 *with tractable* CEqv*.*

In our opinion this difference in the complexity for CSat and CEqv may result in a search for completely new techniques. Nevertheless, note that both CSat and CEqv behave the same on supernilpotent algebras, as CEqv for such algebras has been shown to be tractable in [1].

Now we switch to the problem PolSat. For a fixed algebra its complexity is not bigger than this of CSat. However, as there are examples (e.g. the already mentioned groups $\mathbf{S}_3$ or $\mathbf{A}_4$) where CSat is essentially harder. This is because using (in CSat) additional polynomials we compress (in our opinion artificially inflated) the input of PolSat. For a better understanding of this phenomenon among Malcev algebras first note that, due to Theorem 1.2, PolSat($\mathbf{A}$) $\in \mathsf{P} \not\ni$ CSat($\mathbf{A}$) may happen only if $\mathsf{nr}(\mathbf{A}) \leqslant 2$. Among algebras of nilpotent rank 1, i.e. among nilpotent algebras, all known cases (i.e. both tractable and intractable examples, as well as some more general results, like for supernilpotent algebras) enjoy the same complexity for both PolSat and CSat. This leads to the following:

▶ **Question 4.** *Does there exist a finite nilpotent Malcev algebra with tractable* PolSat *and intractable* CSat*?*

A careful reading of the proof of Corollary 6.5 in [18] shows that the earlier described decomposition $\mathbf{A} = \mathbf{S} \times \mathbf{L}$ requires in fact a weaker assumption that only PolSat(A) (but not necessarily CSat) is not NP-complete. Note that a positive answer to Question 1 can possibly be carried out to PolSat. Thus we believe that the next question, similar to Question 4, has a negative answer.

▶ **Question 5.** *Does there exist a finite algebra from a congruence distributive variety with tractable* PolSat *and intractable* CSat*?*

Summing up we expect that the following conjecture holds.

▶ **Conjecture 6.** *The only examples of finite algebras (from a congruence modular variety) that separate complexity of* PolSat *and* CSat *have to be solvable and of nilpotent rank* 2*.*

When considering complexity of i.e. the polynomial equivalence problem, versus this of CEqv we can repeat Questions 4 and Conjecture 6. An analogous modification of Question 5 is already answered, as both PolEqv and CEqv are co-NP-complete for nontrivial algebras in this realm.

References

1 Erhard Aichinger and Nebojša Mudrinski. Some applications of higher commutators in Mal'cev algebras. *Algebra Universalis*, 63(4):367–403, 2010. doi:10.1007/s00012-010-0084-1.

2 David Mix Barrington, Pierre McKenzie, Cris Moore, Pascal Tesson, and Denis Thérien. Equation satisfiability and program satisfiability for finite monoids. In *25th International Symposium on Mathematical Foundations of Computer Science (MFCS 2000)*, pages 172–181. Springer Berlin Heidelberg, 2000. doi:10.1007/3-540-44612-5_13.

3 Stanley Burris and Hanamantagouda Sankappanavar. *A Course in Universal Algebra With 36 Illustrations*. Springer-Verlag, New York, 1981.

4 Stephen Cook. The complexity of theorem proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing (STOC 1971)*, pages 151–158, 1971. doi:10.1145/800157.805047.

5 Attila Földvári and Gábor Horváth. The complexity of the equation solvability and equivalence problems over finite groups. *International Journal of Algebra and Computation*, 30(03):1–17, 2019. doi:10.1142/S0218196720500137.

6 Ralph Freese and Ralph McKenzie. *Commutator Theory For Congruence Modular Varieties*. London Mathematical Society Lecture Notes, No. 125. Cambridge University Press, 1987.

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

8 Tomasz Gorazd and Jacek Krzaczkowski. Term equation satisfiability over finite algebras. *International Journal of Algebra and Computation*, 20(08):1001–1020, 2010. doi:10.1142/S021819671000600X.

9 David Hobby and Ralph McKenzie. *Structure of Finite Algebras*. Contemporary Mathematics vol. 76. American Mathematical Society, 1988. doi:10.1090/conm/076.

10 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.

11 Gábor Horváth. The complexity of the equivalence and equation solvability problems over meta-abelian groups. *Journal of Algebra*, 433:208–230, 2015. doi:10.1016/j.jalgebra.2015.03.015.

12 Gábor Horváth and Csaba Szabó. The complexity of checking identities over finite groups. *International Journal of Algebra and Computation*, 16(05):931–940, 2006. doi:10.1142/S0218196706003256.

13 Gábor Horváth and Csaba Szabó. Equivalence and equation solvability problems for the alternating group A_4. *Journal of Pure and Applied Algebra*, 216(10):2170–2176, 2012. doi:10.1016/j.jpaa.2012.02.007.

14 Paweł Idziak, Piotr Kawałek, Jacek Krzaczkowski, and Armin Weiß. Equation satisfiability in solvable groups. To appear in *Theory of Computing Systems*, 2021. arXiv:2010.11788.

15 Paweł Idziak, Piotr Kawałek, and Jacek Krzaczkowski. Expressive power, satisfiability and equivalence of circuits over nilpotent algebras. In *43rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2018)*, volume 117, pages 17:1–17:15. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.MFCS.2018.17.

16 Paweł 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 2020)*, pages 578–590, 2020. doi:10.1145/3373718.3394780.

17 Paweł Idziak, Piotr Kawałek, and Jacek Krzaczkowski. Complexity of modular circuits, 2021. arXiv:2106.02947.

18 Paweł Idziak and Jacek Krzaczkowski. Satisfiability in multi-valued circuits. In *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2018)*, pages 550–558, 2018. Full version: to appear in SIAM Journal on Computing (see also 1710.08163). doi:10.1145/3209108.3209173.

19 Piotr Kawałek, Michael Kompatscher, and Jacek Krzaczkowski. Circuit equivalence in 2-nilpotent algebras, 2019. arXiv:1909.12256.

20 Keith Kearnes. Congruence modular varieties with small free spectra. *Algebra Universalis*, 42(3):165–181, 1999. doi:10.1007/s000120050132.

21 Ondrej Klíma. Complexity issues of checking identities in finite monoids. *Semigroup Forum*, 79(3):435–444, 2009. doi:10.1007/s00233-009-9180-y.

22 Michael Kompatscher. The equation solvability problem over supernilpotent algebras with mal'cev term. *International Journal of Algebra and Computation*, 28(06):1005–1015, 2018. doi:10.1142/S0218196718500443.

23 Michael Kompatscher. CSAT and CEQV for nilpotent Maltsev algebras of fitting length > 2, 2021. arXiv:2105.00689.

24 Yuri Matiyasevich. Enumerable sets are diophantine. In *Soviet Mathematics Doklady*, volume 11, pages 354–358, 1970.

25 Bernhard Schwarz. The complexity of satisfiability problems over finite lattices. In *Annual Symposium on Theoretical Aspects of Computer Science (STACS 2004)*, pages 31–43. Springer Berlin Heidelberg, 2004. doi:10.1007/978-3-540-24749-4_4.

26 Armin Weiß. Hardness of equations over finite solvable groups under the exponential time hypothesis. In *Proceedings of 47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, 2020. doi:10.4230/LIPIcs.ICALP.2020.102.