



Libor Barto

Finitely related algebras in congruence modular varieties have few subpowers

Received May 22, 2015

Abstract. We show that every finite algebra which is finitely related and lies in a congruence modular variety has few subpowers. This result, combined with other theorems, has interesting consequences for the complexity of several computational problems associated to finite relational structures: the constraint satisfaction problem, the primitive positive formula comparison problem, and the learnability problem for primitive positive formulas. Another corollary is that it is decidable whether an algebra given by a set of relations has few subpowers.

Keywords. Finitely related algebra, congruence modular variety, Gumm terms, few subpowers, cube terms

1. Introduction

By an *algebra* \mathbf{A} we mean a nonempty set, called the *universe* of \mathbf{A} , together with a set (possibly indexed) of finitary operations on it, called the *basic operations* of \mathbf{A} . An algebra is finite if its universe is finite. The main result of this paper shows that two important properties of finite algebras are equivalent under a certain additional finiteness assumption.

Finitely related algebras

We first discuss the required finiteness condition.

Most interesting properties of an algebra, like its subuniverses, congruences, automorphisms, etc., depend on its term operations rather than on the particular choice of basic operations. Therefore, the *clone* of an algebra, i.e., the set of all its term operations, is an invariant of the algebra which is sufficient for most purposes. Every clone \mathcal{C} on a finite set A is determined by a set \mathcal{R} of relations on A in the sense that \mathcal{C} is equal to the set of all operations which are compatible with every relation in \mathcal{R} [15, 29] (see Section 2 for definitions). If \mathcal{R} can be chosen finite, we call \mathcal{C} *finitely related*. We call a finite algebra *finitely related* if its clone is. Informally, an algebra is finitely related if it has a finite description in terms of relations. Finitely related algebras have also been called *predicately describable*, *finitely definable*, *of finite relational degree*, or *of finite degree*.

L. Barto: Department of Algebra, Faculty of Mathematics and Physics, Charles University in Prague, Sokolovská 83, 18675 Prague 8, Czech Republic; e-mail: libor.barto@gmail.com

Mathematics Subject Classification (2010): Primary 08B10; Secondary 08A70

A classical result of Baker and Pixley [2] implies that every finite algebra with a *near-unanimity* term operation (that is, an operation t of arity at least 3 such that

$$t(a, \dots, a, b, a, \dots, a) = a$$

for any a, b and any position of b) is finitely related. This includes all finite lattices and their *expansions*, i.e., algebras obtained from lattices by adding operations. A recent result of Aichinger, Mayr, and McKenzie [1] is another source of examples: it implies that all finite quasigroups and their expansions (like finite groups, rings and modules) are finitely related. There are both finitely related and non-finitely related (expansions of) semigroups [24, 25, 40]. The simplest example of a non-finitely related algebra is the two-element *implication algebra* $((0, 1); \{\rightarrow\})$, where \rightarrow is the implication regarded as a binary operation (see [24] for an elementary proof).

Finitely related algebras play an important role in several computational problems parametrized by finite sets of relations, including the fixed-template constraint satisfaction problem (CSP).

Congruence modularity and few subpowers

A *variety* is a class of algebras of the same signature defined by a set of identities, where an *identity* is a universally quantified equation. We say that a variety V is *congruence modular* if the congruence lattice of every algebra in V is modular. We denote by CM the class of all algebras which are members of a congruence modular variety.

Congruence modular varieties (and algebras therein) are among the most studied objects in universal algebra. On the one hand, they are in a sense manageable since strong results are applicable to all algebras in CM, in particular, the commutator theory [27]. On the other hand, the class CM contains most of the classical types of algebras: groups, rings, modules, lattices (but not all semigroups or semilattices). In fact, the above mentioned examples belong to one of two important subclasses of CM: CD and CP. An algebra is in CD [CP] if it belongs to a congruence distributive [permutable] variety, where a variety V is *congruence distributive* [*congruence permutable*] if the congruence lattice of every algebra in V is distributive [congruences of each algebra in V permute with respect to the relational product]. Examples of algebras in CD include all algebras with a near-unanimity term operation (e.g., lattices) and also the two-element implication algebra. Expansions of quasigroups are in the class CP.

The property of having few subpowers is of much more recent origin. A finite algebra \mathbf{A} has *few subpowers* if, for some polynomial p , the number of subalgebras of \mathbf{A}^n is less than $2^{p(n)}$. This notion was introduced by Berman, Idziak, Marković, McKenzie, Valeriote, and Willard [13] (building on earlier work by Bulatov, Chen and Dalmau [18, 19, 23]). A number of characterizations were given in this influential paper, among them the following: \mathbf{A} has few subpowers iff, for some polynomial q , each subalgebra of \mathbf{A}^n has a generating set of size $q(n)$. Such “compact” representations of subalgebras of powers were used to devise a polynomial algorithm for solving a large class of CSPs [32]. Another remarkable result based on compact representations is that each finite algebra with few

subpowers is finitely related [1]. (In fact, by [37], a finite idempotent algebra \mathbf{A} has few subpowers iff each expansion of \mathbf{A} is finitely related.)

Interestingly, an equivalent concept to having few subpowers was independently found by Kearnes and Szendrei [35] in connection with the study of relational clones.

The class FS of algebras with few subpowers contains the classes CP and NU (the class of algebras with a near unanimity term operation), and is contained in CM [13] as shown in Figure 1, left (note that our definition of the few subpowers property applies only to finite algebras, but some of the equivalent characterizations, as item (iii) in Theorem 2.8, make sense in general, and the aforementioned inclusions are true for general algebras with such alternative definitions).

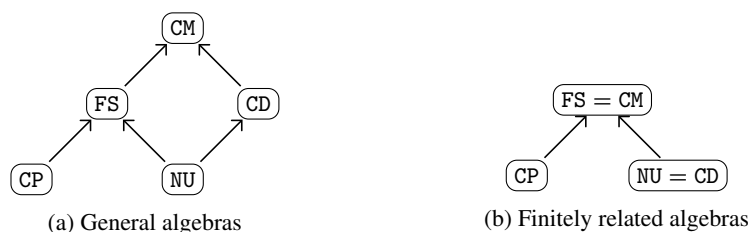


Fig. 1. Comparison of some classes of algebras; arrows depict inclusions.

Main result

The main result of the paper affirmatively answers the Edinburgh conjecture [17], also known as the Valeriote conjecture.

Theorem 1.1. *If \mathbf{A} is a finitely related finite algebra in CM, then \mathbf{A} has few subpowers.*

The proof of Theorem 1.1 covers Sections 3 and 4. This theorem generalizes the main result of [3] that every finitely related finite algebra in CD has a near-unanimity term operation (known as the Zádori conjecture). Indeed, if an algebra has few subpowers and is in CD, then it necessarily has a near-unanimity term operation by [13, 38]. In Figure 1 (right), the classes NU, CP, FS, CD, CM are compared within the class of finitely related algebras.

A combination of Theorem 1.1 with the already mentioned results gives:

Corollary 1.2. *The following are equivalent for a finite algebra \mathbf{A} :*

- (i) \mathbf{A} is finitely related and is in the class CM.
- (ii) \mathbf{A} has few subpowers.

Theorem 1.1 can be viewed from several perspectives. One way to read it is that a finite algebra \mathbf{A} in CM can be approximated by algebras with few subpowers. Indeed, if $\mathcal{R} = \{R_1, R_2, \dots\}$ is a set of relations that determines the clone \mathcal{C} of the algebra \mathbf{A} , and \mathcal{C}_i is the clone determined by $\{R_1, \dots, R_i\}$, then \mathcal{C} is equal to the intersection of the descending chain $\mathcal{C}_1 \supseteq \mathcal{C}_2 \supseteq \dots$ and each \mathcal{C}_i in the chain has few subpowers (this follows from Theorem 1.1 since the class CM is closed under expansions).

Another interpretation of Theorem 1.1 is that it gives a nontrivial implication between two Maltsev conditions under a finiteness assumption. A *Maltsev condition* is, roughly, a condition stipulating the existence of terms satisfying certain identities. Many classes of algebras (e.g., CP, CD, CM, NU, FS) can be characterized by Maltsev conditions and the corresponding terms serve as a useful technical tool for studying these classes.

Recent development has added numerous interesting implications between Maltsev conditions on four finiteness levels: finitely related algebras, finite algebras, algebras in locally finite varieties, and general algebras. Theorem 1.1 can be formulated as “Gumm terms \Rightarrow cube term” for finitely related algebras, and its consequence from [3] as “Jónsson terms \Rightarrow near-unanimity term”. For finite algebras we have, for instance, “Taylor term \Rightarrow cyclic term” [8]. Notable recent examples for algebras in locally finite varieties include “Taylor terms \Rightarrow weak near-unanimity term” [39], “Taylor term \Rightarrow Siggers term” [41, 34]. The implications “Jónsson/Gumm terms \Rightarrow directed Jónsson/Gumm terms”, which was very recently shown to hold for general algebras [33], serves as a useful tool in this paper. See also [36] for further interesting implications.

Finally, Theorem 1.1 can be regarded as a source of examples for non-finitely related algebras. Indeed, every finite algebra in CM which does not have few subpowers is non-finitely related.

Consequences

Theorem 1.1, combined with other results, has interesting consequences for several computational problems associated to a finite relational structure \mathbb{A} . Theorem 1.1 together with [32] implies that the CSP over \mathbb{A} is solvable in polynomial time whenever the corresponding algebra \mathbf{A} is in CM; together with [16, 17], it gives a P/coNP-complete dichotomy for comparison of primitive positive formulas over \mathbb{A} ; together with [32, 22], it classifies learnability of the relation defined by a primitive positive formula over \mathbb{A} . The role of Theorem 1.1 in the last two results is to show that there is no gap between positive (tractability) results for the case that \mathbf{A} has few subpowers and negative (hardness) results for the case that \mathbf{A} is not in CM. The main result also implies that it is decidable (given \mathbb{A} on input) whether \mathbf{A} has few subpowers. The consequences and open problems are discussed in Section 5.

2. Preliminaries

In this section we collect necessary definitions and results. We refer to [12, 21, 42] for the basics of universal algebra, and to [43] for graph theory. Throughout the paper, we use the notation

$$[n] = \{1, \dots, n\}.$$

2.1. Algebras and varieties

An n -ary operation on a set A is a mapping $f : A^n \rightarrow A$. We only consider finitary operations, i.e., n is a natural number. For subsets $A_1, \dots, A_n \subseteq A$, we write

$$f(A_1, \dots, A_n) = \{f(a_1, \dots, a_n) : a_1 \in A_1, \dots, a_n \in A_n\}.$$

An operation is *idempotent* if it satisfies the identity $f(a, \dots, a) = a$.

An *algebra* is a pair $\mathbf{A} = (A; \mathcal{F})$, where A is a nonempty set, called the *universe* of \mathbf{A} , and \mathcal{F} is a set (possibly indexed) of operations on A , called the *basic operations* of \mathbf{A} . We use a boldface letter to denote an algebra and the same letter in the plain type to denote its universe. An algebra is *idempotent* if all of its operations are idempotent. Two algebras have the same *signature* if their operations are indexed by the same set and corresponding operations have the same arities.

A *clone* on A is a set of operations on A which contains the projections (the operations π_i^n defined by $\pi_i^n(a_1, \dots, a_n) = a_i$, where $1 \leq i \leq n$) and is closed under composition. The smallest clone containing all the basic operations of an algebra \mathbf{A} is denoted $\text{Clo}(\mathbf{A})$ and its elements are called *term operations* of \mathbf{A} . A formal expression defining a term operation from the basic operations is called a *term*. If t is a term, then we denote by $t^{\mathbf{A}}$ the corresponding term operation in \mathbf{A} , or we just write t if no ambiguity is imminent.

The *full idempotent reduct* of an algebra \mathbf{A} is the algebra with the same universe whose operations are all the idempotent term operations of \mathbf{A} .

A subset B of the universe of an algebra \mathbf{A} is called a *subuniverse* if it is closed under all operations (equivalently term operations) of \mathbf{A} . Given a nonempty subuniverse B of \mathbf{A} we can form an algebra \mathbf{B} by restricting all the operations of \mathbf{A} to the set B . In this situation, we say that \mathbf{B} is a *subalgebra* of \mathbf{A} and we write $B \leq \mathbf{A}$ or $\mathbf{B} \leq \mathbf{A}$.

The *product* of algebras $\mathbf{A}_1, \dots, \mathbf{A}_n$ of the same signature is the algebra $\mathbf{A}_1 \times \dots \times \mathbf{A}_n$ whose universe is $A_1 \times \dots \times A_n$ on which the basic operations are computed coordinatewise. The product of n copies of an algebra \mathbf{A} is the n -th *power* of \mathbf{A} and is denoted \mathbf{A}^n . A subalgebra, or a subuniverse, of a product of \mathbf{A} is called a *subpower* of \mathbf{A} . If $R \leq \mathbf{A}_1 \times \dots \times \mathbf{A}_n$ and, for each i , the projection of R onto the i -th coordinate is equal to A_i , then we say that R is *subdirect* in the product and write $R \leq_{sd} \mathbf{A}_1 \times \dots \times \mathbf{A}_n$.

An equivalence relation \sim on the universe of an algebra \mathbf{A} is a *congruence* if it is a subuniverse of \mathbf{A}^2 . The corresponding *quotient algebra* \mathbf{A}/\sim has, as its universe, the set of \sim -blocks, which are denoted $[a]_{\sim}$, $a \in A$, and the basic operations of \mathbf{A}/\sim are defined using arbitrarily chosen representatives. The set of congruences of \mathbf{A} forms a lattice, called the *congruence lattice* of \mathbf{A} .

A *variety* is a class of algebras of the same signature closed under forming subalgebras, products (possibly infinite), quotient algebras and isomorphic copies. A fundamental theorem of universal algebra, due to Birkhoff [14], states that a class of similar algebras is a variety if and only if this class can be defined via a set of identities. The smallest variety containing \mathbf{A} is denoted $\text{HSP}(\mathbf{A})$.

2.2. Relational structures

An n -ary relation on a set A is a subset of A^n , where n is a natural number. Alternatively, a relation can be represented as a mapping $A^n \rightarrow \{\text{true}, \text{false}\}$. We will use both formalisms, e.g. $(1, 2, 3) \in R$ and $R(1, 2, 3)$ both mean that the triple $(1, 2, 3)$ is in the ternary relation R .

A relational structure is a pair $\mathbb{A} = (A; \mathcal{R})$, where A is the *universe* of \mathbb{A} and \mathcal{R} is a set of relations on A . We say that relation S on A is *primitively positively definable*, or *pp-definable*, from \mathbb{A} if it can be defined from the relations in \mathbb{A} by a *pp-formula*, that is, a first order formula that uses conjunction, existential quantification, and the equality relation. For example,

$$S(x, y, z) \quad \text{iff} \quad (\exists u, v) R(x, u) \wedge R(u, v) \wedge R(v, y) \wedge (y = z)$$

is a pp-definition of a ternary relation S from a binary relation R . In accordance with the CSP terminology, clauses of a pp-formula will be called *constraints*.

The equality relation on A is denoted $\Delta_A = \{(a, a) : a \in A\}$; it is pp-definable from every relational structure on A . The projection of a relation $R \subseteq A^n$ onto coordinates i_1, \dots, i_k (not necessarily distinct) is denoted

$$\pi_{i_1, \dots, i_k}(R) = \{(a_{i_1}, \dots, a_{i_k}) : \mathbf{a} = (a_1, \dots, a_n) \in R\};$$

it is pp-definable from R (more precisely, from any relational structure containing R). If $R \subseteq A^n$, $B \subseteq A$, $i \in [n]$, then by *fixing the coordinate i of R to B* we mean forming the relation

$$\{(a_1, \dots, a_n) \in R : a_i \in B\}.$$

This relation is pp-definable from R and the unary relation B . If $S, T \subseteq A^2$, then the relational composition of S and T is defined by

$$S \circ T = \{(a, c) \in A^2 : (\exists b \in A) (a, b) \in S, (b, c) \in T\},$$

it is pp-definable from S and T . Projection, coordinate fixing, and relational composition will be performed more generally on sets $R \subseteq A_1 \times \dots \times A_n$, $S \subseteq A_1 \times A_2$, $T \subseteq A_2 \times A_3$.

We say that an operation $f : A^n \rightarrow A$ is *compatible* with a relation $R \subseteq A^m$ if, for any $\mathbf{a}_1, \dots, \mathbf{a}_n \in R$, the tuple $f(\mathbf{a}_1, \dots, \mathbf{a}_n)$ (where f is applied coordinatewise) is in R . In other words, f is compatible with R if R is a subpower of the algebra $(A; \{f\})$. Notice that f is idempotent iff it is compatible with every singleton unary relation on A .

An operation compatible with all relations of a relational structure \mathbb{A} is a *polymorphism* of \mathbb{A} . The set of all polymorphisms of \mathbb{A} is denoted $\text{Pol}(\mathbb{A})$. This set of operations is always a clone on A . More interestingly, every clone on a finite set can be obtained in this way by the following theorem.

Theorem 2.1 ([15, 29]). *For every finite algebra \mathbf{A} there exists a relational structure \mathbb{A} (with the same universe) such that $\text{Pol}(\mathbb{A}) = \text{Clo}(\mathbf{A})$. In this situation, $R \leq \mathbf{A}^n$ iff R is pp-definable from \mathbb{A} (assuming $R \neq \emptyset$).*

A finite algebra is called finitely related if finitely many relations suffice to determine $\text{Clo}(\mathbf{A})$:

Definition 2.2. A finite algebra \mathbf{A} is said to be *finitely related* if there exists a relational structure \mathbb{A} with finitely many relations such that $\text{Pol}(\mathbb{A}) = \text{Clo}(\mathbf{A})$.

2.3. Congruence modularity

Definition 2.3. A variety is called *congruence modular* if all algebras in it have modular congruence lattices. We define CM to be the class of all algebras that belong to some congruence modular variety.

Gumm [30] characterized the class CM by a useful Maltsev condition. The terms involved in the condition are now called *Gumm terms*. A stronger Maltsev condition was given in [33] by means of *directed Gumm terms*.

Theorem 2.4 ([33]). *The following are equivalent for an algebra \mathbf{A} :*

- \mathbf{A} is in CM .
- There exists a natural number m and a sequence of ternary terms p_1, \dots, p_m, q in $\text{Clo}(\mathbf{A})$, called directed Gumm terms of \mathbf{A} , such that the following identities are satisfied:

$$\begin{aligned} p_1(a, a, b) &= a, \\ p_i(a, b, a) &= a && \text{for all } i \in [m], \\ p_i(a, b, b) &= p_{i+1}(a, a, b) && \text{for all } i \in [m-1], \\ p_m(a, b, b) &= q(a, b, b), \\ q(a, a, b) &= b. \end{aligned}$$

Note that from the second and the fifth identity it follows that directed Gumm terms of \mathbf{A} are necessarily idempotent.

2.4. Few subpowers

Definition 2.5. A finite algebra \mathbf{A} is said to have *few subpowers* if there exists a polynomial p such that, for every n , the number of subuniverses of \mathbf{A}^n is less than $2^{p(n)}$.

Theorem 2.8 gives some of the many equivalent characterizations of the few subpowers property. To state the theorem we require the following definitions.

Definition 2.6. A *cube term* of an algebra \mathbf{A} is an n -ary term $t \in \text{Clo}(\mathbf{A})$ such that for each $1 \leq i \leq n$, \mathbf{A} satisfies an identity of the form $t(x_1, \dots, x_i, \dots, x_n) = y$ where $x_i = x$, $\{x_1, \dots, x_n\} \subseteq \{x, y\}$, and x and y are different variables.

Definition 2.7. A pair (\mathbf{E}, \mathbf{D}) is a *cube term blocker* for \mathbf{A} if $\mathbf{E} \leq \mathbf{D} \leq \mathbf{A}$ and for every $t \in \text{Clo}(\mathbf{A})$ there exists a coordinate i such that

$$t(D, \dots, D, E, D, \dots, D) \subseteq E,$$

where E is at the i -th coordinate.

Theorem 2.8 ([13, 37]). *The following are equivalent for a finite algebra \mathbf{A} :*

- (i) \mathbf{A} has few subpowers.
- (ii) There are no cube term blockers for the full idempotent reduct of \mathbf{A} .
- (iii) \mathbf{A} has a cube term.

Proof. The equivalence of (i) and (iii) was proved in [13]. The equivalence of (ii) and (iii) was first proved in [37] and an alternative proof was given in [11]. \square

The following easy relational translation of cube term blockers will be particularly useful.

Lemma 2.9 ([37]). *Let \mathbf{A} be a finite algebra and (\mathbf{E}, \mathbf{D}) a pair of subalgebras of \mathbf{A} with $\mathbf{E} \leq \mathbf{D}$. Then the following are equivalent:*

- (\mathbf{E}, \mathbf{D}) is a cube term blocker for \mathbf{A} .
- For each n , the relation $D^n \setminus (D \setminus E)^n$ is a subuniverse of \mathbf{A}^n .

2.5. Graphs and digraphs

By a *graph* we mean an undirected graph with a finite vertex set, where loops and multiple edges are allowed. A *cut vertex* is a vertex whose removal increases the number of connected components. A graph is *biconnected* if it is connected and has no cut vertices. A *block* of a graph is a maximal biconnected (induced) subgraph. Note that the intersection of the set of vertices of two different blocks is either empty or equal to $\{v\}$ for a cut vertex v . An example is shown in Figure 2. Notice that a connected graph in which every block is a single edge is a tree.

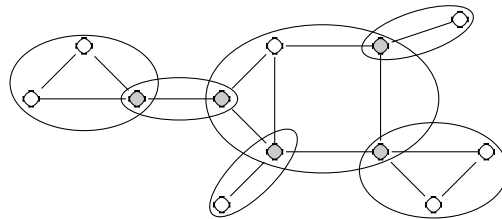


Fig. 2. A graph, its cut vertices (filled-in) and blocks (encircled).

A *walk* in a graph S , or an *S -walk*, from a to b is a sequence of vertices $a = a_1, a_2, \dots, a_k = b$ such that a_i and a_{i+1} are joined by an edge for each $i \in [k - 1]$.

A subset $S \subseteq B \times C$ can be regarded as a bipartite graph without multiple edges whose partite sets are disjoint copies B' (the left partite set) and C' (the right partite set) of $\pi_1(S)$ and $\pi_2(S)$, and $\{a, b\}$ (where $a \in B'$ and $b \in C'$) is an edge iff $(a, b) \in S$. In this context, S is called *linked* if the associated bipartite graph is connected.

A subset $S \subseteq A \times A$ will be sometimes regarded as a directed graph with edge set S . A *directed walk* in S , or a *directed S -walk*, of length $k - 1$ from a to b is a sequence of vertices $a = a_1, a_2, \dots, a_k = b$ such that $(a_i, a_{i+1}) \in S$ for each $i \in [k - 1]$. It is *closed* if $a = b$. A vertex $a \in A$ is a *source* (*sink*, respectively) if it has no incoming (outgoing,

resp.) edge. The *smooth part* of S is the largest $B \subseteq A$ such that $S \cap (B \times B)$ has no sources or sinks in B . It can be described as the set of vertices with a directed S -walk of length $|A|$ from them and a directed S -walk of the same length to them. In particular, the unary relation B is pp-definable from S .

3. Proof of the main theorem

3.1. Reduction to binary structures

In this subsection we show that in order to prove Theorem 1.1 it is enough to consider idempotent algebras determined by *binary relational structures*, i.e., relational structures with at most binary relations. This will make the presentation technically easier. The reduction is based on the following fact (see [3]).

Proposition 3.1. *Let \mathbb{A} be a relational structure whose relations all have arity at most k . Then there exists a binary relational structure $\bar{\mathbb{A}}$ with universe $\bar{A} = A^k$ such that*

$$\text{Pol}(\bar{\mathbb{A}}) = \{\bar{f} : f \in \text{Pol}(\mathbb{A})\},$$

where \bar{f} is defined (if f is n -ary) by

$$\begin{aligned} \bar{f}((a_1^1, \dots, a_k^1), (a_1^2, \dots, a_k^2), \dots, (a_1^n, \dots, a_k^n)) \\ = (f(a_1^1, \dots, a_1^n), f(a_2^1, \dots, a_2^n), \dots, f(a_k^1, \dots, a_k^n)). \end{aligned}$$

Using this proposition we can reduce the main theorem to the following:

Theorem 3.2. *If \mathbb{A} is a finite binary relational structure containing all the singleton unary relations such that the algebra $\mathbf{A} = (A, \text{Pol}(\mathbb{A}))$ is in \mathcal{CM} , then \mathbf{A} has few subpowers.*

Proof of Theorem 1.1 assuming Theorem 3.2. Let \mathbf{A} be a finite, finitely related algebra in \mathcal{CM} and let \mathbb{A} be a relational structure with finitely many relations (say all of them have arity at most k) such that $\text{Pol}(\mathbb{A}) = \text{Clo}(\mathbf{A})$. Let $\bar{\mathbb{A}}$ be the relational structure from Proposition 3.1. By Theorem 2.4, $\bar{\mathbf{A}} = (\bar{A}, \text{Pol}(\bar{\mathbb{A}}))$ is in \mathcal{CM} since $\bar{p}_0, \dots, \bar{p}_m, \bar{q}$ are directed Gumm terms of $\bar{\mathbf{A}}$ whenever p_0, \dots, p_m, q are directed Gumm terms of \mathbf{A} .

Now we add to $\bar{\mathbb{A}}$ all the singleton unary relations, call this structure $\bar{\mathbb{A}}'$, and define $\bar{\mathbf{A}}' = (\bar{A}, \text{Pol}(\bar{\mathbb{A}}'))$. The algebra $\bar{\mathbf{A}}'$ is still in \mathcal{CM} since directed Gumm terms are idempotent. By Theorem 3.2, $\bar{\mathbf{A}}'$ has few subpowers. Therefore, by Theorem 2.8(iii), this algebra has a cube term h , and so does the algebra $\bar{\mathbf{A}}$. By Proposition 3.1, $h = \bar{t}$ for some polymorphism t of \mathbf{A} . The operation t is clearly a cube term of \mathbf{A} , and thus \mathbf{A} has few subpowers, as required. \square

3.2. Proof-sketch of Theorem 3.2

The proof is by contradiction. Let \mathbb{A} be a finite binary relational structure containing all the singleton unary relations, and let $\mathbf{A} = (A; \text{Pol}(\mathbb{A}))$ be its algebra of polymorphisms. We assume that \mathbf{A} is in CM and it does not have few subpowers. Notice that \mathbf{A} is idempotent. For convenience, we add to \mathbb{A} all the unary and binary relations which are pp-definable from \mathbb{A} . By Theorem 2.1, this change does not affect $\text{Clo}(\mathbf{A})$. We fix a sequence p_1, \dots, p_m, q of directed Gumm terms of \mathbf{A} (see Theorem 2.4) and a cube term blocker (\mathbf{E}, \mathbf{D}) for \mathbf{A} (see Theorem 2.9) such that $|D|$ is minimal. Let $F = D \setminus E$ (this does not need to be a subuniverse of \mathbf{A}). All the objects defined in this paragraph will stay fixed throughout the proof.

Since (\mathbf{E}, \mathbf{D}) is a cube term blocker of \mathbf{A} , the relation $D^n \setminus F^n$ is, by Theorem 2.9, a subpower of \mathbf{A} for every positive integer n . It follows from Theorem 2.1 that this relation can be defined by a pp-formula Φ from the structure \mathbb{A} . The proof now goes roughly as follows: We choose a large enough arity n and use Zhuk's technique [44] to obtain a nicer (tree) pp-formula Φ' defining a similar relation (so called CTB-relation as defined in the next subsection) of the same arity n ; then we get an even nicer (comb) pp-formula Ω defining a CTB-relation of arity $\log_2(n-1)$; finally, we reach a contradiction by showing that a CTB-relation of a sufficiently large arity defined by a comb-formula cannot be compatible with directed Gumm terms.

The following subsections give details of the proof using three core technical lemmas which are proved in Section 4.

3.3. CTB-relation

We start with the definition of a CTB-relation (CTB stands for Cube Term Blocking).

Definition 3.3. A relation $R \leq \mathbf{D}^n$ is a *CTB-relation* if $F^n \cap R = \emptyset$ and for every $i \in [n]$ there exists $e_i \in E$ such that $D^{i-1} \times \{e_i\} \times D^{n-i} \subseteq R$.

We set $n = 4^{3^{|\mathbf{A}|}} + 1$ and take an n -ary CTB-relation $U \leq \mathbf{A}^n$, say $U = D^n \setminus F^n$. We take a pp-formula defining U from \mathbb{A} , denote by Φ the quantifier-free part and by x_1, \dots, x_n the free variables of this pp-formula. For an arbitrary sequence y_1, \dots, y_i of variables in Φ , we denote $\Phi[y_1, \dots, y_i]$ the formula obtained by existentially quantifying over the remaining variables, and $\Phi(y_1, \dots, y_i)$ the i -ary relation defined by $\Phi[y_1, \dots, y_i]$. In particular, we have $U = \Phi(x_1, \dots, x_n)$.

Next we modify Φ and get a slightly nicer formula without changing $\Phi(x_1, \dots, x_n)$. (This procedure can be applied to any quantifier-free formula Ξ such that $\Xi(x_1, \dots, x_n)$ is a CTB-relation.) We define a graph $\text{Graph}(\Phi)$ in the following way. Vertices are the variables in Φ and the edges joining x and y are the binary constraints of the form $S(x, y)$ or $S(y, x)$ in Φ . (In the same way, a graph $\text{Graph}(\Xi)$ is associated to any pp-formula Ξ over \mathbb{A} .) For the following definition, we introduce auxiliary notation: Let $x \neq y$ be vertices of $\text{Graph}(\Phi)$ and let G_x be the graph obtained from $\text{Graph}(\Phi)$ by removing the vertex x (and incident edges). We denote $\text{Cut}(x, y)$ the set of vertices in the component of G_x which contains y .

A pp-formula will be called *simple* (with respect to x_1, \dots, x_n) if

- (a) the associated graph is connected, without loops or multiple edges,
- (b) the set of vertices of degree 1 is equal to $\{x_1, \dots, x_n\}$,
- (c) if $x \neq y$ are vertices, then $\text{Cut}(x, y) \cap \{x_1, \dots, x_n\} \neq \emptyset$,
- (d) if $x \neq y$ are vertices such that $\text{Cut}(x, y) \cap \text{Cut}(y, x) \cap \{x_1, \dots, x_n\} = \emptyset$, then $\text{Cut}(x, y) \cap \text{Cut}(y, x) = \emptyset$, and
- (e) the formula has no unary constraints.

Note that (b), (c) and (d) imply that each block of $\text{Graph}(\Phi)$ with more than one edge contains at least three cut vertices of $\text{Graph}(\Phi)$.

Φ can be transformed to a simple quantifier-free pp-formula (over \mathbb{A}) which defines the same relation $\Phi(x_1, \dots, x_n)$ in the following way. First we ensure that x_1, \dots, x_n have degree 1 by renaming x_i as x'_i , adding x_i , and adding the constraint $x_i = x'_i$. Next observe that if two variables x_i and x_j are in different components of $\text{Graph}(\Phi)$, then $\Phi(x_1, \dots, x_n)$ is a product of two relations of smaller arity. Since a CTB-relation is not a product of two relations of smaller arity, all the variables x_1, \dots, x_n must belong to the same component. Thus we can make $\text{Graph}(\Phi)$ connected by removing all the variables (and constraints) in different components. A binary constraint of the form $S(x, x)$ can be replaced by the unary constraint $T(x)$ (where T is defined by $T(x)$ iff $S(x, x)$; recall that \mathbb{A} is closed under pp-definitions of unary and binary relations) and constraints $S_1(x, y), \dots, S_k(x, y), S'_1(y, x), \dots, S'_{k'}(y, x)$ can be replaced by a single constraint $T(x, y)$ (where T is again defined in the obvious way). Similarly, if x, y are different vertices such that $\text{Cut}(x, y) \cap \text{Cut}(y, x) \cap \{x_1, \dots, x_n\} = \emptyset$, then we can replace all the constraints that use variables in $\text{Cut}(x, y) \cap \text{Cut}(y, x)$ by a single constraint $T(x, y)$. Moreover, if $\text{Cut}(x, y) \cap \{x_1, \dots, x_n\} = \emptyset$, then we can replace all the constraints that use variables in $\text{Cut}(x, y)$ by a single unary constraint $T(x)$. Repeating these modifications results in a formula satisfying (a)–(d). Finally, unary constraints can be hidden in binary constraints, for example, constraints $T(x)$ and $S(x, y)$ can be replaced by a single constraint $S'(x, y)$ (with $S' = S \cap (T \times A)$, a pp-definable relation). This ensures (e).

The resulting pp-formula will be called a *simplified form* of the original formula. In the next subsection we assume that Φ was already simplified.

3.4. Tree definition

We describe a construction which transforms the simple formula Φ into a new quantifier-free simple pp-formula Φ' such that $\Phi'(x_1, \dots, x_n)$ is still a CTB-relation, possibly different from $U = \Phi(x_1, \dots, x_n)$. The construction depends on a variable y and a binary constraint in Φ whose scope contains y , say $T(y, z)$, such that $y, z \notin \{x_1, \dots, x_n\}$ (the case where the constraint is of the form $T(z, y)$ is completely analogous).

The construction is divided into three steps. In the first step, we build from Φ a new formula Ψ by adding the constraints $D(x_1), \dots, D(x_n)$, adding a new variable y_* , removing the constraint $T(y, z)$, adding the constraint $T(y_*, z)$, and adding the constraints $C(y)$ and $C(y_*)$, where $C = \Phi(y)$. See Figure 3.

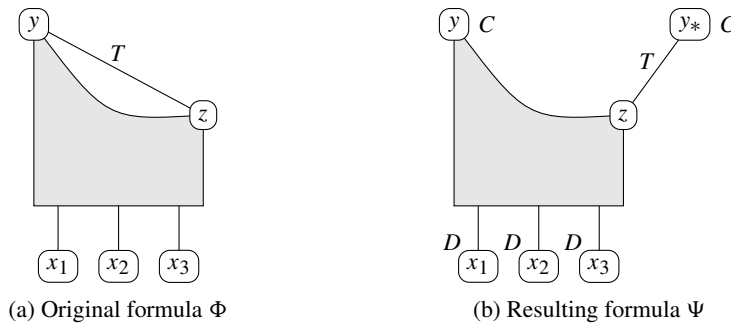


Fig. 3. The first step for $n = 3$.

In the second step, we define a new formula Θ as follows. For each $i \in [l]$, where $l = |A|$, we take a copy Ψ^i of the formula Ψ by renaming each variable w in Ψ as w^i . Then we take the conjunction of Ψ^1, \dots, Ψ^l and identify variables y_*^i and y^{i+1} for each $i \in [l - 1]$. The resulting formula is shown in Figure 4.

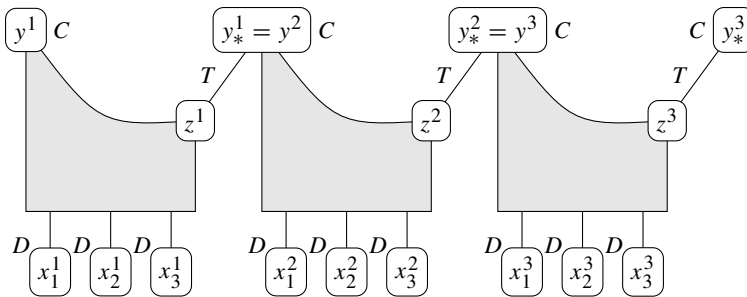


Fig. 4. The second step: formula Θ for $n = l = 3$.

Before describing the third step, we prove a claim which says that if we added to Θ the equality constraints $x_i^1 = x_i^2 = \dots = x_i^l$ for each $i \in [n]$ and existentially quantified the remaining variables, then the resulting formula would define a CTB-relation. The argument is based on the following lemma, proved in Subsection 4.2.

Lemma 3.4. *Let $C \leq A$, let $R \leq C^2 \times D^n$ be such that $\Delta_C \subseteq \pi_{1,2}(R)$ and suppose that there exists $\mathbf{f} = (f_1, \dots, f_n) \in F^n$ such that the smooth part of the digraph $Q = \{(a_1, a_2) : (a_1, a_2, f_1, \dots, f_n) \in R\}$ is nonempty. Then $(\Delta_C \times F^n) \cap R \neq \emptyset$.*

Let

$$V = \Theta(x_1^1, \dots, x_1^l, x_2^1, \dots, x_2^l, \dots, x_n^1, \dots, x_n^l)$$

and let W be the n -ary relation defined by

$$W(a_1, \dots, a_n) \text{ iff } V(\underbrace{a_1, \dots, a_1}_l, \dots, \underbrace{a_n, \dots, a_n}_l).$$

Claim. W is a CTB-relation.

Proof. By construction, $W \leq \mathbf{D}^n$ and also $U \subseteq W$, since any satisfying evaluation of variables in Φ gives a satisfying evaluation of variables in Θ . It is thus enough to check that $W \cap F^n = \emptyset$. Suppose otherwise, that is, there exists $(f_1, \dots, f_n) \in W \cap F^n$.

Let $R = \Psi(y, y_*, x_1, \dots, x_n)$. By construction of Ψ , we see that $\Delta_C \subseteq \pi_{1,2}(R)$ and the projection of $(\Delta_C \times A^n) \cap R$ onto the coordinates $3, 4, \dots, n + 2$ is equal to $U = \Phi(x_1, \dots, x_n)$. Since $U \cap F^n = \emptyset$ by the definition of a CTB-relation, we have $(\Delta_C \times F^n) \cap R = \emptyset$. From the construction of Θ and from $(f_1, \dots, f_1, f_2, \dots, f_2, f_n, \dots, f_n) \in V$ it follows that there exist $b_1, \dots, b_{l+1} \in C$ such that $(b_i, b_{i+1}, f_1, \dots, f_n) \in R$ for every $i \in [l]$. Therefore, the digraph $Q = \{(a_1, a_2) : (a_1, a_2, f_1, \dots, f_n) \in R\}$ contains a walk of length $l = |A|$, which implies that Q contains a closed walk. Hence the smooth part of Q is nonempty. Lemma 3.4 now implies $(\Delta_C \times F^n) \cap R \neq \emptyset$, a contradiction. \square

In the third step, we modify Θ using the data provided by the following lemma, which is proved in Subsection 4.3.

Lemma 3.5. Let $V \leq \mathbf{D}^{ln}$ be such that

$$W = \{(a_1, \dots, a_n) : \underbrace{(a_1, \dots, a_1)}_l, \underbrace{(a_2, \dots, a_2)}_l, \dots, \underbrace{(a_n, \dots, a_n)}_l \in V\}$$

is a CTB-relation. Then for each $i \in [n]$ there exist $m_i \in [l]$ and $C_i^j \leq \mathbf{A}$, $j \in [l] \setminus \{m_i\}$, such that

$$U' = \{(a_1^{m_1}, \dots, a_n^{m_n}) : (a_1^1, \dots, a_1^{l_1}, a_2^1, \dots, a_2^{l_2}, \dots, a_n^1, \dots, a_n^{l_n}) \in R, \\ (\forall i \in [n])(\forall j \in [l], j \neq m_i) a_i^j \in C_i^j\}$$

is a CTB-relation.

Note that we can apply this lemma in our situation, since W is a CTB-relation by the previous claim. The formula Φ' is obtained from Θ as follows. For each $i \in [n]$, we rename the variable $x_i^{m_i}$ as x_i and, for each $i \in [n]$ and $j \in [l]$, $j \neq m_i$, we add the constraint $C_i^j(x_i^j)$. Finally, we replace the resulting formula Φ' by its simplified form. It follows from Lemma 3.5 that $U' = \Phi'(x_1, \dots, x_n)$ is a CTB-relation.

The above construction is used repeatedly in the following claim (with suitable choices of variables and constraints) to obtain a tree-formula defining a CTB-relation.

Definition 3.6. A tree-formula is a pp-formula whose associated graph is a tree.

Claim. There exists a tree-formula defining a CTB-relation of arity n .

Proof. The reasoning follows [44, proof of Theorem 5.2].

To the original formula Φ we assign a triple $(\alpha, n - \beta, \gamma)$ of nonnegative integers that depend on $\text{Graph}(\Phi)$. The first parameter, α , is the greatest number of edges in a block. To define β , we need to introduce further notation. Let B be a block and $w \in B$ a cut vertex.

When we remove $B \setminus \{w\}$ from $\text{Graph}(\Phi)$, we get a disconnected graph, one component of which contains w . We denote the set of vertices in this component by $\text{Cut}(B, w)$ and define

$$\delta(B, w) = |\text{Cut}(B, w) \cap \{x_1, \dots, x_n\}|.$$

Let $\delta(B)$ be the maximum of $\delta(B, w)$ among all cut vertices $w \in B$. Now β is the minimum $\delta(B)$ among blocks B with α edges. Finally, γ is the number of blocks B with α edges and $\beta = \delta(B)$.

If $\alpha = 1$, then $\Phi[x_1, \dots, x_n]$ is a tree-formula defining a CTB-relation and we are done. Otherwise, we take a block B with α edges and $\delta(B) = \beta$, we take a variable y in B and an incident constraint $T(y, z)$ or $T(z, y)$ (where z is in B), and we perform the construction described above. We claim that the triple $(\alpha', n - \beta', \gamma')$ assigned to Φ' is strictly smaller than $(\alpha, n - \beta, \gamma)$ in the lexicographic ordering. We denote by Φ'' the formula obtained by the three-step construction *before* simplification (thus Φ' is a simplified form of Φ'') and we use the notation $\delta''(\dots)$ analogously to $\delta(\dots)$.

The construction removes an edge from B , thus all the blocks of Φ'' coming from B have fewer edges. Every block of $\text{Graph}(\Phi)$ different from B has l isomorphic copies in $\text{Graph}(\Phi'')$ and each remaining block of $\text{Graph}(\Phi'')$ consists of a single edge. Since the simplification procedure only removes blocks or shrinks blocks to edges, we get $\alpha' \leq \alpha$ and every block of $\text{Graph}(\Phi')$ with α edges comes from a block of $\text{Graph}(\Phi)$ different from B with α edges.

Consider a block $C \neq B$ of $\text{Graph}(\Phi)$ with α edges and its isomorphic copies C^1, \dots, C^l in $\text{Graph}(\Phi'')$. Let $t_1 \in C$ denote the cut vertex of $\text{Graph}(\Phi)$ such that $B \subseteq \text{Cut}(C, t_1)$, let t_2, \dots, t_k denote the remaining cut vertices in C , let t_j^i be the corresponding cut vertices of $\text{Graph}(\Phi'')$ in C^i , and let $s \in B$ be the cut vertex such that $C \subseteq \text{Cut}(B, s)$.

By the simplicity of Φ , $\min\{\delta(C, t_1), \dots, \delta(C, t_k)\} \geq 1$ and $k \geq 3$ (see the note after the definition of a simple formula). Since $\text{Cut}(B, s) \supseteq \text{Cut}(C, t_2) \dot{\cup} \dots \dot{\cup} \text{Cut}(C, t_k)$, we get

$$\beta = \delta(B) \geq \delta(B, s) \geq \delta(C, t_2) + \dots + \delta(C, t_k) > \max\{\delta(C, t_2), \dots, \delta(C, t_k)\}.$$

But $\delta(C) \geq \beta$ by the definition of β , so $\delta(C) = \delta(C, t_1)$.

By construction,

$$\sum_{i=1}^l \delta''(C^i, t_j^i) = \delta(C, t_j) \quad \text{for every } j \in \{2, \dots, k\},$$

in particular $\delta''(C^i, t_j^i) \leq \delta(C, t_j)$ for every $i \in [l]$ and $j \in \{2, \dots, k\}$. Therefore, since $\sum_{j=1}^k \delta''(C^i, t_j^i) = n = \sum_{j=1}^k \delta(C, t_j)$, we have $\delta''(C^i, t_1^i) \geq \delta(C, t_1)$. Consequently,

$$\delta''(C^i, t_1^i) \geq \delta(C, t_1) > \max\{\delta(C, t_2), \dots, \delta(C, t_k)\} \geq \max\{\delta''(C^i, t_2), \dots, \delta''(C^i, t_k)\},$$

and thus

$$\delta''(C^i) = \delta''(C^i, t_1^i) \geq \delta(C, t_1) = \delta(C) \geq \beta.$$

It follows that $(\alpha', n - \beta') \leq (\alpha, n - \beta)$ in the lexicographic ordering.

Moreover, if $\delta''(C^i) = \beta$ for some $i \in [l]$, then necessarily $\delta''(C^i, t_1^i) = \delta(C, t_1)$, whence $\delta''(C^i, t_j^i) = \delta(C, t_j)$ for every $j \in \{2, \dots, k\}$, which implies that $\delta''(C^{i'}, t_j^{i'}) = 0$ for every $i \neq i' \in [l]$ and every $j \in \{2, \dots, k\}$. But then each block $C^{i'}$, $i' \neq i$, will be removed during the simplification procedure. In summary, if some copy C^i of C in $\text{Graph}(\Phi'')$ achieves the minimum $\delta''(C^i) = \beta$, then necessarily $\delta(C) = \beta$ and all the other copies of C in $\text{Graph}(\Phi'')$ will be eventually removed. Therefore $(\alpha', n - \beta', \gamma')$ is strictly smaller than $(\alpha, n - \beta, \gamma)$, as claimed.

It follows that in finitely many steps we get a (quantifier-free) tree-formula Φ such that $\Phi[x_1, \dots, x_n]$ is a CTB-relation. \square

3.5. Comb definition

Our goal is to find a comb-formula of sufficiently large arity that represents a CTB-relation. This can be achieved by transforming the resulting tree-formula into a comb-formula in exchange for an exponential loss in arity.

Definition 3.7. A pp-formula over the set $\{x'_1, \dots, x'_{n'}, w_1, \dots, w_{n'}\}$ of variables is a *comb-formula* if the set of edges of its associated graph is equal to

$$\{x'_1 w_1, x'_2 w_2, \dots, x'_{n'} w_{n'}, w_1 w_2, w_2 w_3, \dots, w_{n'-1} w_{n'}\}$$

(see Figure 5).

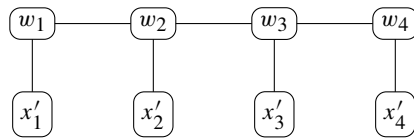


Fig. 5. Graph of a comb-formula for $n' = 4$.

Claim. *There exists a quantifier-free comb-formula Ω with $n' \geq \log_2(n - 1)$ such that $\Omega(x'_1, \dots, x'_{n'})$ is a CTB-relation.*

Proof. Let Φ be a simple quantifier-free tree-formula such that $\Phi(x_1, \dots, x_n)$ is a CTB-relation. Note that $\text{Graph}(\Phi)$ has no vertex of degree 2 and the set of its leaves is equal to $\{x_1, \dots, x_n\}$ (both facts follow from simplicity). We can modify Φ to ensure that each vertex of $\text{Graph}(\Phi)$ has degree 1 or 3 by repeated application of the following procedure: take a variable v of degree $d > 3$; split it into two variables v_1, v_2 so that in $\lfloor d/2 \rfloor$ of the incident constraints the variable v is replaced by v_1 and in the remaining $\lceil d/2 \rceil$ constraints v is replaced by v_2 ; finally, add the equality constraint $v_1 = v_2$.

Let $i = 0$. Starting from an arbitrary leaf we will follow a path (nonintersecting walk), defining x'_i s and w_i s on the way. If we are at the beginning, then we continue in the unique direction. If we are at a vertex of degree 3, then we have two options to continue. We select the one with more leaves ahead (if both options have the same number of leaves ahead, we decide arbitrarily) and before we continue to walk we increment i , define w_i to be the vertex we are at, define x'_i to be any of the leaves in the direction we did not choose

(note that $x'_i \in \{x_1, \dots, x_n\}$). When we arrive at a leaf we stop and set $n' = i$. A simple computation shows that $n' \geq \log_2(n - 1)$.

Choose $f \in F$ arbitrarily and set $S = \{f\}$. Let Ω' be the formula obtained from Φ by adding the unary constraint $S(x_i)$ for each $x_i \notin \{x'_1, \dots, x'_{n'}\}$. Observe that the relation $\Omega'(x'_1, \dots, x'_{n'})$ is obtained from the CTB-relation $\Phi(x_1, \dots, x_n)$ by fixing several coordinates to $\{f\}$ and projecting onto the remaining coordinates, therefore it is a CTB-relation. It is easy to see that the simplified form of Ω' (with respect to $x'_1, \dots, x'_{n'}$) is a comb-formula. This can be visualized as follows: straighten the resulting path and shake the tree; after the simplification, the path becomes $w_1, \dots, w_{n'}$ and the tree below each w_i becomes the edge $x'_i w_i$. \square

3.6. Contradiction

The proof of Theorem 3.2 is concluded using the following lemma, proved in Subsection 4.4.

Lemma 3.8. *If $\Omega(x'_1, \dots, x'_{n'})$ is a CTB-relation, where Ω is a quantifier-free comb-formula, then $n' < 2 \cdot 3^{|A|}$.*

Recall that we have chosen $n = 4^{3^{|A|}} + 1$, thus $\log_2(n - 1) = 2 \cdot 3^{|A|}$, so that this lemma and the last claim contradict each other.

3.7. Remarks on the proof

The structure of the proof is, to some level of detail, the same as in the previous proofs [3, 44] of special cases:

1. obtain a tree definition of a “bad” relation,
2. obtain a comb definition of a “bad” relation,
3. prove that a “bad” relation defined by a comb-formula cannot have large arity.

This approach was used in [3] to prove that a finitely related algebra in CD is in NU. Independently, Zhuk [44] proved a slightly weaker result that a finitely related algebra in NU has a near-unanimity term of “small” arity. As for the proofs, the difference in generality is inessential and only concerns the third item (slightly more work is needed to get the stronger result).

The main difference between [3] and [44] is in carrying out the first item. In [3], this item is quite easily derived from an existing result on the CSP [6]. Zhuk, on the other hand, introduces the construction presented in Subsection 3.4 and uses variants of Lemmas 3.4 and 3.5. The first approach is not yet applicable for the main result in this paper because a corresponding result on the CSP is not available. In fact, we do not even know what the CSP result should be.

4. Technical core

In this section, we fill in the gaps in the proof of Theorem 1.1. We keep some of the assumptions made in the last section: \mathbf{A} is a finite idempotent algebra in CM with directed Gumm terms p_1, \dots, p_m, q , (\mathbf{E}, \mathbf{D}) is a cube term blocker for \mathbf{A} with $|D|$ minimal, and $F = D \setminus E$.

4.1. Tools

Recall from Subsection 2.5 that a set $S \subseteq B \times C$ can be regarded as a bipartite graph with “left” part $\pi_1(S)$ and “right” part $\pi_2(S)$, and that we call S linked if this graph is connected.

We call two elements $a \in B$ (or $a \in C$) and $a' \in B$ (or $a' \in C$) S -linked if there is an S -walk from a to a' . Note that it must be clear from the context whether a and a' are from the copy of $\pi_1(S)$ or $\pi_2(S)$ because it is often the case that B and C are not disjoint (sometimes even $B = C = A$). With this agreement, we define the left (right, resp.) connectivity equivalences on $\pi_1(S)$ ($\pi_2(S)$, resp.):

$$\begin{aligned}\lambda_S &= \{(b, b') \in B^2 : b \in B \text{ and } b' \in B \text{ are } S\text{-linked}\}, \\ \rho_S &= \{(c, c') \in C^2 : c \in C \text{ and } c' \in C \text{ are } S\text{-linked}\}.\end{aligned}$$

The neighborhood of a subset $B' \subset B$ ($C' \subseteq C$, resp.) is denoted by $(B')^{+S}$ ($(C')^{-S}$, resp.):

$$\begin{aligned}(B')^{+S} &= \{c \in C : (\exists b \in B') (b, c) \in S\}, \\ (C')^{-S} &= \{b \in B : (\exists c \in C') (b, c) \in S\}.\end{aligned}$$

The set S is called *rectangular* if it is a disjoint union of sets of the form $B' \times C'$ where $B' \subseteq B$ and $C' \subseteq C$. Observe that S is rectangular iff $(b, c), (b', c), (b', c') \in S$ implies $(b, c') \in S$ for every $b, b' \in B$ and $c, c' \in C$.

Almost all objects in the proofs will be subuniverses of algebras in $\text{HSP}(\mathbf{A})$, with the inconvenient exception of the set F . Several constructions which produce subuniverses are summarized in the following lemma. It will be used extensively but silently. Also, we will always tacitly assume that **all the algebras we talk about are finite**, although it is not always necessary.

Lemma 4.1. *Let $\mathbf{C}_1, \dots, \mathbf{C}_k \in \text{HSP}(\mathbf{A})$.*

- *Every block of a congruence of \mathbf{C}_1 is a subuniverse of \mathbf{C}_1 .*
- *The projection of a subuniverse $R \leq \mathbf{C}_1 \times \dots \times \mathbf{C}_k$ onto coordinates i_1, \dots, i_j is a subuniverse of $\mathbf{C}_{i_1} \times \dots \times \mathbf{C}_{i_j}$.*
- *The set obtained by fixing the i -th coordinate of $R \leq \mathbf{C}_1 \times \dots \times \mathbf{C}_k$ to $B \leq \mathbf{C}_i$ is a subuniverse of \mathbf{R} (and thus also a subuniverse of $\mathbf{C}_1 \times \dots \times \mathbf{C}_k$).*
- *If $S \leq \mathbf{C}_1 \times \mathbf{C}_2$, then λ_S (ρ_S , resp.) is a congruence of $\pi_1(\mathbf{S})$ ($\pi_2(\mathbf{S})$, resp.). Moreover, if $C'_1 \leq \mathbf{C}_1$ ($C'_2 \leq \mathbf{C}_2$, resp.), then $(C'_1)^{+S} \leq \mathbf{C}_2$ ($(C'_2)^{-S} \leq \mathbf{C}_1$, resp.).*
- *If $S_1 \leq \mathbf{C}_1 \times \mathbf{C}_2$ and $S_2 \leq \mathbf{C}_2 \times \mathbf{C}_3$, then $S_1 \circ S_2 \leq \mathbf{C}_1 \times \mathbf{C}_3$.*

Proof. Proofs are straightforward and are omitted. We just note that the first item requires idempotency and that, in the remaining items, if $\mathbf{C}_1 = \cdots = \mathbf{C}_k$, then the constructions are pp-definitions, and thus the claims follow from an (easy) part of Theorem 2.1. \square

The next lemmas are consequences of the definition of a blocker, minimality of $|D|$, and directed Gumm identities. We give them names for easier referencing.

Lemma 4.2 (DDE Lemma). $q(D, D, E) \subseteq E$.

Proof. By the definition of a cube term blocker, we must have $q(E, D, D) \subseteq E$, or $q(D, E, D) \subseteq E$, or $q(D, D, E) \subseteq E$. But the first two inclusions are impossible because of the identity $q(a, a, b) = b$ applied to $a \in E$ and $b \in F$. \square

Lemma 4.3 (Minimality Lemma). *If $C \leq \mathbf{D}$ and $C \cap E \neq \emptyset \neq C \cap F$, then $C = D$.*

Proof. Under the assumptions, $(\mathbf{C} \cap \mathbf{E}, \mathbf{C} \cap \mathbf{D})$ is a cube term blocker. Therefore $C = D$ by the minimality of $|D|$. \square

Recall that a *Maltsev operation* on a set B is a ternary operation t on B such that $t(b, b, a) = a = t(a, b, b)$ for every $a, b \in B$. Note that the term q in any algebra in $\text{HSP}(\mathbf{A})$ automatically satisfies the first identity.

Lemma 4.4 (Rectangularity Lemma). *Let $S \leq \mathbf{B} \times \mathbf{C}$ where $\mathbf{B}, \mathbf{C} \in \text{HSP}(\mathbf{A})$. If q is a Maltsev term in \mathbf{B} , then S is rectangular.*

Proof. If $(b, c), (b', c), (b', c') \in Q$, then

$$Q \ni q((b, c), (b', c), (b', c')) = (q(b, b', b'), q(c, c, c')) = (b, c'). \quad \square$$

An important tool for the proofs are Gumm-absorbing subuniverses and their properties stated below.

Definition 4.5. Let $\mathbf{B} \in \text{HSP}(\mathbf{A})$. A nonempty subuniverse C of \mathbf{B} is a *Gumm-absorbing subuniverse*, written $C \triangleleft_G \mathbf{B}$, if $p_i(C, B, C) \subseteq C$ for every $i \in [m]$.

By the definition of directed Gumm terms and idempotency of \mathbf{B} , every singleton $\{b\}$, $b \in B$, is a Gumm-absorbing subuniverse of \mathbf{B} . A trivial Gumm-absorbing subuniverse of \mathbf{B} is B . Further Gumm-absorbing subuniverses can be obtained from the following lemma.

Lemma 4.6 (Forced Absorption Lemma). *Let $\mathbf{C}_1, \dots, \mathbf{C}_k \in \text{HSP}(\mathbf{A})$.*

- *Every block of a congruence of \mathbf{C}_1 is a Gumm-absorbing subuniverse of \mathbf{C}_1 . (In particular, singletons are Gumm-absorbing subuniverses.)*
- *If $S \triangleleft_G \mathbf{R} \leq \mathbf{C}_1 \times \cdots \times \mathbf{C}_k$, then $\pi_{i_1, \dots, i_j}(S) \triangleleft_G \pi_{i_1, \dots, i_j}(\mathbf{R})$ for any $i_1, \dots, i_j \in [k]$.*
- *The set obtained by fixing the i -th coordinate of $R \leq \mathbf{C}_1 \times \cdots \times \mathbf{C}_k$ to $B \triangleleft_G \pi_i(\mathbf{R})$ is a Gumm-absorbing subuniverse of \mathbf{R} .*
- *If $S \leq_{sd} \mathbf{C}_1 \times \mathbf{C}_2$ and $B \triangleleft_G \mathbf{C}_1$, then $B^{+S} \triangleleft_G \mathbf{C}_2$.*

Proof. Straightforward. \square

The importance of (Gumm-)absorption stems from the fact that it absorbs some connectivity properties. Two such properties are stated in the following “walking lemmas”.

Lemma 4.7 (Bipartite Walking Lemma). *Let $\mathbf{B}, \mathbf{C} \in \text{HSP}(\mathbf{A})$, $Q \triangleleft_G \mathbf{S} \leq \mathbf{B} \times \mathbf{C}$, and $b, b' \in \pi_1(Q)$. If b and b' are S -linked, then b and $q(b, b', b')$ are Q -linked. (A similar claim holds for $c, c' \in \pi_2(Q)$.)*

Proof. Take $c, c' \in C$ such that $(b, c), (b', c') \in Q$ and take an S -walk

$$b = b_1, c_1, b_2, c_2, \dots, c_{k-1}, b_k = b'.$$

Since Q is a Gumm-absorbing subuniverse of S , the pairs $p_i((b, c), (b_j, c_j), (b', c')) = (p_i(b, b_j, b'), p_i(c, c_j, c'))$ and $(p_i(b, b_{j+1}, b'), p_i(c, c_j, c'))$ are in Q for each i, j . Therefore

$$\begin{aligned} b &= p_1(b, b, b') = p_1(b, b_1, b'), p_1(c, c_1, c'), p_1(b, b_2, b'), \dots, p_1(c, c_{k-1}, c'), \\ p_1(b, b_k, b') &= p_1(b, b', b') = p_2(b, b, b'), \dots, p_2(b, b', b') = p_3(b, b, b'), \dots, \\ & \qquad \qquad \qquad p_m(b, b', b') = q(b, b', b') \end{aligned}$$

is a Q -walk from b to $q(b, b', b')$. □

Lemma 4.8 (Directed Walking Lemma). *Let $\mathbf{B} \in \text{HSP}(\mathbf{A})$, $Q \triangleleft_G \mathbf{S} \leq \mathbf{B} \times \mathbf{B}$, and $a, b \in B$. If $(a, a), (b, b) \in Q$ and $(a, b) \in S$, then there are directed Q -walks from a to $q(a, b, b)$ and from $q(b, a, a)$ to b .*

Proof. The sequences

$$\begin{aligned} a &= p_1(a, a, b), p_1(a, b, b) = p_2(a, a, b) = \dots = p_m(a, b, b) = q(a, b, b), \\ q(b, a, a) &= p_m(b, a, a), p_m(b, b, a) = p_{m-1}(b, a, a), \dots, p_1(b, b, a) = b \end{aligned}$$

are directed Q -walks. □

Lemma 4.9 (Edge Absorption Lemma). *Let $\mathbf{B}, \mathbf{C} \in \text{HSP}(\mathbf{A})$, and $Q \triangleleft_G \mathbf{S} \leq \mathbf{B} \times \mathbf{C}$. If S is linked, then q is a Maltsev term in $\pi_1(\mathbf{Q})/\lambda_Q$ and in $\pi_2(\mathbf{Q})/\rho_Q$.*

Proof. By the Bipartite Walking Lemma, a is Q -linked to $q(a, b, b)$ for every $a, b \in \pi_1(Q)$. This implies that a and $q(a, b, b)$ are in the same λ_Q -block, therefore

$$q^{\pi_1(\mathbf{Q})/\lambda_Q}([a]_{\lambda_Q}, [b]_{\lambda_Q}, [b]_{\lambda_Q}) = [a]_{\lambda_Q}.$$

The other Maltsev identity $q^{\pi_1(\mathbf{Q})/\lambda_Q}([b]_{\lambda_Q}, [b]_{\lambda_Q}, [a]_{\lambda_Q}) = [a]_{\lambda_Q}$ follows trivially from $q(b, b, a) = a$. Similarly, q is a Maltsev term in $\pi_2(\mathbf{Q})/\rho_Q$. □

Lemma 4.10 (Vertex Absorption Lemma). *Let $\mathbf{B}, \mathbf{C} \in \text{HSP}(\mathbf{A})$, $S \leq_{sd} \mathbf{B} \times \mathbf{C}$, $B' \triangleleft_G \mathbf{B}$, and assume S is linked. Then $S \cap (B' \times (B')^{+S})$ is linked.*

Proof. For a contradiction, assume that $S \cap (B' \times (B')^{+S})$ is not linked. We inductively define $G_1 = B'$, $G_{2i} = (G_{2i-1})^{+S}$ and $G_{2i+1} = (G_{2i})^{-S}$. By the Forced Absorption Lemma, all these sets are Gumm-absorbing subuniverses of either \mathbf{B} (for odd k) or \mathbf{C} (for even k).

Since S is linked and subdirect in $B \times C$, we have $G_{2i} = B$ and $G_{2i+1} = C$ for all sufficiently large i . Therefore we can take the largest k such that S intersected with $G_k \times G_{k+1}$ (for odd k) or $G_{k+1} \times G_k$ (for even k) is not linked. For simplicity assume that k is odd; the other case is symmetric. Let $Q = S \cap (G_k \times G_{k+1})$. By the Forced Absorption Lemma, $Q \triangleleft_G \mathbf{S}$ and then the Edge Absorption Lemma implies that q is a Maltsev term of \mathbf{G}_{k+1}/ρ_Q .

By the choice of k , the intersection $S \cap (G_{k+2} \times G_{k+1})$ is linked. It follows that there are $b \in G_{k+2}$ and $c, c' \in G_{k+1}$ such that $(b, c), (b, c') \in S$ and $[c]_{\rho_Q} \neq [c']_{\rho_Q}$. There exists $b' \in G_k$ with $(b', c') \in S$ as $c' \in G_{k+1} = G_k^{+S}$. Now

$$S \ni q((b, c), (b, c'), (b', c')) = (q(b, b, b'), q(c, c', c')) = (b', q(c, c', c')).$$

The Q -walk $q(c, c', c'), b', c'$ shows that $[c']_{\rho_Q} = [q(c, c', c')]_{\rho_Q}$. But q is a Maltsev term of \mathbf{G}_{k+1}/ρ_Q , thus $[q(c, c', c')]_{\rho_Q} = [c]_{\rho_Q}$, contradicting $[c]_{\rho_Q} \neq [c']_{\rho_Q}$. \square

The last tool is Theorem 8.1 from [10] (a shorter and cleaner proof is in [8, Theorem 3.5]), which is now usually referred to as the ‘‘Loop Lemma’’. We will use the following special case.

Lemma 4.11 (Loop Lemma). *Let $\mathbf{B} \in \text{HSP}(\mathbf{A})$ and $S \leq_{sd} \mathbf{B} \times \mathbf{B}$. If S is linked, then there exists $b \in B$ such that $(b, b) \in S$.*

4.2. Proof of Lemma 3.4

We require one more technical lemma.

Lemma 4.12. *Let $\mathbf{T} \leq \mathbf{D}^n$ and $\mathbf{f} = (f_1, \dots, f_n) \in F^n \cap T$. Then there exist $F_i \triangleleft_G \pi_i(\mathbf{T})$ with $f_i \in F_i \subseteq F$ (for each $i \in [n]$) such that the following holds. If \sim is a congruence of \mathbf{T} such that q is a Maltsev term of \mathbf{T}/\sim , then every \sim -block intersects $F_1 \times \dots \times F_n$.*

Proof. We define F_1, \dots, F_n inductively by $F_1 = \{f_1\}$, and

$$F_i = \begin{cases} \pi_i(T \cap (F_1 \times \dots \times F_{i-1} \times D^{n-i+1})) & \text{if this set is a subset of } F, \\ \{f_i\} & \text{otherwise.} \end{cases}$$

Each F_i is a Gumm-absorbing subuniverse of $\pi_i(\mathbf{T})$ by the Forced Absorption Lemma.

By induction on $i = 0, 1, \dots, n$, we show that each \sim -block intersects $F_1 \times \dots \times F_i \times D^{n-i}$. This claim is trivial for $i = 0$, so let $i > 0$ and suppose that the claim is true for $i - 1$. For any $\mathbf{a} \in T$ we set

$$C([\mathbf{a}]_{\sim}) = \pi_i([\mathbf{a}]_{\sim} \cap (F_1 \times \dots \times F_{i-1} \times D^{n-i+1})).$$

By the induction hypothesis, $C([\mathbf{a}]_{\sim})$ is nonempty for every $\mathbf{a} \in T$. We need to show that each $C([\mathbf{a}]_{\sim})$ intersects F_i .

If $C([\mathbf{a}]_{\sim}) \subseteq F$ for all \mathbf{a} , then $C([\mathbf{a}]_{\sim}) \subseteq F_i$ by the choice of F_i and we are done. Otherwise pick a \sim -block $[\mathbf{b}]_{\sim}$ and $e \in E \cap C([\mathbf{b}]_{\sim})$. We show that each $C([\mathbf{a}]_{\sim})$ intersects E . Indeed, take any $c \in C([\mathbf{a}]_{\sim})$ and consider

$$\begin{aligned} E \ni q^{\mathbf{A}}(c, e, e) &\in q^{\mathbf{A}}(C([\mathbf{a}]_{\sim}), C([\mathbf{b}]_{\sim}), C([\mathbf{b}]_{\sim})) \\ &\subseteq C(q^{\mathbf{T}/\sim}([\mathbf{a}]_{\sim}, [\mathbf{b}]_{\sim}, [\mathbf{b}]_{\sim})) = C([\mathbf{a}]_{\sim}). \end{aligned}$$

Here $q(c, e, e) \in E$ follows from the DDE Lemma, the inclusion follows from the compatibility of q with T , and the last equality from the assumption that q is a Maltsev term of \mathbf{T}/\sim .

Now $C([\mathbf{f}]_{\sim}) \cap E \neq \emptyset$ and of course $f_i \in C([f]_{\sim}) \cap F$, therefore, by the Minimality Lemma, $C([\mathbf{f}]_{\sim}) = D$. This allows us to show that each $C([\mathbf{a}]_{\sim})$ even contains F : Take any $f \in F$, any \mathbf{a} and any $d \in C([\mathbf{a}]_{\sim})$. We have

$$f = q(d, d, f) \in q(C([\mathbf{a}]_{\sim}), C([\mathbf{f}]_{\sim}), C([\mathbf{f}]_{\sim})) \subseteq C(q([\mathbf{a}]_{\sim}, [\mathbf{f}]_{\sim}, [\mathbf{f}]_{\sim})) = C([\mathbf{a}]_{\sim}). \quad \square$$

We are ready to prove Lemma 3.4. Recall the assumptions:

- $R \leq \mathbf{C}^2 \times \mathbf{D}^n$, where $\mathbf{C} \leq \mathbf{A}$,
- $\Delta_C \subseteq S$, where $S = \pi_{1,2}(R)$, and
- there exists $\mathbf{f} = (f_1, \dots, f_n) \in F^n$ such that the smooth part of the digraph $Q = \{(a_1, a_2) : (a_1, a_2, f_1, \dots, f_n) \in R\}$ is nonempty.

We aim to show that $(\Delta_C \times F^n) \cap R \neq \emptyset$.

First we make two adjustments to R and C so that we can make the following additional assumptions:

- $Q, S \leq_{sd} \mathbf{C} \times \mathbf{C}$, and
- S is linked.

The smooth part C' of Q is a subuniverse of \mathbf{C} since, as noted in Subsection 2.5, it is pp-definable from Q . Thus we can redefine $C := C'$ and $R := R \cap ((C')^2 \times D^n)$ (and S, Q accordingly) to satisfy the first of the additional assumptions. From $\Delta_C \subseteq S$ it follows that $\lambda_S = \rho_S$. Then we can take any λ_S -block C' and redefine C and R as before to satisfy the second assumption.

Let $T = \pi_{3,4,\dots,n+2}(R)$ and

$$R^b = \{((a_1, a_2), (a_3, \dots, a_{n+2})) : (a_1, \dots, a_{n+2}) \in R\} \leq_{sd} \mathbf{S} \times \mathbf{T},$$

thus R^b is essentially R regarded as a subset of $S \times T$.

Clearly, $\mathbf{f} \in F^n \cap T$, so we can apply Lemma 4.12 and obtain $F_1, \dots, F_n \subseteq F$ with $f_i \in F_i \triangleleft_G \pi_i(\mathbf{T})$ for each $i \in [n]$. Let

$$Q' = (T \cap (F_1 \times \dots \times F_n))^{-R^b}.$$

This is a superset of $Q \leq_{sd} \mathbf{C} \times \mathbf{C}$, in particular, $Q' \leq_{sd} \mathbf{C} \times \mathbf{C}$. Moreover, we have $T \cap (F_1 \times \dots \times F_n) \triangleleft_G \mathbf{T}$, so $Q' \triangleleft_G \mathbf{S}$ by the Forced Absorption Lemma. According to the Edge Absorption Lemma, the term q is Maltsev in $\mathbf{C}/\lambda_{Q'}$ and $\mathbf{C}/\rho_{Q'}$. Figure 6 depicts Q', S , and R^b .

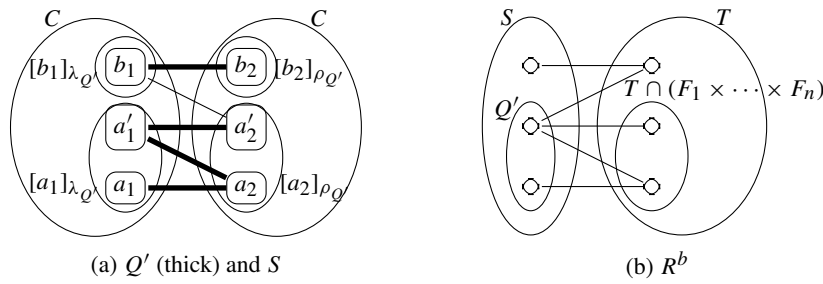


Fig. 6. The proof of Lemma 3.4.

Now we consider Q' , S , and R^b modulo $\lambda_{Q'}$ in the first coordinate and $\rho_{Q'}$ in the second coordinate:

$$\begin{aligned}
 Q'^q &= \{([a_1]_{\lambda_{Q'}}, [a_2]_{\rho_{Q'}}) : (a_1, a_2) \in Q'\}, \\
 S^q &= \{([a_1]_{\lambda_{Q'}}, [a_2]_{\rho_{Q'}}) : (a_1, a_2) \in S\}, \\
 R^{bq} &= \{(([a_1]_{\lambda_{Q'}}, [a_2]_{\rho_{Q'}}), (a_3, \dots, a_{n+2})) : (a_1, \dots, a_{n+2}) \in R\}.
 \end{aligned}$$

Since q is Maltsev in $\mathbf{C}/\lambda_{Q'}$ and $\mathbf{C}/\rho_{Q'}$, it is also Maltsev in $\mathbf{S}^q \leq \mathbf{C}/\lambda_{Q'} \times \mathbf{C}/\rho_{Q'}$. By the Rectangularity Lemma, R^{bq} is rectangular. Moreover, R^{bq} induces an isomorphism between $\pi_1(\mathbf{R}^{bq})/\lambda_{R^{bq}} = \mathbf{S}^q/\lambda_{R^{bq}}$ and $\pi_2(\mathbf{R}^{bq})/\rho_{R^{bq}} = \mathbf{T}/\rho_{R^{bq}}$. Therefore, since q is a Maltsev term in $\mathbf{S}^q/\lambda_{R^{bq}}$, q is also a Maltsev term in $\mathbf{T}/\rho_{R^{bq}}$. By the conclusion of Lemma 4.12, every block of $\rho_{R^{bq}}$ intersects $F_1 \times \dots \times F_n$. From this fact and rectangularity of R^{bq} it follows that $Q'^q = S^q$. Indeed, any $\mathbf{a} \in S^q$ is R^{bq} -adjacent to some $\mathbf{b} \in T$ and then to every element in $[\mathbf{b}]_{\rho_{R^{bq}}}$, in particular, to some element of $F_1 \times \dots \times F_n$. But $Q'^q = (F_1 \times \dots \times F_n)^{-R^{bq}}$ by definitions.

The proof will now be concluded using the Loop Lemma. As $\Delta_C \subseteq S$ and $Q'^q = S^q$, we have $([a]_{\lambda_{Q'}} \times [a]_{\rho_{Q'}}) \cap Q' \neq \emptyset$ for each $a \in C$. It follows that $\lambda_{Q'} = \rho_{Q'}$. The Loop Lemma applied to $Q' \cap (B \times B)$ for an arbitrarily chosen $\lambda_{Q'}$ -block B produces a pair $(a, a) \in Q' = (F_1 \times \dots \times F_n)^{-R^b}$. This witnesses $(\Delta_C \times F^n) \cap R \neq \emptyset$, as required.

4.3. Proof of Lemma 3.5

Lemma 3.5 will be proved by induction. The induction step will be based on the following lemma.

Lemma 4.13. *Let $R \leq \mathbf{D}^{n+1}$ be a relation such that*

$$R^{1=2} = \{(a_1, \dots, a_n) : (a_1, a_1, a_2, a_3, \dots, a_n) \in R\}$$

is a CTB-relation. Then either $\pi_{1,3,4,\dots,n+1}(R)$ or $\pi_{2,3,\dots,n+1}(R \cap (\{f\} \times D^n))$ for some $f \in F$ is a CTB-relation.

Proof. Let $S = \pi_{1,2}(R)$. Since $R^{1=2}$ is a CTB-relation, there exist $e_1, \dots, e_n \in E$ such that $(d_1, d_1, d_2, \dots, d_n) \in D^{n+1}$ is in R whenever $d_i = e_i$ for at least one i . We fix such elements e_1, \dots, e_n . Notice that $\Delta_D \subseteq S$.

The first step is to show that $R \cap F^{n+1} = \emptyset$. Assume the converse and take

$$(f_1, \dots, f_{n+1}) \in R \cap F^{n+1}.$$

As in the last subsection, let

$$Q = \{(a_1, a_2) : (a_1, a_2, f_3, \dots, f_{n+1}) \in R\}.$$

We will show that Q contains (f, f) with $f \in F$. This will contradict the assumption that $R^{1=2}$ is a CTB-relation.

Since $\pi_1(Q)$ contains f_1 and e_1 , we have $\pi_1(Q) = D$ by the Minimality Lemma. Similarly, $\pi_2(Q) = D$, therefore $Q \leq_{sd} \mathbf{D} \times \mathbf{D}$. Also notice that $Q \triangleleft_G S$ by the Forced Absorption Lemma.

We distinguish two cases.

Case 1. $S \subseteq E^2 \cup F^2$.

From $\Delta_D \subseteq S$ it follows that $[a]_{\lambda_S} = [a]_{\rho_S}$ for each $a \in D$. We restrict S and Q to the λ_S -block of f_1 :

$$F' = [f_1]_{\lambda_S} (= [f_2]_{\lambda_S}), \quad S' = S \cap (F' \times F'), \quad Q' = Q \cap (F' \times F')$$

Observe that $F' \subseteq F$ (since $S \subseteq E^2 \cup F^2$), $Q' \triangleleft_G S'$ (since $Q \triangleleft_G S$), $S' \leq_{sd} \mathbf{F}' \times \mathbf{F}'$, $Q' \leq_{sd} \mathbf{F}' \times \mathbf{F}'$ (since $Q \leq_{sd} \mathbf{D} \times \mathbf{D}$), and S' is linked. By the Edge Absorption Lemma, q is a Maltsev term in $\mathbf{F}'/\lambda_{Q'}$ and in $\mathbf{F}'/\rho_{Q'}$.

Let

$$\begin{aligned} R^b &= \{((a_1, a_2), (a_3, \dots, a_{n+2})) : (a_1, \dots, a_{n+2}) \in R\} \leq_{sd} \mathbf{S} \times \pi_{3, \dots, n+1}(\mathbf{R}), \\ R'^b &= R^b \cap (S' \times (S')^{+R^b}), \\ Q'^q &= \{([a_1]_{\lambda_{Q'}}, [a_2]_{\rho_{Q'}}) : (a_1, a_2) \in Q'\}, \\ S'^q &= \{([a_1]_{\lambda_{Q'}}, [a_2]_{\rho_{Q'}}) : (a_1, a_2) \in S'\}, \\ R'^{bq} &= \{(([a_1]_{\lambda_{Q'}}, [a_2]_{\rho_{Q'}}), (a_3, \dots, a_{n+2})) : ((a_1, a_2), (a_3, \dots, a_{n+2})) \in R^b\}. \end{aligned}$$

From $\{(e_1, e_1)\}^{+R^b} = \pi_{3, \dots, n+1}(R)$ (see the first paragraph of this proof) it follows that R^b is linked. By the Forced Absorption Lemma, $F' \triangleleft_G \mathbf{D}$, and using the same lemma again, $S' \triangleleft_G \mathbf{S}$. The Vertex Absorption Lemma now implies that R^b is linked, thus R'^{bq} is linked.

The term q is Maltsev in the projections of \mathbf{S}'^q , therefore it is a Maltsev term in \mathbf{S}'^q . Then, by the Rectangularity Lemma, R'^{bq} is rectangular, and since it is also linked, we get $R'^{bq} = S'^q \times \pi_{3, \dots, n}(R'^{bq})$. In particular, $Q'^q = \{(f_3, \dots, f_n)\}^{-R'^{bq}} = S'^q$. Now we apply the Loop Lemma in the same way as in the last paragraph of the last subsection and get $f \in F'$ with $(f, f) \in Q' \subseteq Q \subseteq F \times F$, a contradiction.

Case 2. $S \not\subseteq E^2 \cup F^2$.

In this case S contains (f, e) (or (e, f)) for some $e \in E, f \in F$ and also $(e, e) \in S$, therefore $e, f \in \{e\}^{-S}$ (or $e, f \in \{e\}^{+S}$) and then $\{e\}^{-S} = D$ (or $\{e\}^{+S} = D$) by the Minimality Lemma. It follows that S is linked. Recall that $Q \leq_{sd} \mathbf{D} \times \mathbf{D}$ and $Q \triangleleft_G \mathbf{S}$, so, by the Edge Absorption Lemma, q is a Maltsev term of \mathbf{D}/λ_Q .

If $\lambda_Q \subseteq E^2 \cup F^2$, then $q^{\mathbf{D}}([\!f\!]_{\lambda_Q}, [e]_{\lambda_Q}, [e]_{\lambda_Q}) \subseteq [\!f\!]_{\lambda_Q} \subseteq F$ for some (actually all) $e \in E$ and $f \in F$, contrary to the DDE Lemma. Thus $\lambda_Q \not\subseteq E^2 \cup F^2$. But then a λ_Q -block intersecting E and F must be the whole D by the Minimality Lemma, hence $\lambda_Q = D^2$. In other words, Q is linked.

Since $(f_1, f_2) \in Q$ and there is a Q -walk from f_2 to an element of E , we can find $f, f' \in F$ and $e \in E$ such that $(f, e), (f, f') \in Q$ (or $(e, f), (f', f) \in Q$). But then $\{f\}^{+Q} = D$ (or $\{f\}^{-Q} = D$) by the Minimality Lemma. In particular, $(f, f) \in Q$, a contradiction.

In both cases, we have proved that $R \cap F^{n+1} = \emptyset$. It may happen that $\pi_{1,3,\dots,n+1}(R)$ is a CTB-relation, in which case we are done. So, assume the converse, that is, $\pi_{1,3,\dots,n+1}(R) \cap F^n \neq \emptyset$. Take $f'_1, f'_3, \dots, f'_n \in F$ and $e'_2 \in D$ witnessing this, i.e. $(f'_1, e'_2, f'_3, \dots, f'_{n+1}) \in R$. Since $R \cap F^{n+1} = \emptyset$, we have $e'_2 \in E$. We will show that

$$Z = \pi_{2,3,\dots,n+1}(R \cap (\{f'_1\} \times D^n)) = \{(d_2, \dots, d_{n+1}) : (f'_1, d_2, \dots, d_{n+1}) \in R\}$$

is a CTB-relation by showing that $(d_2, \dots, d_{n+1}) \in Z$ whenever $d_2 = e'_2$ or $d_i = e_i$ for some $i \in \{3, \dots, n+1\}$.

Let $(d_3, \dots, d_{n+1}) \in D^{n-1}$ be any tuple such that $d_i = e_i$ for some $i \in \{3, \dots, n+1\}$. Let

$$Q'' = \{(a_1, a_2) : (a_1, a_2, d_3, d_4, \dots, d_{n+1}) \in R\}.$$

Clearly $Q'' \triangleleft_G S$ and $\Delta_D \subseteq Q''$. Since also $(f'_1, e'_2) \in S$, by the Directed Walking Lemma there is a directed Q'' -walk from f'_1 to $q(f'_1, e'_2, e'_2)$. As $q(f'_1, e'_2, e'_2) \in E$ by the DDE Lemma, it follows that there exist $f \in F$ and $e \in E$ such that $(f, e) \in Q''$. Now, repeated application of the Minimality Lemma gives $F \times D \subseteq Q''$: First, the set $\{f\}^{+Q''}$ contains e and f , so $\{f\}^{+Q''} = D$. Then, for every $e' \in E$, $\{e'\}^{-Q''}$ contains f and e' , thus $\{e'\}^{-Q''} = D$. Finally, for every $f' \in F$, $\{f'\}^{+Q''}$ contains f' and (every) $e' \in E$, therefore $\{f'\}^{+Q''} = D$.

The last paragraph proves that $(d_2, \dots, d_{n+1}) \in Z$ whenever $d_i = e_i$ for some $i \in \{3, \dots, n+1\}$. To finish the proof it is enough to show that each tuple of the form $(e'_2, d_3, \dots, d_{n+1})$ (where $d_i \in D$) is in Z , equivalently $(f'_1, e'_2, d_3, \dots, d_{n+1})$ is in R .

Let $(d_3, \dots, d_{n+1}) \in D^{n-1}$ be arbitrary. The subuniverse

$$\{a : (f'_1, e'_2, a, f'_4, \dots, f'_{n+1}) \in R\} \leq \mathbf{D}$$

contains e_3 and f'_3 , therefore it is equal to D by the Minimality Lemma. In particular, it contains d_3 . Then, the subuniverse

$$\{a : (f'_1, e'_2, d_3, a, f'_5, \dots, f'_{n+1}) \in R\}$$

contains e_4 and f'_4 , so it contains d_4 . In this way, eventually $(f'_1, e'_2, d_3, \dots, d_{n+1}) \in R$. This finishes the proof of Lemma 4.13. \square

We are ready to prove a slightly stronger version of Lemma 3.5: Let l_1, \dots, l_n be positive integers and $V \leq \mathbf{D}^{l_1+\dots+l_n}$ be such that

$$W = \{(a_1, \dots, a_n) : (\underbrace{a_1, \dots, a_1}_{l_1}, \underbrace{a_2, \dots, a_2}_{l_2}, \dots, \underbrace{a_n, \dots, a_n}_{l_n}) \in V\}$$

is a CTB-relation. Then for each $i \in [n]$ there exist $m_i \in [l_i]$ and $C_i^j \leq \mathbf{A}$, $j \in [l_i] \setminus \{m_i\}$, such that

$$U' = \{(a_1^{m_1}, \dots, a_n^{m_n}) : (a_1^1, \dots, a_1^{l_1}, a_2^1, \dots, a_2^{l_2}, \dots, a_n^1, \dots, a_n^{l_n}) \in V, \\ (\forall i \in [n])(\forall j \in [l_i], j \neq m_i) a_i^j \in C_i^j\}$$

is a CTB-relation.

The claim is proved by induction on $l_1 + \dots + l_n$. If $l_1 = \dots = l_n = 1$, then the claim is trivially true. Assume that $l_i > 1$ for some i , for simplicity, $l_1 > 1$. Let

$$R = \{(a'_1, a_1, a_2, \dots, a_n) : (\underbrace{a_1, a'_1, \dots, a'_1}_{l_1-1}, \underbrace{a_2, \dots, a_2}_{l_2}, \dots, \underbrace{a_n, \dots, a_n}_{l_n}) \in V\}.$$

Clearly $R^{1=2} = W$, so we can apply Lemma 4.13. If $\pi_{1,3,\dots,n+1}(R)$ is a CTB-relation, then we set $C_1^1 = D$, $l'_1 = l_1 - 1$, and $V' = \pi_{2,3,\dots,l_1+\dots+l_n}(V)$. If $\pi_{2,3,\dots,n+1}(R \cap (\{f\} \times D^n))$ is a CTB-relation, then we set $C_1^2 = C_1^3 = \dots = C_1^{l_1} = \{f\}$, $l'_1 = 1$ and $V' = \pi_{1,l_1+1,l_1+2,\dots,l_1+\dots+l_n}(V \cap (D \times \{f\}^{l_1-1} \times D^{l_2+\dots+l_n}))$. In both cases, the m_i s and the remaining C_i^j s are obtained by applying the induction hypothesis to l'_1, l_2, \dots, l_n and the relation V' . It is easy to see that U' is then a CTB-relation, as required.

4.4. Proof of Lemma 3.8

Assume that a CTB-relation R is defined by

$$R(x_1, \dots, x_n) \text{ iff } (\exists w_1, \dots, w_n) S_1(x_1, w_1) \wedge \dots \wedge S_n(x_n, w_n) \\ \wedge T_1(w_1, w_2) \wedge T_2(w_2, w_3) \wedge \dots \wedge T_{n-1}(w_{n-1}, w_n),$$

where $S_i \leq \mathbf{A}^2$ for each $i \in [n]$ and $T_i \leq \mathbf{A}^2$ for each $i \in [n - 1]$. Figure 7 shows an example of such a definition. We need to prove that $n < 2 \cdot 3^{|A|}$. Striving for a contradiction we assume the converse.

We will need only a part from the properties of the CTB-relation R , namely $R \leq \mathbf{D}^n$ and there are $f \in F, e_1, \dots, e_n \in E$ such that

$$(f, f, \dots, f) \notin R \quad \text{and} \quad (f, \dots, f, e_i, f, \dots, f) \in R,$$

where e_i is at the i -th position. We fix such elements f, e_1, \dots, e_n .

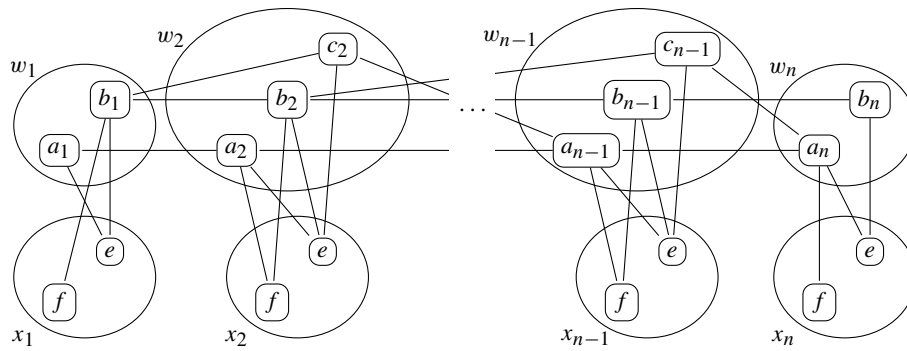


Fig. 7. A comb-formula defining $\{e, f\}^n \setminus \{(f, \dots, f)\}$.

The following terminology will be useful. An (i, j) -path from a to a' , where $i \leq j \in [n]$ and $a, a' \in A$, is a tuple $(a = a_i, a_{i+1}, \dots, a_j = a')$ of elements of A such that $(a_k, a_{k+1}) \in T_k$ for every $k \in \{i, \dots, j - 1\}$. Such an (i, j) -path is supported by (b_i, \dots, b_j) if $(b_k, a_k) \in S_k$ for every $k \in \{i, \dots, j\}$. Observe that a tuple (d_1, \dots, d_n) is in R iff there exists a $(1, n)$ -path supported by (d_1, \dots, d_n) . In Figure 7, the $(1, n)$ -path $(b_1, c_2, a_3, a_4, \dots, a_n)$ is supported by any tuple from $\{e, f\} \times \{e\} \times \{e, f\}^{n-2}$.

For each $i \in \{2, \dots, n - 1\}$ we define two subuniverses G_i, H_i of \mathbf{A} : $a \in G_i$ iff there exists a $(1, i)$ -path to a supported by (f, \dots, f) , and $a \in H_i$ iff there exists an (i, n) -path from a supported by (f, \dots, f) . The sets G_i and H_i are indeed subuniverses of \mathbf{A} , because they can be pp-defined from S_j s, T_j s and singletons. For every $i \in \{2, \dots, n - 1\}$, these subuniverses are nonempty (since $(f, \dots, f, e_{i+1}, f, \dots, f) \in R$ and $(f, \dots, f, e_{i-1}, f, \dots, f) \in R$) and disjoint (since $(f, \dots, f) \notin R$). In Figure 7, we have $G_i = \{b_i\}$, $H_i = \{a_i\}$.

Now we use the fact that the arity of R is large. There are $3^{|A|} - 2^{|A|+1} + 1$ ordered pairs of disjoint nonempty subsets of A and at least $n/2 - 1$ even integers $i \in \{2, \dots, n - 1\}$, so, since $n/2 - 1 \geq 3^{|A|} - 1 > (3^{|A|} - 2^{|A|+1} + 1)$, there must be two different even k, l such that $(G_k, H_k) = (G_l, H_l)$. In particular, there exist $k, l, 1 < k < l - 1 < n$, such that $G_k = G_l$ and $H_k = H_l$. We fix such k, l and denote

$$e = e_{k+1}, \quad G = G_k = G_l, \quad H = H_k = H_l.$$

For each $d \in D$ we define a subuniverse $Q(d) \leq \mathbf{A}^2$ so that $(a, b) \in Q(d)$ iff there exists a (k, l) -path from a to b supported by (f, d, f, f, \dots, f) . In Figure 7, if $a_2 = a_{n-1} = a$ and $b_2 = b_{n-1} = b$, then we can take $k = 2, l = n - 1$ and get $G = \{b\}$, $H = \{a\}$, $Q(f) = \{(a, a), (b, b)\}$, $Q(e) = Q(f) \cup \{(b, a)\}$.

Let $Q \leq \mathbf{A}^2$ be the union of $Q(d)$ over $d \in D$. The sets $Q(d)$ and Q are indeed subuniverses of \mathbf{A}^2 since they can be pp-defined from subuniverses of \mathbf{A} . Moreover, they have the following algebraic properties:

(Alg 1) $Q(d) \triangleleft_G Q$ for every $d \in D$.

(Alg 2) For any term t , say of arity z , and any $d_1, \dots, d_z \in D$,

$$t^{\mathbf{A}^2}(Q(d_1), \dots, Q(d_z)) \subseteq Q(t^{\mathbf{A}}(d_1, \dots, d_z)).$$

(Alg 3) If $(a, b) \in Q(d_1) \cap Q(d_2)$, then $(a, b) \in Q(d)$ for any d in the subuniverse of \mathbf{D} generated by d_1, d_2 .

All three properties are consequences of a simple observation: If t is a z -ary term and, for each $i \in [z]$, (a_k^i, \dots, a_l^i) is a (k, l) -path supported by (d_k^i, \dots, d_l^i) , then the sequence

$$(t(a_k^1, \dots, a_k^z), \dots, t(a_l^1, \dots, a_l^z))$$

is a (k, l) -path supported by $(t(d_k^1, \dots, d_k^z), \dots, t(d_l^1, \dots, d_l^z))$.

We regard Q and $Q(d)$ s as digraphs (not as bipartite graphs). They satisfy the following:

(Dig 1) $Q(e)$ contains an edge from G to H .

(Dig 2) $Q(f) \cap (G \times G)$ has no sources in G , $Q(f) \cap (H \times H)$ has no sinks in H .

(Dig 3) $Q(f)$ has no edge from G to outside of G and no edge from outside of H to H .

There exists a $(1, n)$ -path (a_1, \dots, a_n) supported by $(f, \dots, f, e_{k+1} = e, f, \dots, f)$. Then $a_k \in G, a_l \in H$, and $(a_k, a_l) \in Q(e)$ by definitions, and (Dig 1) is proved. To prove the first part of (Dig 2), consider a vertex $a \in G = G_l$. By the definition of G_l , there exists a $(1, l)$ -path $(a_1, \dots, a_{l-1}, a_l = a)$ supported by (f, \dots, f) , thus $a_k \in G_k = G$ and $(a_k, a) \in Q(f)$. This shows that $Q(f) \cap (G \times G)$ has no sources in G . Similarly, $Q(f) \cap (H \times H)$ has no sinks in H . Finally, if $a \in G = G_k$ and $(a, b) \in Q(f)$, then there exists a $(1, k)$ -path $(a_1, \dots, a_k = a)$ supported by (f, \dots, f) and a (k, l) -path $(a = a_k, a_{k+1}, \dots, a_l = b)$ supported by (f, \dots, f) . Then (a_1, \dots, a_l) is a $(1, l)$ -path supported by (f, \dots, f) , hence $a_l = b \in G_l = G$. This proves the first part of (Dig 3); the second part is analogous.

Let r_1, r_2 be integers such that

$$r_1, r_2 \geq |A|, \quad r_1 + r_2 + 1 = r_3 |A|!, \quad \text{where } r_3 \geq |A|.$$

For $d \in D$, set

$$Q'(d) = \underbrace{Q(f) \circ \dots \circ Q(f)}_{r_1} \circ Q(d) \circ \underbrace{Q(f) \circ \dots \circ Q(f)}_{r_2}$$

and similarly

$$Q' = \underbrace{Q(f) \circ \dots \circ Q(f)}_{r_1} \circ Q \circ \underbrace{Q(f) \circ \dots \circ Q(f)}_{r_2} = \bigcup_{d \in D} Q'(d).$$

It is easily seen that the primed versions of properties (Alg 1–3) and (Dig 3) are satisfied. Moreover,

(Dig 1') $Q'(e)$ contains an “ f -looped edge” from G to H , that is, there exist $g \in G$ and $h \in H$ such that $(g, h) \in Q'(e)$ and $(g, g), (h, h) \in Q'(f)$, and

(Dig 4) $Q'(f)$ is transitive.

To prove these properties, consider the $|A|!$ -fold composition of $Q(f)$ with itself:

$$Q'' = \underbrace{Q(f) \circ \cdots \circ Q(f)}_{|A|!}.$$

Clearly, $Q'(f)$ is the r_3 -fold composition of Q'' with itself. From the Dirichlet principle it follows that every vertex $a \in A$ which is contained in a closed $Q(f)$ -walk (equivalently, in a closed Q'' -walk) is contained in a closed $Q(f)$ -walk of length at most $|A|$, and so $(a, a) \in Q''$. For the same reason, every Q'' -walk from a to b of length at least $|A|$ contains an element in a closed Q'' -walk. Since this element has a Q'' -loop, we can modify the walk so as to get a Q'' -walk from a to b of any length greater than or equal to $|A|$. In particular, if two elements a and b are connected by a Q'' -walk of length $2r_3$, then they are connected by a Q'' -walk of length r_3 . Therefore, $Q'(f)$ is transitive.

To prove (Dig 1'), take $(g', h') \in Q(e)$, $g' \in G$ and $h' \in H$ guaranteed by (Dig 1). By (Dig 2) and the Dirichlet principle, we can find a $Q(f)$ -walk of length r_1 from a vertex $g \in G$ contained in a closed $Q(f)$ -walk to the vertex g' . Similarly, there exists a $Q(f)$ -walk of length r_2 from h' to a vertex $h \in H$ which is contained in a closed $Q(f)$ -walk. It follows that $(g, h) \in Q'(e)$ and $(g, g), (h, h) \in Q'(f)$.

Armed with all these digraphs and their properties, we are ready to finish the proof. We would like to have the situation shown in Figure 8.

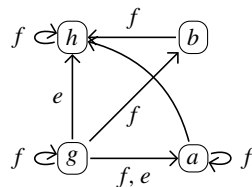


Fig. 8. The final argument for the proof of Lemma 3.8. Label d means that the edge is in $Q'(d)$.

We start by fixing $g \in G$ and $h \in H$ as in (Dig 1') and setting

$$a = q(g, h, h), \quad b = q(h, a, a).$$

Since $(g, g), (h, h) \in Q'(f)$, $(g, h) \in Q'(e) \subseteq Q'$, and $Q'(f) \triangleleft_G Q'$ (see (Alg 1)), there is a directed walk in $Q'(f)$ from g to $q(g, h, h) = a$ by the Directed Walking Lemma. From (Dig 4) it follows that $(g, a) \in Q'(f)$. The edge $(g, a) = q((g, g), (g, h), (g, h))$ also lies in $Q'(q(f, e, e))$ by (Alg 2). According to (Alg 3), (g, a) is in $Q(d)$ for any d in the subuniverse D' of \mathbf{D} generated by f and $q(f, e, e)$. Since $q(f, e, e) \in E$ by the DDE Lemma, we get $D' = D$ by the Minimality Lemma. In particular $(g, a) \in Q'(e)$.

Using (Alg 2), $(g, h) \in Q'(e)$, $(g, a) \in Q'(e) \cap Q'(f)$, and $q(e, e, f) = f$, we get

$$(g, b) = (g, q(h, a, a)) = q((g, h), (g, a), (g, a)) \in Q'(q(e, e, f)) = Q'(f).$$

Thus $b \in G$ by (Dig 3).

On the other hand, $Q'(f) \triangleleft_G Q'$ by (Alg 1), $(a, a) = q((g, g), (h, h), (h, h)) \in Q'(f)$, $(h, h) \in Q'(f)$, and $(a, h) = q((g, h), (h, h), (h, h)) \in Q'$, so we can apply the second part of the Directed Walking Lemma to obtain a $Q'(f)$ -walk from $b = q(h, a, a)$ to h . Then $b \in H$ by (Dig 3). We have proved $b \in G \cap H = \emptyset$, a contradiction.

5. Consequences and open problems

CSP

Let \mathbb{A} be a relational structure with a finite universe. The *constraint satisfaction problem over \mathbb{A}* , denoted $\text{CSP}(\mathbb{A})$, is the decision problem asking whether a given input pp-sentence over \mathbb{A} is true. The main open problem in the area is the dichotomy conjecture [26] stating that, for every \mathbb{A} , $\text{CSP}(\mathbb{A})$ is tractable (i.e., solvable in polynomial time) or NP-complete. It is known [20] that the computational complexity of $\text{CSP}(\mathbb{A})$ depends only on the variety generated by the algebra $\mathbf{A} = (A; \text{Pol}(\mathbb{A}))$. A precise borderline between polynomial solvability and NP-completeness was conjectured and the hardness part proved in [20]. What remains is to show that $\text{CSP}(\mathbb{A})$ is tractable whenever \mathbf{A} satisfies a nontrivial idempotent Maltsev condition. This was verified for algebras in meet semi-distributive varieties [9] and for algebras \mathbf{A} with few subpowers [32]. The next natural step was to concentrate on the class CM. However, Theorem 1.1 shows that there are no new relational structures to consider and we can use the “few subpowers algorithm” whenever $\mathbf{A} \in \text{CM}$.

This does not mean that the CSPs with $\mathbf{A} \in \text{CM}$ are solved in a satisfactory way. For instance, the few subpowers algorithm heavily uses algebraic operations similar to cube terms. It could be useful to have a different algorithm based on (directed) Gumm terms, or even an algorithm which does not use any operations at all. One reason is that it could make the proof of Theorem 1.1 shorter and cleaner—see Subsection 3.7. More importantly, it could lead to more general tractability results: an algorithm for $\mathbf{A} \in \text{CD}$ from [7], which is different from algorithm for $\mathbf{A} \in \text{NU}$ from [26], was an important step toward the general result for algebras in meet semi-distributive varieties from [9].

Open Problem 5.1. Find a different algorithm for solving $\text{CSP}(\mathbb{A})$, where $(A; \text{Pol}(\mathbb{A})) \in \text{CM}$.

A natural next step is to consider algebras in varieties omitting types **1** and **5** from tame congruence theory [31]. Indeed, $\mathbf{A} \in \text{CM}$ iff \mathbf{A} is in a variety which omits types **1** and **5** and has “no tails”.

Primitive positive formula comparison

Let again \mathbb{A} be a relational structure and $\mathbf{A} = (A; \text{Pol}(\mathbb{A}))$ the corresponding algebra. The *pp-formula equivalence problem over \mathbb{A}* asks whether two given pp-formulas over \mathbb{A} define the same relations. A similar problem is the *pp-formula containment problem over \mathbb{A}* that asks whether the first formula defines a subrelation of the relation defined by the second formula.

In [17], the computational complexity of these problems is completely resolved, showing a P/coNP-complete/ Π_2^P -complete trichotomy, modulo the conjectured borderline for CSPs and the Edinburgh conjecture: If \mathbf{A} does not satisfy any nontrivial idempotent Maltsev condition, then the problems are Π_2^P -complete. If this is not the case and $\text{CSP}(\mathbb{A})$ is tractable, then the problems are in coNP. If \mathbf{A} is not in CM, then these problems are coNP-hard. Finally, if \mathbf{A} has few subpowers, then both problems are solvable in polynomial time (this was already proved in [16]). Therefore, Theorem 1.1 gives the P/coNP-hard dichotomy.

Learnability

The aim of the *learning problem for pp-formulas over \mathbb{A}* is to learn (in some sense) the relation defined by an unknown pp-formula, given access to an oracle which can answer simple queries like “is the tuple \mathbf{a} in the relation?” (see [32, 22] for precise definitions of the learning models). A positive learnability result for \mathbf{A} with few subpowers was given in [32] and a negative learnability result for \mathbf{A} outside the class CM was proved in [22] (under some standard cryptographic assumptions). Theorem 1.1 thus closes the gap.

Deciding few subpowers for relational structures

By Theorem 1.1, the problem FEWSUB of deciding whether a given relational structure \mathbb{A} determines an algebra \mathbf{A} with few subpowers is equivalent to the problem of deciding whether \mathbf{A} is in CM (for a given \mathbb{A}). It is quite easy to see that the latter problem is decidable in exponential time: we can search among the ternary operations for (directed) Gumm terms (see [28] for details). Recently, Kazda (personal communication) has observed that FEWSUB is in NP since local characterizations of congruence modularity (as in [28]) can be encoded as CSP instances. However, the exact complexity is open:

Open Problem 5.2. *Determine the computational complexity of FEWSUB.*

We remark that Kazda’s idea and [28] place the corresponding problem for NU (or CD) in the class P.

Algebraic questions

The results in [3, 44] give upper bounds for the least arity of a near-unanimity operation of $\mathbf{A} \in \text{NU}$ depending on $|A|$ and the maximum arity of a relation in \mathbb{A} (the bounds in [44] are tighter for nonbinary structures; this is caused by using the reduction to binary structures via [3, Proposition 3.1]). Examples essentially achieving these bounds were given in [44]. It is not clear how to obtain reasonable bounds for arities of cube terms from the results in this paper.

Open Problem 5.3. *Find a reasonable (or even essentially optimal) upper bound for the least arity of a cube term of $\mathbf{A} = (A; \text{Pol}(\mathbb{A}))$ depending on $|A|$ and the maximum arity of a relation in \mathbb{A} .*

Corollary 1.2 says, schematically, “ $\text{CM} + \text{finitely related} \Leftrightarrow \text{FS}$ ”, and we also know that “ $\text{CD} + \text{finitely related} \Leftrightarrow \text{NU}$ ” [3]. Are there other instances of this phenomenon?

Open Problem 5.4. *For which (important) Maltsev conditions P is there a Maltsev condition Q such that “ $P + \text{finitely related} \Leftrightarrow Q$ ”, or at least a condition Q stronger than P such that “ $P + \text{finitely related} \Rightarrow Q$ ”?*

The equivalence “ $\text{NU} \Leftrightarrow \text{CD}$ ” for finitely related finite algebras is generalized to “absorption \Leftrightarrow Jónsson-absorption” for finitely related finite algebras in [4]. Moreover, the difference between Jónsson-absorption and absorption for general finite algebras is captured in the main result of [5] which roughly says “absorption \Leftrightarrow Jónsson-absorption + no bad cube term blockers”. Are there analogues of these results for (directed) Gumm-absorption instead of Jónsson-absorption?

Open Problem 5.5. *Does there exist a useful notion of “cube-absorption” so that “cube-absorption \Leftrightarrow Gumm-absorption” is true for all finitely related finite algebras? Is there an analogue to “absorption \Leftrightarrow Jónsson-absorption + no bad cube term blockers” for cube-absorption?*

Acknowledgments. Research supported by the Grant Agency of the Czech Republic under the grant No. 13-01832S.

References

- [1] Aichinger, E., Mayr, P., McKenzie, R.: On the number of finite algebraic structures. *J. Eur. Math. Soc.* **16**, 1673–1686 (2014) [Zbl 06363150](#) [MR 3262454](#)
- [2] Baker, K., Pixley, A.: Polynomial interpolation and the Chinese remainder theorem for algebraic systems. *Math. Z.* **143**, 165–174 (1975) [Zbl 0292.08004](#) [MR 0371782](#)
- [3] Barto, L.: Finitely related algebras in congruence distributive varieties have near unanimity terms. *Canad. J. Math.* **65**, 3–21 (2013) [Zbl 1283.08009](#) [MR 3004455](#)
- [4] Barto, L., Bulín, J.: Deciding absorption in relational structures. *Algebra Universalis* **78**, 3–18 (2017) [Zbl 06787888](#) [MR 3687194](#)
- [5] Barto, L., Kazda, A.: Deciding absorption. *Int. J. Algebra Comput.* **26**, 1033–1060 (2016) [Zbl 1352.08006](#) [MR 3536445](#)
- [6] Barto, L., Kozik, M.: Congruence distributivity implies bounded width. *SIAM J. Comput.* **39**, 1531–1542 (2009) [Zbl 1205.68529](#) [MR 2580538](#)
- [7] Barto, L., Kozik, M.: Constraint satisfaction problems of bounded width. In: *FOCS’09, Proc. 50th Symposium on Foundations of Computer Science*, 595–603 (2009) [Zbl 1292.68088](#) [MR 2648455](#)
- [8] Barto, L., Kozik, M.: Absorbing subalgebras, cyclic terms, and the constraint satisfaction problem. *Logic. Methods Computer Sci.*, Special issue: Selected papers of the Conference “Logic in Computer Science (LICS) 2010”, 27 pp. (2012/14) [Zbl 1239.08002](#) [MR 3285464](#)
- [9] Barto, L., Kozik, M.: Constraint satisfaction problems solvable by local consistency methods. *J. ACM* **61**, 3:1–3:19 (2014) [Zbl 1295.68126](#) [MR 3167919](#)
- [10] Barto, L., Kozik, M., Niven, T.: The CSP dichotomy holds for digraphs with no sources and no sinks (a positive answer to a conjecture of Bang–Jensen and Hell). *SIAM J. Comput.* **38**, 1782–1802 (2008/09) [Zbl 1191.68460](#) [MR 2476276](#)

- [11] Barto, L., Kozik, M., Stanovský, D.: Mal'tsev conditions, lack of absorption, and solvability. *Algebra Universalis* **74**, 185–206 (2015) [Zbl 1323.08004](#) [MR 3374664](#)
- [12] Bergman, C.: *Universal Algebra: Fundamentals and Selected Topics*. Pure Appl. Math., CRC Press, Boca Raton, FL (2012) [Zbl 1229.08001](#) [MR 2839398](#)
- [13] Berman, J., Idziak, P., Marković, P., McKenzie, R., Valeriote, M., Willard, R.: Varieties with few subalgebras of powers. *Trans. Amer. Math. Soc.* **362**, 1445–1473 (2009) [Zbl 1190.08004](#) [MR 2563736](#)
- [14] Birkhoff, G.: On the structure of abstract algebras. *Proc. Cambridge Philos. Soc.* **31**, 433–454 (1935) [JFM 61.1026.07](#) [MR 2052301](#)
- [15] Bodnarčuk, V. G., Kalužnin, L. A., Kotov, V. N., Romov, B. A.: Galois theory for Post algebras. I, II. *Kibernetika (Kiev)* **1969**, no. 3, 1–10, and no. 5, 1–9 [Zbl 0212.31802](#) [MR 0300895](#)
- [16] Bova, S., Chen, H., Valeriote, M.: On the expression complexity of equivalence and isomorphism of primitive positive formulas. *Theory Computing Systems* **50**, 329–353 (2012) [Zbl 1288.68080](#) [MR 2875302](#)
- [17] Bova, S., Chen, H., Valeriote, M.: Generic expression hardness results for primitive positive formula comparison. *Information and Computation* **222**, 108–120 (2013) [Zbl 1295.68121](#) [MR 3000965](#)
- [18] Bulatov, A., Chen, H., Dalmau, V.: Learning intersection-closed classes with signatures. *Theoret. Computer Sci.* **382**, 209–220 (2007) [Zbl 1127.68044](#) [MR 2348225](#)
- [19] Bulatov, A., Dalmau, V.: A simple algorithm for Mal'tsev constraints. *SIAM J. Comput.* **36**, 16–27 (2006) [Zbl 1112.08002](#) [MR 2231638](#)
- [20] Bulatov, A., Jeavons, P., Krokhin, A.: Classifying the complexity of constraints using finite algebras. *SIAM J. Comput.* **34**, 720–742 (2005) [Zbl 1071.08002](#) [MR 2137072](#)
- [21] Burris, S. N., Sankappanavar, H. P.: *A Course in Universal Algebra*. Grad. Texts in Math. 78, Springer, New York (1981) [Zbl 0478.08001](#) [MR 0648287](#)
- [22] Chen, H., Valeriote, M.: Learnability of solutions to conjunctive queries: the full dichotomy. in: *Proc. 28th Conference on Learning Theory*, 326–327 (2015)
- [23] Dalmau, V.: Generalized majority-minority operations are tractable. *Logic. Methods Computer Sci.* **2**, 4:1, 14 pp. (2006) [Zbl 1127.68039](#) [MR 2295778](#)
- [24] Davey, B. A., Jackson, M. G., Pitkethly, J. G., Szabó, C.: Finite degree: algebras in general and semigroups in particular. *Semigroup Forum* **83**, 89–110 (2011) [Zbl 1243.08003](#) [MR 2824076](#)
- [25] Dolinka, I.: Finite regular bands are finitely related. *Bull. Austral. Math. Soc.* **87**, 1–9 (2013) [Zbl 1275.20061](#) [MR 3011935](#)
- [26] Feder, T., Vardi, M. Y.: The computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory. *SIAM J. Comput.* **28**, 57–104 (1998) [Zbl 0914.68075](#) [MR 1630445](#)
- [27] Freese, R., McKenzie, R.: *Commutator Theory for Congruence Modular Varieties*. London Math. Soc. Lecture Note Ser. 125, Cambridge Univ. Press, Cambridge (1987) [Zbl 0636.08001](#) [MR 0909290](#)
- [28] Freese, R., Valeriote, M. A.: On the complexity of some Maltsev conditions. *Int. J. Algebra Comput.* **19**, 41–77 (2009) [Zbl 1172.08004](#) [MR 2494469](#)
- [29] Geiger, D.: Closed systems of functions and predicates. *Pacific J. Math.* **27**, 95–100 (1968) [Zbl 0186.02502](#) [MR 0234893](#)
- [30] Gumm, H. P.: Congruence modularity is permutability composed with distributivity. *Arch. Math. (Basel)* **36**, 569–576 (1981) [Zbl 0465.08005](#) [MR 0629294](#)
- [31] Hobby, D., McKenzie, R.: *The Structure of Finite Algebras*. Contemp. Math. 76, Amer. Math. Soc., Providence, RI (1988) [Zbl 0721.08001](#) [MR 0958685](#)

- [32] Idziak, P., Marković, P., McKenzie, R., Valeriotte, M., Willard, R.: Tractability and learnability arising from algebras with few subpowers. In: Proc. 22 Annual IEEE Symposium on Logic in Computer Science (LICS 2007), IEEE Computer Soc. Press, 213–222 (2007)
- [33] Kazda, A., Kozik, M., McKenzie, R., Moore, M.: Absorption and directed Jónsson terms. Submitted (2015)
- [34] Kearnes, K., Marković, P., McKenzie, R.: Optimal strong Malcev conditions for omitting type 1 in locally finite varieties. *Algebra Universalis* **72**, 91–100 (2014) [Zbl 1305.08008](#) [MR 3229955](#)
- [35] Kearnes, K. A., Szendrei, Á.: Clones of algebras with parallelogram terms. *Int. J. Algebra Comput.* **22**, 1250005, 30 pp. (2012) [Zbl 1286.08002](#) [MR 2900858](#)
- [36] Kozik, M., Krokhin, A., Valeriotte, M., Willard, R.: Characterizations of several Maltsev conditions. *Algebra Universalis* **73**, 205–224 (2015) [Zbl 1319.08002](#) [MR 3350327](#)
- [37] Marković, P., Maróti, M., McKenzie, R.: Finitely related clones and algebras with cube terms. *Order* **29**, 345–359 (2012) [Zbl 1310.08002](#) [MR 2926316](#)
- [38] Marković, P., McKenzie, R.: Few subpowers, congruence distributivity and near-unanimity terms. *Algebra Universalis* **58**, 119–128 (2008) [Zbl 1155.08001](#) [MR 2386525](#)
- [39] Maróti, M., McKenzie, R.: Existence theorems for weakly symmetric operations. *Algebra Universalis* **59**, 463–489 (2008) [Zbl 1186.08003](#) [MR 2470592](#)
- [40] Mayr, P.: On finitely related semigroups. *Semigroup Forum* **86**, 613–633 (2013) [Zbl 1278.20077](#) [MR 3053784](#)
- [41] Siggers, M. H.: A strong Malcev condition for locally finite varieties omitting the unary type. *Algebra Universalis* **64**, 15–20 (2010) [Zbl 1216.08002](#) [MR 2745482](#)
- [42] Szendrei, Á.: Clones in Universal Algebra. *Sém. Math. Sup. 99*, Presses Univ. de Montréal (1986) [Zbl 0603.08004](#) [MR 0859550](#)
- [43] West, D. B.: *Introduction to Graph Theory*. 2nd ed., Prentice Hall (2005) [Zbl 1121.05304](#)
- [44] Zhuk, D. N.: The existence of a near-unanimity function is decidable. *Algebra Universalis* **71**, 31–54 (2014) [Zbl 1350.08001](#) [MR 3162419](#)