A NEW LINE OF ATTACK ON THE DICHOTOMY CONJECTURE

GÁBOR KUN AND MARIO SZEGEDY

E-mail: kungaborcs.elte.hu, szegedycs.rutgers.edu.

The first author's research was supported by the "Lendület: Groups and Graphs" Grant, by the Hungarian National Foundation for Scientific Research (OTKA) Grant No. PD 109731, ERC Advanced Research Grant No. 227701 and by Marie Curie IIF Fellowship Grant No. 627476, while the second author is supported by NSF expedition grant 0832787.

Abstract. The well known dichotomy conjecture of Feder and Vardi states that for every finite family Γ of constraints CSP(Γ) is either polynomially solvable or NP-hard. Bulatov and Jeavons reformulated this conjecture in terms of the properties of the algebra $Pol(\Gamma)$, where the latter is the collection of those n-ary operations ($n = 1, 2, \ldots$) that keep all constraints in Γ invariant. We show that the algebraic condition boils down to whether there are arbitrarily resilient functions in $Pol(\Gamma)$. Equivalently, we can express this in the terms of the PCP theory: CSP(Γ) is NP-hard iff every long code test created from Γ that passes with zero error admits only juntas[1]. Then, using this characterization and a result of Dinur, Friedgut and Regev, we give an entirely new and transparent proof to the Hell-Nešetřil theorem, which states that for a simple, connected and undirected graph H, the problem CSP(H) is NP-hard if and only if H is non-bipartite.

We also introduce another notion of resilience (we call it strong resilience), and we use it in the investigation of CSP problems that 'do not have the ability to count.' We show that CSP problems without the ability to count are exactly the ones with strongly resilient term operations. This gave already a handier tool to attack the conjecture that CSP problems without the ability to count have bounded width, or equivalently, that they can be characterized by existential k-pebble games: Barto and Kozik already proved this conjecture using a variant of our characterization. This is considered a major step towards the resolution of the dichotomy conjecture.

Finally, we show that a yet stronger notion of resilience, when the term operation is asymptotically constant, holds for the class of *width one* CSPs.

What emerges from our research, is that certain important algebraic conditions that are usually expressed via identities have equivalent analytic definitions that rely on asymptotic properties of term operations.

[1]For us "junta" means that a constant number of the variables have constant influence on the outcome. Also, Γ needs to be a core (See Definition 8).

Contents

1. Introduction 3
1.1. Robust approximation 7
2. A connection between two theories 7
3. More on the algebraic theory 9
4. Term Operations of Algebras 12
5. An analytic look at term operations 14
6. Asymptotic resilience 15
7. The Hell-Nešetřil theorem 19
8. Subclasses of CSPs 22
8.1. Bounded width classes 23
9. Strong Resilience 25
9.1. Polynomials 26
9.2. Composition 27
9.3. Proof of $(I) \to (II)$ 28
9.4. Proof of $(II) \to (III)$ 32
9.5. Proof of $(III) \to (I)$ 35
10. Width One, asymptotic invariance and symmetric functions 35
11. Asymptotic Invariance 37
References 42
12. Appendix 45
12.1. Unique Games 45
12.2. Convergence 46

1. Introduction

Constraint satisfaction problems (CSP) are the pinnacles in NP not only because they have multiple interpretations in logic, combinatorics, and complexity theory, but also for their immense popularity in various branches of science and engineering, where they are looked at as a versatile language for phrasing search problems. This said, it is even more remarkable that some basic complexity questions about them remain unanswered.

To a finite domain D, variables $\{x_1, x_2, \ldots\}$ ranging in D, and a set Γ of finitary relations on D we can associate a problem CSP(Γ), whose instances consist of a finite set of *constraints* of the form $(x_{i_1}, \ldots, x_{i_k}) \in R_j$ for some $R_j \in \Gamma$. The size of the instance (usually denoted by n) is by definition the number of different variables involved in its constrains.

As one might expect, for the tractability of CSP(Γ) the relations in Γ matter. For instance, general Boolean CSPs are NP-hard, but if all constraints are Horn clauses (i.e. disjunctions of literals, at most one of which is negative), then the problem is polynomially solvable. Other polynomially solvable cases include linear equations over finite fields and the set of all Boolean constraints that involve at most two variables.

The central question of the field is how the complexity of CSP(Γ) depends on Γ. Due to a beautiful result of Schaefer [51] we know, that in the Boolean case CSP(Γ) is either NP-hard or polynomial time solvable for every Γ. His *Dichotomy Theorem* also gives a full description of the polynomial time solvable families.

A fundamental question, raised by Feder and Vardi [26], asks if this theorem generalizes for arbitrary finite domain. Their *Dichotomy Conjecture* would imply the dichotomy of Monotone Monadic SNP ([26, 38], see also [39]). This is perhaps the largest natural subclass of NP expected to have dichotomy. That the entire class NP does not have dichotomy (unless P=NP) was proved by Ladner [40].

In [26] it is established that it is sufficient to settle the dichotomy conjecture when Γ contains a single binary relation, i.e. a directed graph, H. With a slight abuse of notation we denote this problem by CSP(H). A problem instance now simply becomes a directed graph G whose vertices we want to map to the vertices of H such that edges go into edges. This is a graph homomorphism problem. What if G is undirected? In this case dichotomy holds by a pioneering theorem due to Hell and Nešetřil (1990):

Theorem 1 (Hell-Nešetřil). *Assume that H is a simple, connected, undirected graph. Then CSP(H) is polynomial time solvable if and only if H is bipartite. Otherwise CSP(H) is NP-complete.*

Remark 2. The graph homomorphic view can be extended to arbitrary *relational structures*. Relational structures have a *type*, i.e. a list of relational names with associated arities. A relational structure of type $\mathcal{R}$ is an ordered pair $\mathbf{R} = \langle D, \Gamma \rangle$, where D is a non-empty set and Γ is a family of relations with names and associated arities as required by $\mathcal{R}$. Any CSP problem can be cast as a homomorphism problem $\langle E, \Upsilon \rangle \stackrel{?}{\to} \langle D, \Gamma \rangle$ between two relational structures of the same type. We refer the reader interested in the homomorphic view to an excellent survey written by Hell and Nešetřil, which also puts our current result into that context [30]. We, for the most part, stick to the language that is more familiar to computer scientists, which talks about constraints and assignments.

There is a beautiful algebraic theory due to Jeavons and his coauthors [16, 17, 33, 13, 14], that looks at maps from D^m to D ($m = 1, 2, \ldots$), which keep all relations in Γ invariant (said to be *compatible* with Γ). The set of these *compatible operations* is denoted by $Pol(\Gamma)$. The theory heavily relies on the fact that a composition $f(g_1, \ldots, g_m)$ of operators g_i that are compatible with Γ, is also compatible with Γ, hence $Pol(\Gamma)$ is closed under composition. Finally, to apply the tools of algebra, $Pol(\Gamma)$ is often viewed as an *algebra* (see Section 4).

We can also look at these operations in an entirely different way. For fixed m the condition that $f : D^m \to D$ keeps all relations in Γ can be interpreted so that f passes the long code test associated with Γ with zero error.

This dual interpretation of $Pol(\Gamma)$ allows us to connect the algebraic theory of CSPs with Fourier analytic techniques that were successfully used in the theory of probabilistically checkable proofs.

To demonstrate the strong interaction between the theories we reprove the theorem of Hell and Nešetřil in a transparent way. We rely on theorems of Bulatov and Jeavons as well as on the Fourier analytic results of Dinur, Friedgut and Regev.

Although PCP testing and the algebraic theory of $Pol(\Gamma)$ are apparently related, to exploit this relation we have to find a third "theory," which connects the two. This is our main contribution. We give here an example to the kind of issues the new study deals with.

Consider the following algebras on universe $D_2 = \{0, 1\}$, each with a single trinary operation:

$$
\begin{array}{lcll}
\mathbf{A}_{proj} & = & \langle D_2, \pi \rangle, & \text{where } \pi(a,b,c) = a \\
\mathbf{A}_{\oplus} & = & \langle D_2, \oplus \rangle, & \text{where } \oplus(a,b,c) = a+b+c \mod 2 \\
\mathbf{A}_{maj} & = & \langle D_2, \text{maj} \rangle, & \text{where maj}(a,b,c) = \text{majority of } a,\ b \text{ and } c \\
\mathbf{A}_{max} & = & \langle D_2, \max \rangle, & \text{where } \max(a,b,c) = a \vee b \vee c.
\end{array}
$$

A deep branch of algebra, called Tame Congruence Theory, finds fundamental differences between these algebras. (The theory of Bulatov, Jeavons and Krokhin exploits exactly this classification.)

Here we give a very different aspect in which these algebras differ. Let $cl(\mathbf{A})$ be the set of all *term operations* of the algebra $\mathbf{A}$, i.e. of those that we can obtain from the operations of $\mathbf{A}$ by any kind of compositions and identifying variables.

In the case of $\mathbf{A}_{proj}$ the output of any term operation, no matter how many variables it has, is always controlled by a single variable. In contrast, the other algebras are very different. They have term operations, where any single variable has very little say on the output statistics: for every $\epsilon > 0$ there is a term, that if Alice controls a single variable of her choice, and, independently, Bob sets the remaining variables randomly and independently (for the sake of simplicity think only of the uniform distribution on D_2 here, but we will eventually need to consider all distributions), it is nearly impossible to tell from the output statistics how Alice has set her variable. (ϵ plays a role in the notion of "nearly:" the statistical difference of the output distributions that arise from Alice's two different settings of her input bit is at most ϵ).

Although all $\mathbf{A}_{\oplus}$, $\mathbf{A}_{maj}$ and $\mathbf{A}_{max}$ are resilient in the above sense, there are fine differences between them. Consider, for instance, $\mathbf{A}_{\oplus}$, whose all term operations are parity functions. For a parity function, if Alice can choose the value of the first bit *after* looking at the setting of the other bits (in other words, if we allow dependence between Alice's and Bob's bits), she can set the output anyhow she wants to.

In contrast, in the case of the iterated majority function (which is in $cl(\mathbf{A}_{maj})$), it is easy to see that Alice cannot influence the output even when she is given the power of dependence from Bob.

Finally, we find an even stronger notion of resilience to Alice's attempt to change the output in the case of the iterated max operation (which is in $cl(\mathbf{A}_{max})$). This is just a giant OR function, which outputs almost always one for a random input (the fraction of zero outputs can be made arbitrarily small by making the iteration large enough).

Our new study tells exactly what properties of algebras give the various types of asymptotic behavior of term operations we have seen for the above examples.

Applying our results to (the special case of) $Pol(\Gamma)$, we can give characterizations of three different classes of CSPs. The first class is known as Block Projective CSPs: This is the class that does not have "interesting" polymorphisms, provably NP-hard, and contains all known NP-hard instances. The second class is the set of CSPs to which some linear equations can be reduced. The class goes under the name "CSPs with ability to count." The third class we can characterize in our framework is the class of width one CSPs. This well studied group still contains interesting problems, like *st*-connectivity.

With one leg our characterizations stand on the algebraic theory of CSPs, and with the other leg they rest on concepts familiar from PCP theory such as resilience to noise (random or adversarial) and the long code tests. The table in Figure 1 gives a summary of our results.

Algebraic Condition		*Analytic Condition on* $Pol(\Gamma)$
Block Projective	$\leftrightarrow$	Lacks Asymptotically Resilient Terms
$\neg$ Block Projective	$\leftrightarrow$	Has Asymptotically Resilient Terms
No Ability to Count	$\leftrightarrow$	Has Strongly Resilient Terms
Width One	$\rightarrow$	Has Asymptotically Constant Terms

FIGURE 1. Classes of CSPs with new characterization

Our paper also contributes a little bit to the theory of higher order dynamical systems: We characterize maps from D^n to D whose high iterates are resilient to small noise. That is, for any measure on D, if k is large enough, then no matter how we change a fixed constant number of inputs *before* other input bits are set, the distribution of the function values of the k times iterated map will be decreasingly influenced as k tends to infinity. Raghavendra [49] already used our ideas in the study of the linearity of fractional/1-approximate polymorphisms.

Our results shed new light to important classes of CSPs. We demonstrate this by showing that our new characterization of the Block projective class gives a new and more modular proof to the Hell-Nešetřil theorem (Sections 6, 7). Our characterization of the class of CSPs without the ability to count (Section 9) gives a new tool to tackle the conjecture of Feder and Vardi [26], Larose and Zádori [41], and Bulatov [10] (proved all to be equivalent by Larose, Zádori and Valeriote [42]) that CSP problems without the ability to count have bounded width. Barto and Kozik [2] already proved this conjecture using a variant of our characterization. This is considered a major step towards the resolution of the dichotomy conjecture.

1.1. **Robust approximation.** We have to mention an important connection to approximation of CSP's. We say that an algorithm robustly decides a CSP problem if it can distinguish $(1-\varepsilon)$-satisfiable CSP instances from less than $1-r(\varepsilon)$-satisfiable instances, where $r(\varepsilon) \to 0$ as $\varepsilon \to 0$. This class of CSP problems has an algebraic characterization. Our paper contributed to the understanding of these classes, too. Barto and Kozik [3] showed recently that a CSP is robustly decidable in poly-time if it has bounded width, and robust decision is NP-complete otherwise. (This was a conjecture of Guruswami and Zhou.) They have used Semidefinite Programming in the tractable case. O'Donnell, the first author, Tamaki, Yoshida and Zhou [22] studied CSP problems robustly decidable by a linear program and showed that width one CSPs belong to this class. (Dalmau and Krokhin also proved this independently [19].)

2. A connection between two theories

The theory of Probabilistically Checkable Proofs, or in short PCP theory, and the algebraic theory of CSPs both use a machinery, that we describe here from the two different angles.

First we describe the machinery from the PCP point of view. Fix n, and let $\mathscr{F}$ be a family of functions of type $f : D^n \to D$. The most frequently considered families are:

Linear Functions[2]:	$\mathscr{L} = \{f \mid f \text{ is linear over } GF(\lvert D\rvert)\}$,
Long Code:	$\mathscr{P} = \{f(x_1, \dots x_n) = x_i \mid 1 \le i \le n\}$.

In PCP theory we want to test for *membership in* $\mathscr{F}$. Given a function $f : D^n \to D$, which either belongs to the family $\mathscr{F}$ or is far from $\mathscr{F}$ (in the sense that for every $g \in \mathscr{F}$ the probability $Pr_x(f(x) \neq g(x)) > \epsilon$), the tester needs to decide with high probability, using a small number of black box queries to f, if it is in $\mathscr{F}$ or far from it.

It is equally important to tell *how* we test. In PCP theory each known test[3] is associated with a relation R on D (or with a set of relations, in which case we run tests associated with each relations, separately). Let $R \subseteq D^k$ be a k-ary relation on D and let π be a probability distribution on k tuples $(x^{(1)}, \dots, x^{(k)}) \in (D^n)^k$ that obey the property:

$$(x_i^{(1)}, \dots, x_i^{(k)}) \in R \qquad \text{for } 1 \le i \le n. \tag{1}$$

[2] in this case $|D|$ is assumed to be a prime power

[3] Hastad's test requires a little modification of the framework.

Then $\text{Test}_{R,\pi}$ is a procedure that takes a function $f : D^n \to D$ as its input, selects a k-tuple $(x^{(1)}, \ldots, x^{(k)}) \in (D^n)^k$ according to π, and accepts if and only if

$$(f(x^{(1)}), \ldots, f(x^{(k)})) \in R. \tag{2}$$

Take Dinur's test of the Long Code on $D = \{0, 1\}$ for an example. She used relations: $b = \neg a$ and $a \vee b \vee c$. The first relation is automatically provided to hold everywhere by a technique known as *folding*. For the second relation Dinur used a certain non-trivial probability distribution π on triples satisfying $x_i^{(1)} \vee x_i^{(2)} \vee x_i^{(3)} = 1$ for all i. By Fourier analytic techniques she verified that the test checks the long code in the following strong sense: If the acceptance probability is $1-\varepsilon$ then f must coincide with some word of the long code on $1 - O(\varepsilon)$ fraction of randomly and uniformly chosen elements of D^n. Dinur's analysis also gives, that if her test is accepted with probability one, then f is an element of the long code.

We are now warmed up to a different point of view of the above machinery. For relation R on D define relation R^n on D^n in the usual way: $(x^{(1)}, \ldots, x^{(k)}) \in R^n$ if and only if Equation (1) holds. Then Equation (2) holds for every k-tuple in R^n if and only if f is a homomorphism from R^n to R.

Term operation
Compatible operation

Definition 3. Let $\Gamma = \{R_1, \ldots, R_l\}$ be a set of finitary relations on D. A function $f : D^n \to D$ is a *term operation* of Γ with arity n if f is a homomorphism from R_i^n to R_i for every $1 \le i \le l$ (i.e. satisfies Equation (1) $\to$ Equation (2) for all R_i). Term operations are sometimes called *compatible operations.*

$Pol(\Gamma)$

Definition 4. The set of all term operations of arbitrary arities associated with a constraint family Γ form $Pol(\Gamma)$.

Example 5. Let $D = \{0, 1, 2\}$, and let Γ contain all ternary relations that express linear equations over $GF(3)$, for instance, $x + 0y + 2z = 1$. Then $f(x_1, x_2, x_3) = x_1 - x_2 + x_3 \mod 3$ is a term operation for Γ. Indeed, consider a relation $ax + by + cz = d \mod 3$. Assume that $ax_i + by_i + cz_i = d$ for $1 \le i \le 3$. Then $a(x_1 - x_2 + x_3) + b(y_1 - y_2 + y_3) + c(z_1 - z_2 + z_3) = d - d + d = d$. One can similarly show that all functions of the form $f(x_1, \ldots, x_n) = \sum_{i=1}^n \alpha_i x_i$, $\sum_{i=1}^n \alpha_i = 1 \mod 3$ are in $Pol(\Gamma)$. In fact, $Pol(\Gamma)$ consists exactly of these functions.

Remark 6. It is easy to see that composition of functions in $Pol(\Gamma)$ remains in $Pol(\Gamma)$, and that $Pol(\Gamma)$ contains all projections (i.e. members of $\mathscr{P}$). When we compose we do not have to make the variable sets of the inner functions disjoint. A consequence of this, and of the fact that we have all projections in $Pol(\Gamma)$, is that identifying variables does not take us out of $Pol(\Gamma)$. For instance, if $f(x, y, z) \in Pol(\Gamma)$, then so is $f(x, y, x)$.

$Pol(\Gamma, n)$

Definition 7. $Pol(\Gamma, n)$ is the set of the n-ary members of $Pol(\Gamma)$.

$Pol(\Gamma, n)$ is exactly the family that we can analyze with the Fourier analytic techniques of the PCP theory.

3. More on the algebraic theory

The fact that the algebraic theory of CSPs and long code tests talk about the same objects, raises a lot of questions. Why this connection has not been utilized thus far? The answer perhaps is that the testing theory deals with *analytical* properties of functions that *nearly* satisfy the tests, while the algebraic theory of CSPs deals with *algebraic* properties of functions that keep *all* relations. One of our contributions is that we positively demonstrate, that it is worthwhile to take an analytic approach to functions that keep *all* relations. When these functions are examined both from analytic and algebraic viewpoints, nontrivial conclusions like the Hell-Nešetřil theorem can be obtained. Identities

The connection has another great benefit, namely it lends more sense to rewriting algebraic identities into analytic form. Let us explain: Bulatov, Jeavons and Krokhin essentially conjectured that CSP(Γ) is tractable iff there is a compatible operation which can be distinguished from the projections by its identities. E.g a majority operation satisfies $f(y, x, x) = f(x, x, x)$, $f(x, y, x) = f(x, x, x)$, $f(x, x, y) = f(x, x, x)$, the i^{th} identity shows that this can not be a projection to the i^{th} coordinate, since this coordinate is x on one side and y on the other side. The above is just a special case. Before getting closer to algebra we have to use two technical assumptions.

Core

Definition 8. We only want to deal with the case when Γ is a core, i.e. every homomorphism $\langle D, \Gamma\rangle \to \langle D', \Gamma|_{D'}\rangle$ ($D' \subseteq D$) is an automorphism. Every structure has a unique core (up to isomorphism), and a structure and its core define the same CSP language. That Γ is a core is always assumed in the literature to make algebraic methods to work.

Example 9. Let $D = S_3$ (the set of permutations of $\{1, 2, 3\}$) and let Γ consist of a single trinary relation: $(x, y, z) \in R$ if and only if $z = xy$ in S_3. What is the complexity of $CSP(R)$? A closer examination reveals that this problem is equivalent to the same problem for the two element cyclic group. Why? Consider an input instance to $CSP(R)$. We claim that if the instance is satisfiable, then it is also satisfiable with an assignment, where all values are from $\{e, (1,2)\} \subset S_3$. Indeed, let φ be the homomorphism that takes even elements of S_3 to e, and odd elements into $(1,2)$, and apply φ to the value of each variable of the original assignment to get a new assignment. If $(x, y, z) \in R$, i.e. $z = xy$ in S_3, then $(\varphi(x), \varphi(y), \varphi(z)) \in R$ too, since $\varphi(z) = \varphi(x)\varphi(y)$, because φ is a group homomorphism from S_3 to Z_2.

The core of $\Gamma = \{R\}$ is $\{e\}$, however, (and not $\{e, (1,2)\}$) which makes $CSP(R)$ (even more) obvious: all instances can be satisfied just by assigning e to all variables.

Example 10. Let us add the unary relation S to the previous example that selects the odd elements of S_3 (thus we can say: "variable x takes an odd permutation"). Now $\{e, (1,2)\}$ will be a core for $\Gamma = \{R, S\}$, and so will be $\{e, (1,3)\}$ and $\{e, (2,3)\}$. These cores are isomorphic.

Idempotency

Definition 11. f is idempotent if $f(x, \dots, x) = x$ for every $x \in D$.

In $Pol(\Gamma)$ we only want to consider idempotent operations. The simple reason is that the complexity of a core CSP problem depends only on its idempotent operations, and this assumption simplifies the algebraic theory a lot. A way to make $Pol(\Gamma)$ idempotent (for a core Γ, without changing the complexity of $CSP(\Gamma)$) is by adding all unary relations $x = c$ for all $c \in D$.

WNUs

By a result of McKenzie and Maróti, if there is a compatible operation which can be distinguished from the projections by its identities, then there is also a special type, called *weak near-unanimity* (WNU) term. An idempotent operation f is a WNU if for every $x, y \in X$ it satisfies $f(y, x, \dots, x) = f(x, y, x, \dots, x) = \cdots = f(x, \dots, x, y)$.

The following theorem uses the WNU condition of Maróti and McKenzie [43], while condition (1) is stated in the combinatorial terminology of Nešetřil, Siggers and Zádori [44].

Theorem 12. *For any constraint family Γ the following are equivalent.*

1. Γ *is not block-projective, i.e. there exist no disjoint subsets* S_1, S_2 *of* D *such that* $S_1 \cup S_2$ *is a subalgebra and for every compatible, idempotent operation* f *there exists a* k *such that if* $x_1, \dots, x_n \in S_1 \cup S_2$ *then* $f(x_1, \dots, x_n) \in S_i$ *iff* $x_k \in S_i$ *for* $i = 1, 2$.
2. *There exists a compatible WNU term operation.*

In the next sections we add to the above equivalent conditions a new one: There exists a compatible WNU term operation iff there exists a sequence of term operations that are arbitrarily resilient to small noise. This is part of our larger project of translating algebraic conditions into analytic ones.

It may occur that in $Pol(\Gamma)$ there is no WNU operation, but CSP(Γ) is not NP-complete, however this cannot happen when Γ is a core. So we restrict ourselves to cores as promised.

Theorem 13. *If Γ is a core and has no compatible WNU operation then* CSP(Γ) *is NP-complete.*

The Dichotomy Conjecture of Bulatov, Jeavons and Krokhin states that Theorem 13 can be reversed in the following sense:

Conjecture 14 (Algebraic Dichotomy Conjecture [14])**.** Let Γ be a core. If Γ admits compatible WNU operation then CSP(Γ) is tractable, else it is NP-complete.

Example 15. This gives a remark for Dinur's test: her test implies the NP-hardness of CSP($\neg a = b,\ a \wedge b \wedge c$) by Theorem 13. While this is not earth-shattering, the Algebraic Dichotomy Conjecture also immediately suggests that the $\neg a = b$ folding is essential for the test to work.

Example 16. Considering Unique games (see Appendix) it is easy to see that $Pol(\mathrm{UG}_D)$ contains a lot of WNUs. Our Theorem 113 for instance immediately implies that

$$f(x_1, \ldots, x_n) = \begin{cases} a & \text{if there is a unique } a \in D \text{ such that } |\{i \mid x_i = a\}| \text{ is maximal,} \\ x_1 & \text{otherwise.} \end{cases}$$

is in $Pol(\mathrm{UG}_D)$. This is a WNU, and in fact a majority operation.

Bulatov, Jeavons and Krokhin used the term *Polymorphism* for functions (of arbitrary number of variables) that are compatible with all relations in Γ, and they denoted this set of functions by $Pol(\Gamma)$. They have proved that $Pol(\Gamma)$ determines the complexity of Γ, i.e. different problems with the same Pol are inter-reducible in polynomial time. The approach has been applied in several contexts, in particular, this is how Bulatov solved the problem for $|D| = 3$ [7]. Another application of their theory by Bulatov proves dichotomy, when Γ is a set of list homomorphisms [8]. The original goal of the algebraic theory was to deal with decision problems, though it proved to be successful in other cases. Bulatov an Dalmau proved a dichotomy theorem for counting the solutions of CSPs [12], Bodirsky and Nešetřil [6] managed to extend the theory to (omega-categorical) countably infinite target structures, Chen partly managed for quantified CSPs [15].

$|D| = 3$

list homomorphisms

counting classes

The harder part of the algebraic dichotomy conjecture is the tractable part: how does an algebraic condition lead to tractability? Jeavons, Cohen and Gyssens [17] proved that the existence of a semilattice operation implies tractability, Cohen, Cooper and Jeavons [16] proved it in case of the existence of a so-called near-unanimity operation (a generalization of majority operations, still stricter than WNUs), Bulatov and Dalmau [11] in case of the existence a so-called Maltsev term (what shows that the algebra is "somewhat grouplike"): the algorithms are generalizations of the ones solving Horn-formulas, 2-SAT and linear system of equations, respectively. But to solve a general tractable CSP problem we need to combine these algorithms (and also, of course, find new ones). There are very few results that combine algorithms of different nature: Bulatov's result for list homomorphisms is one such example [8], and Dalmau's result for CSPs that have an operation on every pair behaving like a group or a majority operation [23] is another. The latter result was generalized in a "truly algebraic" manner in [5].

near-unanimity

Maltsev

combined

4. Term Operations of Algebras

Perhaps it sounds like we are splitting hairs, but $Pol(\Gamma)$ is not an algebra, for its operations are not named. Every algebra has to have a type $\mathcal{F}$, which is a set of *function symbols*. Each symbol $f \in \mathcal{F}$ comes with an arity n_f, which is a non-negative integer.

Definition 17. An algebra $\mathbf{A}$ of type $\mathcal{F}$ is an ordered pair $\langle D, F \rangle$, where D is a non-empty set and F is a family of operations indexed by $\mathcal{F}$ such that $f^{\mathbf{A}} \in F$ is an n_f-ary operation on D.

Example 18. Groups have type $\mathcal{G} = \{\times, \ ^{-1}, 1\}$. Function symbols $\times$, $^{-1}$ and 1 have arities 2, 1 and 0, respectively.

In contrast, syntactically $Pol(\Gamma)$ is just an (unordered) set of finitary functions of the type $D^n \to D$.

Definition 19. We denote the set of all finitary operations $D^n \to D$ $(n = 0, 1, 2, \ldots)$ by $\mathrm{Op}[D]$.

The difference between an algebra and a subset of $\mathrm{Op}[D]$ is similar to the difference between codes that are discussed as a collection (set) of code words, and ones that are discussed together with the encoding function.

With algebras we do not have such a choice. There are several good reasons to name the operations of an algebra. First, the notion of direct product requires that we can identify the "same" operations. For instance, it is essential that we match multiplication with multiplication and addition with addition when taking the direct product of two rings. Second, we also want sub-algebras of an algebra to have the same type as the original one.

Definition 20. Let D be a finite universe. A *clone* on D is a set $\mathcal{C} \subseteq \mathrm{Op}[D]$ that satisfies:

(1) For all $n \geq i \geq 1$ all projections $p_i(x_1, \dots x_n) = x_i$ are in $\mathcal{C}$.
(2) $\mathcal{C}$ is closed under composition: if $f(x_1 \dots, x_n)$ and $g_1, \dots, g_n$ are in $\mathcal{C}$, then so is $f(g_1, \dots, g_n)$.
(3) If $f \in \mathcal{C}$, then identifying some variables of f with each other does not take out of $\mathcal{C}$.

Our most prominent example for clones is $Pol(\Gamma)$ for some family Γ. The clone properties of $Pol(\Gamma)$ are stated in Remark 6. We now give another important example for clones:

Term operations for algebras

Definition 21. Let $\mathbf{A} = \langle A, F \rangle$ be an algebra. A *term operation* of $\mathbf{A}$ is an element of $\mathrm{Op}[D]$ that can be created from projections and operations in $\mathbf{A}$ by compositions and identification of variables.

Lemma 22. *The set of all term operations of an algebra* $\mathbf{A}$ *form a clone.*

Example 23. Let $G = \langle A, \times, {}^{-1}, 1 \rangle$ be a group. Then the set of all functions that we can create using a finite number of variables, such as e.g. $((x \times y) \times (x^{-1} \times z)) \times y^{-1}$, form the clone $cl(G)$.

Definition 24. An algebra $\mathbf{A}$ is *idempotent* if for every $f \in \mathbf{A}$ we have $f(x, \dots, x) = x$.

Lemma 25. *All term operations of an idempotent algebra are idempotent.*

We would like to stress here that the objects about which we gather useful information in this article, are clones.

Yet, we will not mention clones in the rest of the article for the following reasons:

(1) The main applications of our results are for $Pol(\Gamma)$.
(2) General clones can be conveniently interpreted as term operations of algebras.

Remark 26. Indeed, when we study $Pol(\Gamma)$, or any other clone, algebra is an indispensable tool. The way we introduce algebra is that first we turn the clone into an algebra by indexing all its operations with themselves. (We may also index only a subset of operations if the rest can be interpreted as term operations.) In the sequel, we will either talk about $Pol(\Gamma)$, and call its elements term operations (or compatible operations), or we consider an algebra together with the set of its term operations. The later type of discussion is not only more general, but also more convenient, when we use algebraic tools, and this will be our choice e.g. throughout Section 9.

In Section 9 we will need the notion of subalgebra:

Definition 27. $\mathbf{A}' = \langle A', F' \rangle$ is a subalgebra of $\mathbf{A} = \langle A, F \rangle$ if $A' \subseteq A$, they have the same type $\mathcal{F}$, and for every symbol $f \in \mathcal{F}$ operation $f^{\mathbf{A}'}$ is the restriction of $f^{\mathbf{A}}$ to A'. This assumes that $\mathbf{A}'$ is closed under all operations of $\mathbf{A}$.

5. An analytic look at term operations

The goal of this article is to create a one to one correspondence between known classes of CSPs and analytic properties of their set of compatible operations, or more generally, certain properties of algebras and the asymptotic behavior of their term operations. In this section we develop the necessary concepts for these studies.

Let D be a finite domain, $f : D^n \to D$ and $\mu_1, \ldots, \mu_n$ be distributions on D. By $f(\mu_1, \ldots, \mu_n)$, or shortly by $f(\vec{\mu})$, we denote the distribution on D that we obtain by plugging independent D-valued random variables into f such that the ith variable is distributed as μ_i.

Definition 28 (Resilience)**.**

$$Resil(f, l, \mu) = \sup_{\mu_1, \ldots, \mu_n} \delta(f(\mu, \mu, \ldots, \mu), f(\mu_1, \ldots, \mu_n)),$$

Resilience

where δ refers to the statistical difference, $\delta(\mu, \nu) = \frac{1}{2} \sum_{x \in D} |\mu(x) - \nu(x)|$, of distributions and $\mu_1, \ldots, \mu_n$ runs through all sequences of distributions on D with the properties that at most l of the μ_is are different from μ and the support of each μ_i is contained in the support of μ. We call $Resil(f, l, \mu)$ the *resilience* of f.

Influence

Definition 29 (Influence)**.** Let D be a finite domain and μ be a measure on D. The *influence* of the i^{th} variable of $f : D^n \to D$ is $\text{Inf}_{i,\mu}(f) = Prob_{\mu^{n+1}}(f(x) \neq f(x'))$, where x, x' runs through all random input-pairs that differ only in the ith coordinate: μ^{n+1} gives a natural measure on such pairs.

max inf

Definition 30. The *maximal influence*, $\max \inf_\mu(f)$ is $\max_i \text{Inf}_{i,\mu}(f)$.

Invariance

Definition 31 (Invariance)**.** Let D be a finite domain and μ be a measure on D. The *invariance*, inv_μ of $f : D^n \to D$, is the smallest $0 \leq \epsilon \leq 1$ for which there is a $c \in D$ for which $Prob_{x \in \mu^n}(f(x) \neq c) \leq \epsilon$.

We can now study $Pol(\Gamma)$, or in general, term operations for an algebra $\mathbf{A}$ from an analytic point of view:

An algebra $\mathbf{A}$ with term operations $cl(\mathbf{A})$ is

Asymptotically Resilient: if

$$\forall \varepsilon, \mu, l\, \exists f \in cl(\mathbf{A}): \quad Resil(f, l, \mu) < \varepsilon.$$

Strongly Resilient: if

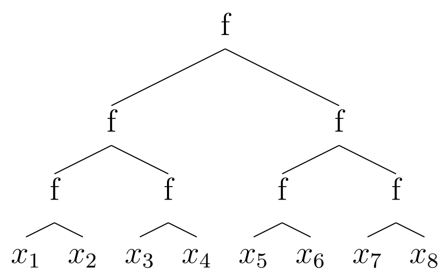
$$\forall \varepsilon, \mu\, \exists f \in cl(\mathbf{A}): \quad \max \inf_{\mu} f < \varepsilon.$$

Asymptotically Invariant: if

$$\forall \varepsilon, \mu\, \exists f \in cl(\mathbf{A}): \quad \mathrm{inv}_{\mu} f < \varepsilon.$$

It turns out that $\mathbf{A}$ has the asymptotically or the strongly resilient property, respectively, if and only if there is a function in $\mathbf{A}$ whose set of iterates have this property. The iterates of a function $f : D^n \to D$ are $f^1 = f$, $f^{i+1} = f(f^i, \ldots, f^i)$ for $i > 1$. The arity of f^k is n^k, and we can visualize it as an n-ary tree of depth k built of fs.

$f^3 =$



Definition 32 (Asymptotic Resilience of a function)**.** We say that a function $f : D^n \to D$ is *asymptotically resilient* if for every distribution μ on D and every $l, \varepsilon > 0$ we have $Resil(f^k, l, \mu) < \varepsilon$ for any sufficiently large k.

Definition 33 (Strong Resilience of a function)**.** A function $f : D^n \to D$ is strongly resilient if for every measure μ on D: $\max \inf_{\mu}(f^k) \to 0$ when $k \to \infty$.

Definition 34 (Asymptotic Invariance of a function)**.** A function $f : D^n \to D$ is *asymptotically invariant* (or *asymptotically constant*) if for every measure μ on D the invariance of f^k tends to 0 when $k \to \infty$.

6. Asymptotic resilience

Most functions are asymptotically resilient, but e.g. projections are not. Instead of giving further examples we describe *all* asymptotically resilient idempotent functions. (With a little extra effort one can give a similar characterization without the idempotency condition.)

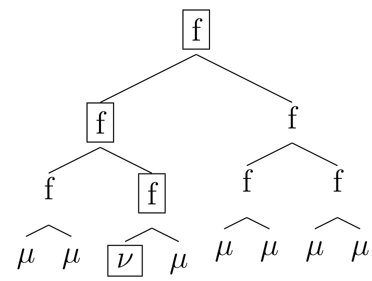


FIGURE 2. If f is a WNU, as we go up in the tree, the variable controlled by Alice has less and less effect

Theorem 35. *Let $f : D^n \to D$ be idempotent. The following are equivalent:*

(1) *f is asymptotically resilient.*
(2) *$Resil(f^k, 1, \mu)$ goes to zero as k goes to infinite for every fixed μ.*
(3) *f generates a WNU (including that itself is a WNU).*
(4) *There do not exist pairs of disjoint subsets $S_1, S_2 \subseteq D$ and $1 \le k \le n$ with the following property: $S_1 \cup S_2$ is closed under f and if $x_1, \ldots, x_n \in S_1 \cup S_2$ then $f(x_1, \ldots, x_n) \in S_i$ iff $x_k \in S_i$ for $i = 1, 2$.*

Proof. (1) implies (2) by the definition of asymptotic resilience. For (2) $\to$ (1) we prove:

Lemma 36. *If $Resil(f^k, 1, \mu)$ goes to zero as k goes to infinite, then so does $Resil(f^k, l, \mu)$ for every $l \ge 1$.*

Proof. We proceed by induction on l. The case $l = 1$ is trivial. Let $l \ge 2$ and $\varepsilon > 0$, and let k', k'' be such that $r(f^{k'}, 1, \mu) < \varepsilon/l$ and $r(f^{k''}, l-1, \mu) < \varepsilon$ (by induction), respectively. Let $g = f^{k'}$ and $h = f^{k''}$. For $k = k' + k''$ we have:

$$f^k = h(g, \ldots, g).$$

Let L be any subset of l inputs for f^k. We will show that f^k is ϵ-resilient with respect to L. We distinguish between two cases:

Case 1: Each g in $h(g, \ldots, g)$ gets at most one input from L. The output of those that get an input from L is ε/l-close to the distribution $h(\mu^{n^{k'}})$ by the choice of k'. We then use Proposition 38.

Case 2: There is a g in $h(g, \ldots, g)$ which gets at least two inputs from L. In that case at most $(l-1)$ of the g's involve inputs from L, and we use that $r(h, l-1, \mu) < \varepsilon$. □

The equivalence of (3) and (4) was proved by McKenzie and Maróti [43], and we have discussed it in Section 3. It is easy to see that (1) implies (4): in fact if (4) does not hold then $Resil(f^k, 1, \mu) = 1$ for every k. In the rest of the section we prove that (3) implies (2).

For the sake of simplicity we assume that f is a WNU itself, of arity n (if f only *generates* a WNU, the proof needs only a minor adjustment). For our argument we fix μ. Let $\mu_k = f^k(\mu^{n^k})$ (recall that the arity of f^k is n^k). We would like to estimate the statistical difference of μ_k and $f^k(\mu^{i-1}\nu\mu^{n^k-i})$ for any $1 \le i \le n^k$ and any ν, whose support is contained in the support of μ. Let $\alpha_k = \max_{i,\nu} \delta(\mu_k, f^k(\mu^{i-1}\nu\mu^{n^k-i}))$. What we need to show is that $\alpha_k \to 0$. By the following propositions and its corollary it is straightforward that α_k is non-increasing:

Proposition 37. *The variation distance of two distributions cannot increase under any map $F : X \to Y$.*

Proof. If μ and ν are the two distributions we can write: $\delta(\mu,\nu) = \frac{1}{2}\sum_{y\in Y} |\mu(y) - \nu(y)| \geq \frac{1}{2}\sum_{x\in X} \left|\sum_{y\in F^{-1}(x)}(\mu(y) - \nu(y))\right| = \delta(F(\mu), F(\nu))$. □

Corollary 38. *Let $f : D^n \to D$ be arbitrary and $\mu_1,\dots,\mu_n$, $\nu_1,\dots,\nu_n$, be two sequences of distributions on D. Then*

$$\delta(f(\mu_1,\dots,\mu_n), f(\nu_1,\dots,\nu_n)) \le \sum_{i=1}^{n} \delta(\mu_i,\nu_i).$$

Proof. The corollary follows from $\delta(\prod_{i=1}^n \mu_i, \prod_{i=1}^n \nu_i) \le 1 - \prod_{i=1}^n (1 - \delta(\mu_i,\nu_i)) \le \sum_{i=1}^n \delta(\mu_i,\nu_i)$. □

We now want to go a step further and to show that α_{k+1}/α_k is upper bounded by a constant (i.e. independent of k) less than 1. It is easy to see that Proposition 37 can be strengthened if we find an $x \in X$, $y_0, y_1 \in F^{-1}(x)$ such that $\mu(y_0) - \nu(y_0) \ge 0$, $\nu(y_1) - \mu(y_1) \ge 0$:

$$\text{(3)}\qquad \delta(F(\mu), F(\nu)) \le \delta(\mu,\nu) - \min\{\mu(y_0) - \nu(y_0), \nu(y_1) - \mu(y_1)\}.$$

At this point we exploit that f is a WNU, and certain identities hold for its output. Before describing what we get from this we need a technical definition:

Definition 39. Let μ and ν be probability distributions on X. We define

$$\min \frac{\mu}{\nu} = \min_{x:\nu(x)\neq 0} \frac{\mu(x)}{\nu(x)}.$$

Lemma 40. *For every WNU term f and probability distributions μ and ν on D:*

$$\delta(f(\mu^{i-1}\nu\mu^{n-i}), f(\mu^n)) \leq \left(1 - \frac{\delta(\mu,\nu)^{n-1}}{|D|^n} \min\frac{\mu}{\nu}\right)\delta(\mu,\nu)$$

for every $1 \leq i \leq n$.

Proof. (of Lemma 40) There are $x, y \in D$ such that

$$\mu(x) - \nu(x) \geq \delta(\mu,\nu)/|D| \tag{4}$$

$$\nu(y) - \mu(y) \geq \delta(\mu,\nu)/|D|. \tag{5}$$

Without loss of generality assume that $i = 1$. Define:

$$\begin{aligned}
p_1 &= Prob_{\mu^n}(y, x, \ldots, x) &&= \mu(y)\mu(x)^{n-1},\\
q_1 &= Prob_{\nu\mu^{n-1}}(y, x, \ldots, x) &&= \nu(y)\mu(x)^{n-1},\\
p_2 &= Prob_{\mu^n}(x, y, x, \ldots, x) &&= \mu(x)\mu(y)\mu(x)^{n-2},\\
q_2 &= Prob_{\nu\mu^{n-1}}(x, y, x, \ldots, x) &&= \nu(x)\mu(y)\mu(x)^{n-2}.
\end{aligned}$$

From (4) and (5) we obtain that

$$\begin{aligned}
\mu(x) &\geq \frac{\delta(\mu,\nu)}{|D|},\\
\mu(y) &\geq \frac{\delta(\mu,\nu)}{|D|}\min\frac{\mu}{\nu},\\
p_1 - q_1, q_2 - p_2 &> 0.
\end{aligned}$$

Let $a = f(y, x, \ldots, x) = f(x, y, x \ldots, x)$. From (3):

$$\begin{aligned}
\delta(f(\mu^{i-1}\nu\mu^{n-i}), f(\mu^n)) &\leq \delta(\mu^{i-1}\nu\mu^{n-i}, \mu^n) - \min\{p_1 - q_1, q_2 - p_2\}\\
&\leq \delta(\mu,\nu) - \frac{\delta(\mu,\nu)^n}{|D|^n}\min\frac{\mu}{\nu}.
\end{aligned}$$

□

Lemma 40 gives that $\alpha_{k+1}/\alpha_k \leq 1 - \frac{\alpha_k^{n-1}}{|D|^n}\min_{\tilde{\mu}_k}\min\frac{\mu_k}{\tilde{\mu}_k}$, where $\tilde{\mu}_k$ ranges among distributions of the form $f^k(\mu^{i-1}\nu\mu^{n^k-i})$. Indeed, use that $f^{k+1}(\mu^{i-1}\nu\mu^{n^{k+1}-i})$ can be written as f on many copies of μ_k and one copy of $f^k(\mu^{i'-1}\nu\mu^{n^k-i'})$. An easy analysis shows that this improvement is sufficient, because $\min_{\tilde{\mu}_k}\min\frac{\mu_k}{\tilde{\mu}_k}$ remains bounded from below by $\min\frac{\mu}{\nu}$. This follows from the more general:

$$\min\frac{\prod_i \mu_i}{\prod_i \nu_i} = \prod_i \min\frac{\mu_i}{\nu_i}; \quad \min\frac{F(\mu)}{F(\nu)} \geq \min\frac{\mu}{\nu}.$$

□

7. The Hell-Nešetřil theorem

Recall that the Hell-Nešetřil theorem [29] states that for a simple, connected, undirected graph H the complexity of CSP(H) is polynomial if H is bipartite and NP-complete otherwise. The first part of the theorem is trivial: every bipartite graph with at least one edge has a retraction to any of its edges. This implies that a graph G has a homomorphism into H iff it has a homomorphism into a single edge, i.e. G is bipartite. The interesting, and combinatorially quite involved part is the NP-completeness of CSP(H) when H is non-bipartite. This was the first dichotomy theorem with a really sophisticated proof using gadget reductions. Later Bulatov [9] streamlined the proof using the algebraic theory, though his proof still has some ad hoc part. Barto, Kozik and Niven [4] extended the theorem proving dichotomy for digraphs with no sink and source.

Here we give a proof based on our notion of asymptotic resilience. According to Theorem 13 in [14] and using the fact that the core of a non-bipartite graph is also non-bipartite it is sufficient to prove that:

Lemma 41. *Let H be a simple, connected, undirected, non-bipartite graph. Then $Pol(H)$ has no asymptotically resilient term.*

Proof. We explore the analytic properties of $Pol(H,n)$.

Let us denote the vertex set of H by D (faithfully to our prior notations), and the edge set of H with E. The power set of a set is denoted by $P()$. Let

Graph	Vertices	Edges
H^n	D^n	$(\vec{v},\vec{w}) : (v_i,w_i) \in E$ for all $1 \leq i \leq n$;
$P(H)$	$P(D)$	(S,T): $(s,t) \in E$ for all $s \in S$ and $t \in T$.

The *stationary measure* on the vertices, μ, (edges, μ_E) of H assigns frequencies to every node (edge), with which that node (edge) is visited by an infinite random walk. It is well known that the stationary measure on the edges of a simple, connected, undirected, non-bipartite graph is uniform. This implies that the stationary measure on the vertices is proportional to the degree of each node.

It is immediate that the stationary measure on the vertices (edges) of H^n is μ^n (μ_E^n), where

$$\mu^n((v_1,\dots,v_n)) = \prod_{i=1}^{n} \mu(v_i); \qquad \mu_E^n((e_1,\dots,e_n)) = \prod_{i=1}^{n} \mu_E(e_i).$$

We would like to find analytic properties of $Pol(H,n)$. What we show is that, independently of n, for any $f \in Pol(H,n)$ we find a constant number of coordinates that (jointly) have non-negligible influence on the value of f. This holds when H is connected, non-bipartite.

Lemma 42. *Let $H = (D, E)$ be a simple, connected, undirected,* non-bipartite *graph, and let μ and μ_E be the stationary measure on its vertices and edges. Then for every $\varepsilon > 0$ there exists an integer $l = l(\varepsilon, H)$ such that if $f : D^n \to D$ is a homomorphism then there is a mapping $s : D^n \to P(D)$ such that:*

1. $\text{Prob}_{(v,w)\in\mu_E^n}((s(v), s(w))$ *is not an edge in* $P(H)) \leq \epsilon$*;*
2. $\text{Prob}_{v\in\mu^n}(f(v) \notin s(v)) \leq \epsilon$*.*
3. *The mapping s depends on at most l coordinates, i.e. there is an $\mathbf{s} : D^l \to P(D)$ such that $s(r, t) = \mathbf{s}(r)$ holds for all r, t, where r is a partial assignment to the selected l coordinates, and t is an assignment to the complementary $n - l$ coordinates.*

Proof. Since $f : D^n \to D$ is a graph homomorphism, the inverse image, $f^{-1}(K)$, of an independent set $K \subseteq D$ is independent in H^n. We use a theorem of Dinur, Friedgut and Regev to show that $f^{-1}(K)$ has a special structure:

Theorem 43. [24] *Let $H = (D, E)$ be a simple, undirected, connected, non-bipartite graph with stationary measures μ and μ_E on its vertices and edges. Then for every $\delta > 0$ there exists a positive integer $j = j(\delta)$ such that to every independent set I in H^n we can associate a set of coordinates L_I and an "almost independent" set I^* that spans less than δ fraction of the edges (according to measure μ_E^n) and depends only on coordinates in L_I, such that*

$$\mu^n(I \setminus I^*) \leq \delta. \tag{6}$$

For an independent set $I \subseteq D^n$ let L_I and I^* as in Theorem 43. We choose δ later. Let $\text{Ind}(H)$ be the system of all independent sets of H, and define the following set of coordinates:

$$L = \bigcup_{K\in\text{Ind(H)}} L_{f^{-1}(K)} \qquad l = |L|.$$

We define $s : D^n \to P(D)$ via an "inverse" function $S : D \to P(D^n)$ as:

$$S(x) = \bigcap_{\substack{K\in\text{Ind(H)}\\ x\in K}} f^{-1}(K)^* \tag{7}$$

$$s(v) = \{x \in D |\ v \in S(x)\}. \tag{8}$$

Lemma 44. *The mapping s depends only on its coordinates in L.*

Proof. Notice that for any $K \in \text{Ind(H)}$ membership in $f^{-1}(K)^*$ depends only on the coordinates in $L_{f^{-1}(K)} \subseteq L$. Thus whether $w \in S(x)$ or not depends only on those coordinates of w that are in L. This, in turn, implies that $s(w)$ depends only on those coordinates of w that are in L. □

We now set $\delta = \varepsilon/|\text{Ind(H)}|$.

Lemma 45. *Condition (1) of Lemma 42 holds for s.*

Proof. If $(s(v), s(w))$ is not an edge in $P(H)$ then there are $x \in s(v)$, $y \in s(w)$ such that $\{x, y\} \in \text{Ind}(H)$, which in turn by Definitions (7) and (8) implies that $v, w \in f^{-1}(\{x, y\})^*$. The total measure of (v, w) edges contained in $f^{-1}(\{x, y\})^*$ is at most δ. This multiplied by the number of $\{x, y\} \in \text{Ind(H)}$ is at most ϵ, as needed for Condition (1). □

Lemma 46. *Condition (2) of Lemma 42 holds for f.*

Proof. Let us call $v \in D^n$ *faulty* if for some $K \in \text{Ind}(H)$ it belongs to $f^{-1}(K) \setminus f^{-1}(K)^*$. The probability that v is faulty is then at most $\delta|\text{Ind}(H)|$. It is obvious from our definitions, that when v is not faulty, then $f(v) \in s(v)$. □

□

Lemma 42 provides sufficient information about members of $Pol(H)$ to show that $Pol(H)$ is not asymptotically resilient:

Assume for a contradiction that we have an asymptotically resilient $f_0 \in Pol(H)$. Let l be the integer given by Lemma 42 for the choice of $\varepsilon = \frac{1}{7}$. For large enough k the iterate $f = f_0^k$ is a homomorphism $H^n \to H$ such that the following holds: for any l coordinates $1 \le a_1 < a_2 < \cdots < a_l \le n$ we have, that no matter how Alice sets the input variables on $x_{a_1}, \ldots, x_{a_l}$, if Bob gives random (according to the stationary distribution for H) values to the other variables, the output is $\frac{1}{14}$-indistinguishable, in the statistical difference sense, from the distribution that arises when Bob gives random values to all variables.

This immediately implies that if r and r' are any two elements of D^l then

$$\delta(f(r, \mu^{n-l}), f(r', \mu^{n-l})) < \frac{1}{7}. \tag{9}$$

Let $s : H^n \to P(H)$ be the "ϵ-almost homomorphism" that "ϵ-covers f" and depends on l coordinates, provided by Lemma 42. Then $s(r, t) = \mathbf{s}(r)$ for $r \in D^l$, $t \in D^{n-l}$. (For simplicity of notation we assumed that the crucial coordinates are the first l ones.)

Definition 47. We call $r \in D^l$ bad if the stationary measure of those $t \in D^{n-l}$ for which $f(r,t) \notin \mathbf{s}(r)$ is at least $\frac{1}{3}$.

Lemma 48. *The measure of bad vertices in H^l is at most 3ε.*

Proof. Follows from Markov's inequality, when applied to

$$\mathrm{E}_{r \in \mu^l}\left(\mathrm{Prob}_{t \in \mu^{n-l}}(f(r,t) \notin \mathbf{s}(r))\right) = \mathrm{Prob}_{(r,t)\in\mu^n}(f(r,t) \notin \mathbf{s}(r)) \leq \varepsilon.$$

□

Lemma 49. *For any two good $r, r' \in D^l$ we have $\mathbf{s}(r) \cap \mathbf{s}(r') \neq \emptyset$.*

Proof. We assume that $\mathbf{s}(r) \cap \mathbf{s}(r') = \emptyset$, and get a contradiction with Equation (9) by showing that the statistical difference of $f(r, \mu^{n-l})$ and $f(r', \mu^{n-l})$ (both are distributions on D) is too large ($\geq \frac{1}{3}$). Indeed, use the definition of statistical difference that $\delta(\alpha, \beta) = \max_A |\mathrm{Prob}_\alpha(A) - \mathrm{Prob}_\beta(A)|$, for distributions $\alpha = f(r, \mu^{n-l})$, $\beta = f(r', \mu^{n-l})$, and event $A = \mathbf{s}(r)$. The probability that $f(r, \mu^{n-l}) \in \mathbf{s}(r)$ is at least $2/3$. Also, the disjointness assumption implies that

$$\mathrm{Prob}(f(r', \mu^{n-l}) \in \mathbf{s}(r)) \leq 1 - \mathrm{Prob}(f(r', \mu^{n-l}) \in \mathbf{s}(r)) \leq 1 - \frac{2}{3} = \frac{1}{3},$$

which leads to the desired contradiction. □

The following is trivial:

Lemma 50. *Let G be a simple, connected, undirected, non-bipartite graph. The stationary measure of a set of edges incident to a set of vertices with stationary measure h is at most $2h$.*

We apply this lemma to H^l. Since the stationary measure of bad vertices is $\leq 3\varepsilon$, the measure of edges not incident to any bad vertex is at least $1 - 6\varepsilon$. By Lemma 49 the two endpoints of any such edge are mapped into intersecting sets by s. Since H has no loops, a pair of two intersecting sets always forms a non-edge in $P(H)$. We have arrived at a contradiction, since $1 - 6\varepsilon > \varepsilon$, contradicting that

$$\mathrm{Prob}_{(r,r')\in\mu_E^l}((\mathbf{s}(r), \mathbf{s}(r')) \text{ is not an edge in } P(H)) \leq \varepsilon.$$

□

8. Subclasses of CSPs

One way of attacking the dichotomy conjecture is to first solve it for algebraically "pure" subclasses of CSPs, and then combine these methods of solutions for more complex families. This approach is supported by a deep algebraic theorem that says that $Pol(\Gamma)$ in a sense consists of a limited number of "atomic" components. One such "pure" component is particularly puzzling.

Feder and Vardi have studied CSP problems that no linear system of equations (over a finite field) can be reduced to using gadget reductions: they called these CSP problems without the ability to count. We will denote this class by Λ.

Unlike in the case of linear equations, one might expect here some local algorithms to solve the problem.

It has turned out in the work of Larose, Valeriote and Zádori [42], that Λ can be well understood in algebraic terms. They use a branch of algebra called Tame Congruence Theory, a localization theory for finite algebras. The localization process of this theory corresponds to gadget reductions of CSP problems. In fact this theory has started to play a role much earlier: the algebraic characterization of CSP problems reducible to 3-SAT given by this theory has led to the algebraic dichotomy conjecture (Conjecture 14).

The work of Larose, Zádori and Valeriote is more involved: they manage to characterize Λ in terms of having locally no algebra that has only group operations and no algebra with only projections. Luckily, when we define Λ, we can circumvent the explanation of the Tame Congruence Theory due to a fairly recent characterization of Λ via its compatible WNUs. A non-obvious algebraic theorem of Maróti and McKenzie implies:

Definition of Λ

Theorem 51. [43] CSP(Γ) $\in \Lambda$ *if and only if there is an* $k_0 \in \mathbb{N}$ *such that for all* $k \geq k_0$ $Pol(\Gamma)$ *contains a WNU of arity* k*.*

In this write-up we take this theorem as a definition of Λ.

8.1. **Bounded width classes.** A recurring theme in combinatorics and computer science is whether consistent local solutions can be patched together into a global solution. The notion of *bounded width* intends to capture those CSPs for which local solutions can be made global.

First some definitions: A partial assignment σ with support $X \subseteq \mathbb{N}$ assigns a value from D to each variable x_i, $i \in X$. We say that σ with support on X and σ' with support on Y are *consistent* if they assign the same values to variables in $X \cap Y$. A CSP instance is *satisfied* by a partial assignment σ with support on X, if σ satisfies all constraints that take variables only from X.

Definition 52. An instance of CSP(Γ) is (k,l)-consistent ($k < l$) if there exist sets Ξ and Ξ' of partial solutions such that:

(1) Every $\sigma \in \Xi$ has support size k;
(2) Every $\sigma \in \Xi'$ has support size l;

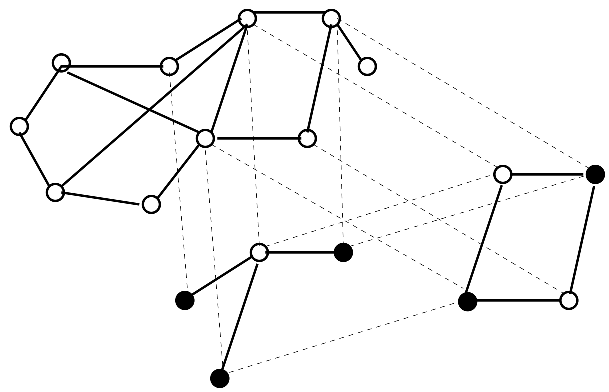


FIGURE 3. Two partial assignments to the two-coloring problem of an input graph. The assignments are consistent with each other in an intersection of size three

(3) Every $\sigma \in \Xi$ satisfies the instance;
(4) Every $\sigma \in \Xi'$ satisfies the instance;
(5) For every $|X| = k$, $X \subseteq Y$ and $|Y| = l$, and $\sigma \in \Xi$ with support X, there exists a partial assignment $\sigma' \in \Xi'$ with support Y that is consistent with σ.
(6) For every $|X| = k$, $X \subseteq Y$ and $|Y| = l$, and $\sigma' \in \Xi'$ with support Y, there exists a partial assignment $\sigma \in \Xi$ with support X that is consistent with σ'.

Width k

Definition 53 (Width k)**.** CSP(Γ) has width k if and only if there is some (fixed) $l > k$ such that any (k, l)-consistent instance is (globally) satisfiable.

Bounded width

Definition 54 (Bounded width)**.** CSP(Γ) has bounded width (or constant width) if and only if it has width k for some fixed k.

The notion of local consistency emerged independently in graph theory [31], finite model theory [37] and algebra [17]. This was a successful direction of research in the last years: Foniok, Nešetřil and Tardif [45, 27] studied CSP problems with good characterizations in the category of relational structures with homomorphisms (with finitely many obstructions, these are called finite dualities). Rossman [50] proved the well-known Homomorphism Preservation Theorem in model theory. Dalmau, Kolaitis and Vardi [20, 36, 37] have found the connection with logic, Datalog and existential pebble games, see also Atserias [1]. Hell, Nešetřil and Zhu [31] proved that the k-consistency of a given input can be characterized by obstructions of treewidth at most $(k + 1)$.

Barto and Kozik proved that the only reason for a tractable CSP having a local, but not having a global solution is that it can solve linear equations over a finite field:

Theorem 55. [2] *Every problem in Λ has bounded width.*

9. Strong Resilience

In order to solve Theorem 55 one needed to somehow exploit the existence of WNUs in $Pol(\Gamma)$. There are only a few cases known where complexity results are shown via *general* WNUs. Considering our proof of the Hell-Nešetřil theorem via asymptotic resilience, a key idea was to "boost" the power of a WNU to obtain a term that has statistically noticeable properties. We want to achieve the same with the condition in Theorem 51.

We show that the notion of strong resilience (defined in Section 5) captures Λ at once. This new characterization (eventually a trivially equivalent one) has been an important part of the proof of Theorem 55.

A concept of "immunity" will be useful.

Definition 56. Let $f : D^n \to D$ be a function. A subset $D' \subseteq D$ is *invariant* under f, if f maps $\underbrace{D' \times \cdots \times D'}_{n \text{ times}}$ into D'.

$\daleth(f)$

Definition 57. We denote by $\daleth(f)$ the collection of $D' \subseteq D$ that is invariant under f.

Definition 58 (1-immune)**.** Let $f : D^n \to D$ be a function, $1 \leq i \leq n$, $D' \in \daleth(f)$. We say that f is *immune* with respect to (i, D') if there are constants $c_1, \ldots, c_n, c \in D'$ such that

$$f(c_1, \ldots, c_{i-1}, x, c_{i+1}, \ldots, c_n) = c \text{ for every } x \in D'.$$

Furthermore, f is *1-immune*

w.r.t.	$D' \in \daleth(f)$,	if	$\forall i$	it is 1-immune w.r.t. (i, D').
w.r.t.	i,	if	$\forall D' \in \daleth(f)$	it is 1-immune w.r.t. (i, D').
unconditionally,		if	$\forall i \forall D' \in \daleth(f)$	it is 1-immune w.r.t. (i, D').

Theorem 59. *Let* $\mathbf{A} = \langle D, F\rangle$ *be an algebra,* $|D| < \infty$*. The following are equivalent:*

(I) $\mathbf{A}$ *has arbitrary arity WNUs from a threshold (as term operations).*
(II) *There is a 1-immune WNU term operation in* $\mathbf{A}$*.*
(III) *There is a strongly resilient term operation in* $\mathbf{A}$*.*

The proof of these equivalences will occupy the rest of this section. We will also rely on:

Theorem 60 (McKenzie and Maróti [43])**.** $\mathbf{A}$ *has arbitrary arity WNUs from a threshold (as term operations) if and only if* $\mathbf{A}$ *has a WNU of arity divisible by* $|D|!$*.*

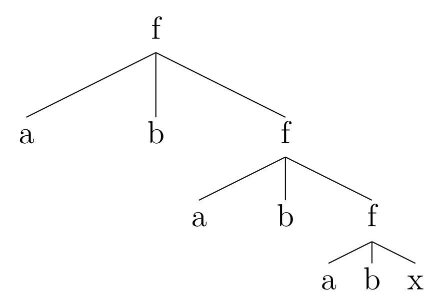
9.1. **Polynomials.** So far we have been dealing with terms. Polynomials arise from terms by setting some variables to constants.

Definition 61 (Unary Polynomial)**.** Let f be an n-ary term operation in an algebra $\mathbf{A} = \langle D, F\rangle$. We say that P is a unary polynomial created from f if for some $1 \le i \le n$ and for some constants $c_1, \ldots, c_n \in D$:

$$P(x) = f(c_1, \ldots, c_{i-1}, x, c_{i+1}, \ldots, c_n).$$

Unary polynomials are functions from D into D. We compose polynomials just like terms. Here is the composition of a unary polynomial, $f(a, b, x)$, with itself three times:

$P^3(x) =$



See Figure 9.1 how to express P^3 of the above example.

An important, and fairly obvious lemma about unary polynomials will occur in later applications:

Lemma 62. *Let* P *be a unary polynomial on* D*. Then* $Q = P^{|D|!}$ *is a polynomial for which* $Q^2 = Q$*.*

Definition 63 (General Polynomial)**.** Let f be an n-ary term of an algebra $\mathbf{A} = \langle D, F\rangle$. We say that $P(x_1, \ldots, x_k)$ is a k-ary polynomial created from f if for some $1 \le i_1 \le \ldots \le i_k \le n$ and some constants

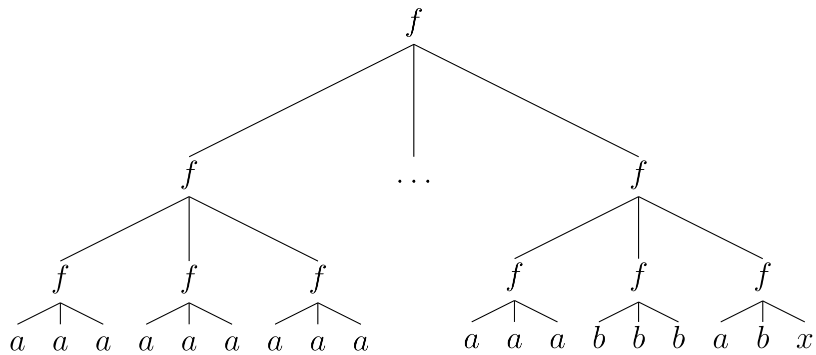


FIGURE 4. For an idempotent algebra an iterated polynomial made from a term can be expressed as a polynomial made from the iteration of the term

$f(x, y, z) \odot g(x, y) =$

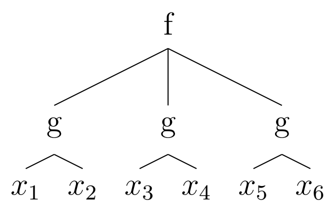


FIGURE 5. The composition of two functions is defined on disjoint variables

$c_1, \dots, c_n \in D$:

$$P(x_1, \dots, x_k) = f(c_1, \dots, c_{i_1-1}, x_1, c_{i_1+1}, \dots, c_{i_k-1}, x_k, c_{i_k+1}, \dots, c_n).$$

9.2. **Composition.** We have used the composition of a term operation f with itself, and the composition of two term operations, but we would like to make the notion a little bit more explicit. Let f be an n-ary, and g be an m-ary term operation. We define $f \odot g$ as the composition $f(\underbrace{g, \dots, g}_{n \text{ times}})$. Each inner occurrence of g receive different variables (Figure 5).

The variables of $f \odot g$ are naturally ordered in a lexicographic manner. For instance, if f is ternary and g is binary, then the variables of $f \odot g$ are ordered (for notational convenience) as

$$f(g(x_1, x_2), g(x_3, x_4), g(x_5, x_6)).$$

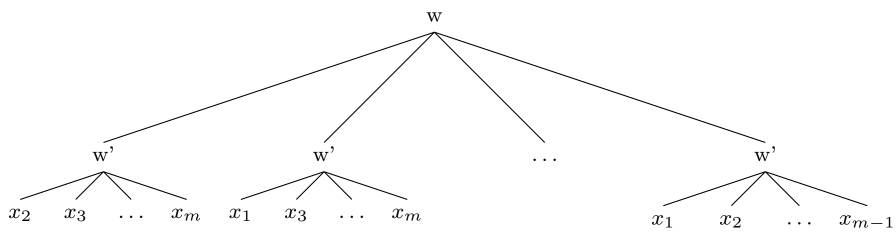


FIGURE 6. 1-immune expression made from a minority form w and WNU w'

9.3. **Proof of** $(I) \rightarrow (II)$**.** In this section we prove $(I) \rightarrow (II)$, i.e. that if $\mathbf{A}$ has arbitrary size WNUs (from a threshold) then $\mathbf{A}$ has a 1-immune WNU term operation. Undoubtedly, the best case is when $\mathbf{A}$ has a WNU term operation, w, which is a majority form (i.e. $w(x, y \dots y) = \dots w(y \dots y, x) = y$ for all $x, y \in A$). Then we do not have to look further, since w is already 1-immune.

Proposition 64. *A majority form is always 1-immune.*

This might make us to think that the worst case is when we find only *minority* forms, i.e. WNUs for which $w(x, y \dots y) = \dots w(y \dots y, x) = x$ holds for all $x, y \in A$. Luckily, there is a solution in this case, too, but, if w is n-ary, we need another (arbitrary) $(n-1)$-ary WNU:

Lemma 65. *Let w be an n-ary minority WNU and w' be an arbitrary $(n-1)$-ary WNU term operation of the algebra $\mathbf{A}$. Then $\mathbf{A}$ has a 1-immune WNU term operation.*

Proof. Consider

$$w(w'(x_2, x_3, \dots, x_n), w'(x_1, x_3, \dots, x_n), \dots, w'(x_1, x_2, \dots, x_{n-1}))$$

with $x_1, \dots, x_n$ as variables! It is straightforward that this is a majority term operation, hence 1-immune. □

Comb construction

Definition 66 (Comb construction)**.** The sequence

$$w'(x_2, x_3, \dots, x_n), w'(x_1, x_3, \dots, x_n), \dots, w'(x_1, x_2, \dots, x_{n-1})$$

of expressions we call the *Comb construction.*

Lemma 65 proves $(I) \rightarrow (II)$ whenever there is a minority term operation in $\mathbf{A}$ (since if there is a minority term operation there is also an arbitrarily large one). We generalize this idea. The general construction will be more complex, and has the following outline:

(1) Pick a WNU term operation w_0 with arity divisible by $|D|!$, with the additional property that no other WNU with arity divisible by $|D|!$ has more invariant subsets.

(2) Define the algebra $\mathbf{A}' = \langle D, w_0 \rangle$.
(3) Notice that since w_0 is (obviously) a term of $\mathbf{A}'$, and has arity divisible by $|D|!$, then by Theorem 60 the algebra $\mathbf{A}'$ has WNUs of arbitrary length from a threshold t_0.
(4) Pick a WNU term operation w in $\mathbf{A}'$ with arity $n \geq t_0 + 1$ and divisible by $|D|!$. Note that since the only operation in $\mathbf{A}'$ is w_0, it must hold that w is invariant with respect to all those sets $D' \subseteq D$ that are invariant under w_0.
(5) Select large enough K and L (to be determined later), and set $M = KL + K - 1$.
(6) Pick any WNU term w' of arity $n^L - 1$ of $\mathbf{A}'$. This w' will play the role of w' in Theorem 65.
(7) Create a term operation from w and w', by appropriately identifying the variables of $w^M \odot w'$ with each other. In the identification process we use the Comb construction, multiple times.

Items (1)$-$(4) are self-explanatory, perhaps the only comment we have to make is that the cumbersome procedure to pick w is only required to ensure that the final construction will be 1-immune with respect to *all* of the invariant subsets of the final construction. If we just wanted it to be immune with respect to D, an arbitrary (large enough) WNU term operation for w would do, and the proof could start at (5).

Next we describe the construction of item (7).

We index the variables of w^M with sequences $\vec{a} = a_1 a_2 \dots a_M$, where $a_i \in \{1, \dots, n\}$ for $1 \leq i \leq M$. (Recall that n denotes the arity of w.)

We define an equivalence relation $\mathfrak{R}$ on the variables of w^M: $\mathfrak{R}$

Definition 67 (Equivalence relation $\mathfrak{R}$)**.** $x_{\vec{a}}$ and $x_{\vec{b}}$ are equivalent with each other if and only if $a_i = b_i$ for all $i \not\equiv 0 \mod K$.

Proposition 68. *Every equivalence class of $\mathfrak{R}$ has n^L elements.*

$w^{K,L} \circledast w'$

Definition 69. We compose w^M with w' in the following way: We prepare the Comb construction on w' for every equivalence class of $\mathfrak{R}$ separately, and we replace the n^L variables of each class with the expressions of the construction. For each class we use a new set of variables. Note that the arity restriction of the Comb construction is satisfied, since the arity of w' is $n^L - 1$. The resulting formula will have n^M variables. We denote this composition by $w^{K,L} \circledast w'$.

Lemma 70. *$w^{K,L} \circledast w'$ is a WNU.*

Lemma 71. *$\daleth(w^{K,L} \circledast w') = \daleth(w) = \daleth(w_0) \subseteq \daleth(w')$.*

Proof. Since w, w' and $w^{K,L} \circledast w'$ are all term operations of $\mathbf{A}'$, they are generated (in term sense, see Definition 21) by w_0 alone. Hence any $D' \subseteq D$, which is invariant under w_0, is also invariant under w, w' and $w^{K,L} \circledast w'$.

We prove that in the case of w and $w^{K,L} \circledast w'$ the converse holds, too. This follows from the fact that both w and $w^{K,L} \circledast w'$ have n^M variables and $|D|!$ divides n. Among the term operations of $\mathbf{A}$ with arity divisible by $|D|!$, by its choice, w_0 had the greatest number of invariant subsets. Since w and $w^{K,L} \circledast w'$ are trivially term operations of $\mathbf{A}$, they can not have more. □

$\gimel$ **Definition 72.** We denote $\gimel(w^{K,L} \circledast w') = \gimel(w) = \gimel(w_0)$ by $\gimel$.

Remark 73. $\gimel$ can be alternatively defined as the collection of (the universes of) the subalgebras of $\mathbf{A}'$.

We are left with the last big question: How should we pick K and L and why will the resulting $w^{K,L} \circledast w'$ be 1-immune. What we will show is that:

Lemma 74. *There are K and L such that for every $D' \in \gimel$ the following holds: No matter which equivalence class of $\mathfrak{R}$ we are given, it is possible to set the variables of w^M to constants from D' in all but the above equivalence class, so that we get a polynomial W (obviously, with n^L variables) with the following property: there is a unary polynomial P with constants from D' such that for every $1 \leq i \leq n^L$:*

$$\forall x, y \in D' : \quad W(y, \dots, y, \underbrace{x}_{i}, y, \dots, y) = P(x).$$

Furthermore, the replacement in the definition of W respects the equivalence classes: two variables that are from the same equivalence class of $\mathfrak{R}$ receive the same constants.

Corollary 75. *$w^{K,L} \circledast w'$ is 1-immune.*

Proof. (of the corollary) Let K and L be as required by Lemma 74. Pick the class of $\mathfrak{R}$ under which the variable lies with respect to which we wish to show immunity, and pick $D' \in \gimel$, too. All constants mentioned below will be from D'. Consider the expression

$$W(w'(x_2, x_3, \dots, x_{n^L}), w'(x_1, x_3, \dots, x_{n^L}), \dots, w'(x_1, x_2, \dots, x_{n^L - 1})),$$

where W is given by the lemma. It is easy to see that there is a polynomial with coefficients from D' made from $w^{K,L} \circledast w'$, equivalent to the above expression, that contains the variable for which we wish to show immunity. (Here we used the idempotency of $\mathbf{A}$, and that the replacement hidden in W respects the equivalence classes of $\mathfrak{R}$.)

Thus all we need to show is that the above expression is immune with respect to all of its variables. Consider any variable, say x_1 w.l.g., and set all other variables to some $c \in D'$. Then W will receive $w'(c, \ldots, c) = c$, and $w'(x_1, c, \ldots, c)$, the latter $(n^L - 1)$ times. $w'(x_1, c, \ldots, c) \in D'$ by Lemma 71. By Lemma 74 the value of W will be $P(c)$, independent of x_1. □

We are left to prove Lemma 74. We show that from every WNU we can construct a polynomial W, which behaves something like a minority form. Moreover, we can embed this construction into w^M. The embedding must obey certain structural restrictions, related to the classes of $\mathfrak{R}$, that will easily follow from the proof.

We need the notion of *symmetric* polynomials of a WNU.

Definition 76 (Symmetric polynomials of a WNU)**.** Let w be an n-ary WNU. We define *symmetric* polynomials of w recursively:

(1) w itself is a symmetric polynomial of depth 1.
(2) If P and Q are symmetric polynomials of depth k and l, respectively, then $P \odot Q$ is a symmetric polynomial of depth $(k + l)$.
(3) If $P(x_1, \ldots, x_k)$ and Q are symmetric polynomials of depth k and l, respectively, and c is a constant, then $P(c, \ldots, c, Q)$ is a symmetric polynomial of depth $(k + l)$.

An example to a symmetric polynomial of depth 3 is

$$R(x_1, x_2, x_3) = w(w(w(x_1, c, c), w(x_2, c, c), w(x_3, c, c)), d, d). \tag{10}$$

Lemma 77. *Symmetric polynomials of a WNU term behave like WNUs with respect to their variables.*

We are now ready to state the main lemma:

Lemma 78. *Let w be an n-ary WNU term operation of an algebra $\langle D, F\rangle$. Then there is a* symmetric*, unary polynomial P, and a symmetric polynomial W, both of w, such that*

$$W(x, y, \ldots, y) = P(x) \qquad \textit{for all } x, y \in D.$$

Proof. Consider a symmetric unary polynomial P of w with minimal range $P(D)$, where $P(D) = \{P(x) \mid x \in D\}$. Set $r = |P(D)|$. Without loss of generality we can assume that $P(P(x)) = P(x)$, otherwise we can replace P with $P^{r!}$. Let $W_0 = (P \odot w)^{r!}$. We will show that $W = W_0 \odot P$ satisfies the conditions of the lemma. W_0 is a symmetric polynomial and its output values are all from $P(D)$. If we restrict all input values to $P(D)$ we get a WNU on $P(D)$. We need to show that

$$W_0(x, a, \ldots, a) = x \qquad \text{for every } x, a \in P(D).$$

Let a be an arbitrary element of D and consider the unary polynomial $Q_a = (P \odot w)(x, a, \ldots, a)$. Clearly Q_a is symmetric, unary, and it maps $P(D)$ into $P(D)$. Moreover, since the image of P was minimal, Q_a maps $P(D)$ *onto* $P(D)$ (!). Hence, applying Lemma 62, $Q_a^{r!}$ is the identity on $P(D)$. On the other hand, the idempotency of w and the fact that P acts as identity on $P(D)$ imply that

$$Q_a^{r!}(x) = W_0(x, a, \ldots, a) \qquad \text{for every } x, a \in P(D).$$

Thus for any $x, a \in P(D)$ we have $W_0(x, a, \ldots, a) = x$, as needed. □

To prove Lemma 74 we apply Lemma 78 to the algebra $\langle D', w_0\rangle$ and w. Our choice for K and L will be $K = \text{depth}(P) + 1$ and $L = r!$. These parameters arise in the construction of W in Lemma 78. It follows easily from the symmetry of P that $W = (P \odot w)^{r!} \odot P$ embeds into w^{KL+K-1}, with all the structural requirements (with respect to the classes of $\mathfrak{R}$), as needed.

Remark 79. The construction allows more flexibility: Denoting the above values for K and L as K_0 and L_0, any other K and L would work, where $K_0|K$ and $L_0|L$. Thus, we can pick K and L that works for all $D' \in \beth$.

Example 80. Suppose that the polynomial R in Expression (10) behaves like a minterm with respect to the variables x_1, x_2 and x_3. (Imagine that R is W that Lemma 78 gave us.) R has depth three, so we want to embed it into $w \odot w \odot w$. Let us index the 27 variables of $w \odot w \odot w$ by elements of $[3]^3$, where a triplet denotes the choices in the path from the top to the bottom. If we follow how R is constructed from w, we notice that the original variables of R map to $x_{111}, x_{121}, x_{131}$ in the embedding. We get back R from $w \odot w \odot w$ if we set $x_{112}, x_{113}, x_{122}, x_{123}, x_{132}, x_{133}$ to c, and all the remaining variables to d (using the idempotency of w).

The set $\{x_{111}, x_{121}, x_{131}\}$ represents one equivalence class of $\mathfrak{R}$. For the other classes pick the triplets

$$\{\{x_{\alpha 1\beta}, x_{\alpha 2\beta}, x_{\alpha 3\beta}\} \mid \alpha, \beta \in [3]\}.$$

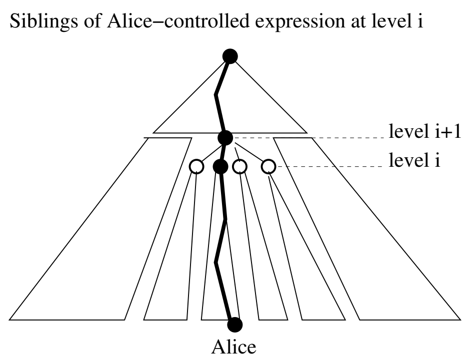
Then we have that: 1. The way we set the constants in R respects these classes. 2. For any given class we can embed R into $w \odot w \odot w$ so that its variables will embed into that class and the embedding respects all other classes.

9.4. **Proof of** $(II) \to (III)$**.** Let f be a 1-immune n-ary WNU term. We show that f is strongly resilient. Let μ be a distribution on D. By Lemma 116 there is an infinite sequence of positive integers $l(1), l(2), \ldots$

such that $\nu = \lim_{l(i)\to\infty} f^{l(i)}(\mu^{n^{l(i)}})$ exists, and the support S of ν is f-invariant.

We need to show that for a sufficiently large M the expression f^M has the property that for any $1 \leq j \leq n^M$ if Alice controls x_j and Bob sets the remaining variables randomly and independently according to μ, then for most choices of Bob, the output is the same, when Alice's choice is from the support of μ.

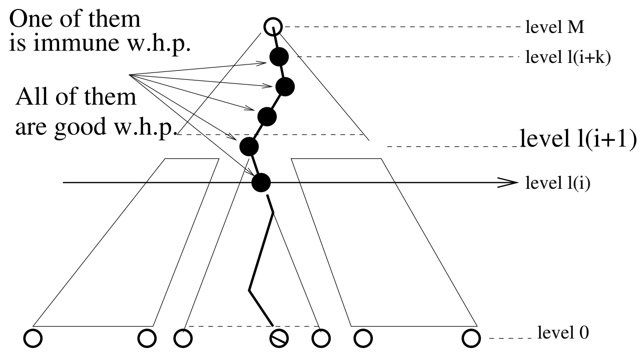
Bob's choices for the inputs of f^M he controls form an event space. When in this section we talk about events and their probabilities, we always mean them in this event space. f^M can be viewed as a tree, where every node corresponds to a sub-expression. The sub-expressions that contain Alice's input correspond to a path leading from x_j to the top. We refer to the bottom level of the tree as level zero and to the top level as level M.



Definition 81. An evaluation of Bob's variables is called immunizing at level i if the siblings of the (only) node at level i whose subtree contains x_j (i.e. the variable controlled by Alice) are evaluated to constants from S that make the parent f immune with respect to (j, S).

Lemma 82. *There is a threshold i_0 such that for every $i \geq i_0$ the probability that Bob's evaluation is immunizing at level $l(i)$ is $\geq \delta = (\frac{1}{2}\min_{s\in S}\nu(s))^{n-1}$.*

Proof. Expressions $f^{l(i)}(\mu^{n^{l(i)}})$ corresponding to separate nodes at level $l(i)$ are independent, and when i_0 is large enough, and $i \geq i_0$, their distribution is $\frac{1}{2}\min_{s\in S}\nu(s)$-close to ν. 1-immunity of f (over S) means that siblings of the Alice-influenced expression can be set to magic constants (from S) that the remaining input, assuming it is from S, does not have a say in the output at all. The probability that any given sibling evaluates to its magic constant has probability at least $\frac{1}{2}\min_{s\in S}\nu(s)$, and independence concludes the lemma. □



Given any $\epsilon > 0$, set k large enough that $1-(1-\delta)^k \geq 1-\frac{\epsilon}{2}$ ($\delta = (\frac{1}{2}\min_{s\in S}\nu(s))^{n-1}$) and $M \geq l(i_0+k)$. We apply the lemma for every level $l(i_0+1), \ldots, l(i_0+k)$, and use independence to get that

Lemma 83. *If i_0 is sufficiently large then the probability that Bob's assignment for at least one of the levels $l(i_0), \ldots, l(i_0+k)$ is immunizing is at least $1-\frac{\epsilon}{2}$.*

Definition 84. An assignment of Bob is *good* at level i if for *every* choice of Alice the (only) expression at level i dependent on Alice's choice, regardless of her choice, evaluates from S.

Lemma 85. *For every $\epsilon' > 0$ if i is large enough, the probability that Bob's assignment is good at the $l(i)$th level is at least $1-\epsilon'$.*

Proof. Since f is a WNU it is asymptotically resilient. So, if i is large enough and $1 \leq j \leq n^{l(i)}, c \in \mathrm{supp}(\mu)$ are arbitrarily fixed then the distribution $f(\mu^{j-1}, c, \mu^{n^{l(i)}-j})$ is arbitrarily close to $f(\mu^{n^{l(i)}})$. This, in turn is arbitrarily close to ν. Therefore the probability, that $f(\mu^{j-1}, c, \mu^{n^{l(i)}-j})$ is not in the support of ν, can be made arbitrarily small by setting i to be larger than a suitable threshold. Select the threshold large enough that the above probability is not larger than $\frac{\epsilon'}{|D|}$. Use the union bound for a fixed j, where c runs through the elements of $\mathrm{supp}(\mu)$. □

¿From the the above lemma and the union bound we get:

Lemma 86. *For every $\epsilon > 0$ if i_1 is large enough, then the probability that Bob's assignment is good for* all *of the levels $l(i_1+1), \ldots, l(i_1+k)$ is at least $1-\epsilon/2$.*

It is straightforward that if Bob's assignment is both good and immunizing at some level, then the top expression takes a value that does not depend on the choice of Alice. We select $M > \max\{i_0, i_1\} + k$, and we conclude from Lemmas 83 and 86 using the union bound.

9.5. **Proof of** $(III) \to (I)$**.** The proof of this direction relies on results in Tame Congruence Theory, a localization theory for finite algebras. This branch of algebra started with Pálfy [47] and was developed by Hobby and McKenzie [32], see Kiss [34, 35] for an introduction. We say that two algebras are term equivalent if they have the same base set and the same term operations.

Definition 87. A k-ary operation f on the set D is called affine if there is an Abelian group $G = \langle D, +, -, 0\rangle$ and $\varphi_1, \dots, \varphi_k \in End(G), a \in G$ such that $f(x_1, \dots, x_k) = a + \sum_{i=1}^{k} \varphi_i(x_i)$. ($End(G)$ denotes the set of endomorphisms of G, i.e. homomorphisms from G to itself.)

Remark 88. If x_i has nonzero influence according to the uniform measure then $\varphi_i \neq 0$, hence the influence of x_i is at least $\frac{1}{2}$.

A factor (divisor) of an algebra is a homomorphic image of a subalgebra. We combine the results of Maróti, McKenzie [43] and Valeriote [53] in the following theorem.

Theorem 89. *If a finite, idempotent algebra* $\mathbf{A}$ *does not have arbitrary size WNUs from a threshold (as term operations) then it has a factor which admits only affine term operations.*

Now we can prove $(III) \to (I)$. Assume that $\mathbf{A}$ has a strongly resilient term f_0, and suppose for a contradiction that it has a subalgebra S and a homomorphism φ on S such that $\varphi(S)$ has only affine operations. Consider the uniform probability measure ν on $\varphi(S)$ and a measure μ on S such that $\nu(T) = \mu(\varphi^{-1}(T))$ for every $T \subseteq \varphi(S)$. Now, since f_0 is strongly resilient, for a large enough k the term $f = f_0^k$ is such that the influence of every variable is less than $\frac{1}{2}$ according to μ. This gives a term on $\varphi(S)$ such that the influence of every variable is less than $\frac{1}{2}$ according to ν. But an affine operation does not have this property by Remark 88, a contradiction. This completes the proof of $(III) \to (I)$.

10. Width One, asymptotic invariance and symmetric functions

The case of width one CSP's is much better understood than bounded width CSP's. We have several characterizations: in terms of local consistency algorithms by Feder and Vardi [26], while Hell, Nešetřil and Zhu [31] gave one in terms of structure homomorphisms showing that these are exactly the structures with tree dualities, see also Szabó and Zádori [52]. Not surprisingly, one characterization is via $Pol(\Gamma)$.

Definition 90. If R is a k-ary relation on D, we say that $(H_1, \ldots, H_k)$ ($H_i \subseteq D$ for $1 \le i \le k$) has Property R^D if for every $1 \le i \le k$ and for every $x \in H_i$ there are $x_1, \ldots, x_k \in D$ such that $x_j \in H_j$ for $1 \le j \le k$ and

$$(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k) \in R.$$

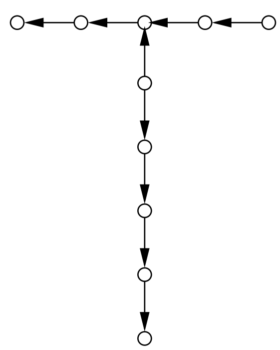
Definition 91 (Set function)**.** $\varphi : P(D) \to D$ is a *set function for* Γ if and only if φ keeps all $R \in \Gamma$ in the following sense: Whenever R is a k-ary relation in Γ and $(H_1, \ldots, H_k)$ ($H_i \subseteq D$ for $1 \le i \le k$) has Property R^D, then $f(\varphi(H_1), \ldots, \varphi(H_k)) \in R$.

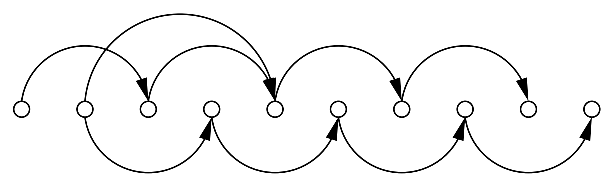
Definition 92 (Set operation)**.** $f : D^n \to D$ is a *set operation* if and only if the value of $f(x_1, \ldots, x_n)$ depends only on the set $\{x_1, \ldots, x_n\}$.

Remark 93. A set operation is obviously a WNU. A set function φ gives rise to n-ary set operations for arbitrary n by setting $f(x_1, \ldots, x_n) \stackrel{def}{=} \varphi(\{x_1, \ldots, x_n\})$.

Theorem 94 (Dalmau, Pearson [21])**.** CSP(Γ) *is width 1 if and only if there is a set function in* $Pol(\Gamma)$*.*

Corollary 95. *If* CSP(Γ) *is width 1 then* $Pol(\Gamma)$ *contains set operations of arbitrary length.*





Example 96. Let H be a directed acyclic graph on vertex set D and assume that D can be linearly ordered in such a manner that (i) all edges point forward (if (a, a') is an edge then $a < a'$) (ii) there are no two edges, (a, a') and (b, b') of H such that $b < a$ and $b' > a'$. Then the function $\max(x_1, \ldots, x_n)$ is a set operation in $Pol(H)$. Indeed, max is determined only by the set $\{x_1, \ldots, x_n\}$ (and independent of the ordering and multiplicities), so all we need to prove is that it is in $Pol(H)$. Assume that $(x_1, x_1'), \ldots, (x_n, x_n')$ are edges of H, and let $x = \max(x_1, \ldots, x_n)$, $x' = \max(x_1', \ldots, x_n')$. To show that (x, x') is an edge of H, it is sufficient to find an i such that $x = x_i$ and $x' = x_i'$. Let $x = x_j$ and $x' = x_k'$. Because of the properties of the max we have that

$x_k \leq x_j$ and $x'_j \leq x'_k$. By the properties of H both inequalities cannot be sharp. If $x_j = x_k$, we can set $i = k$, otherwise we can set $i = j$.

We conclude that for any H satisfying conditions (i) and (ii) CSP(Γ) is width one.

Example 97. Consider unique games. Since we have WNUs of arbitrary length in $Pol(\mathbf{UG}_D)$ (see Example 16), we conclude that unique games are in Λ. Does CSP($\mathbf{UG}_D$) have width one? No! Consider any $f \in Pol(\mathrm{UG}_D)$. By Theorem 113 it is possible to find α that depends only on the equality structure of $x_1, \ldots, x_n$ such that $f(x_1 \ldots, x_n) = x_{\alpha(x_1 \ldots, x_n)}$. In particular, focus on inputs for which $x_1 = \ldots = x_{n-1} \neq x_n$. Since these inputs have the same equality structure, there is an i such that for these inputs $f(x_1 \ldots, x_n) = x_i$. Let a and b be two distinct elements of D. Then, no matter what i is, f takes different values on $(a, \ldots, a, b)$ and on $(b, \ldots, b, a)$, so f cannot be a set operation. (Note that unique games are in fact width two.)

11. Asymptotic Invariance

In this section we show that asymptotic invariance is somewhere between width one and the class that contains exactly those CSP's that are solvable by the basic linear program (for more about the latter class see [22]). We can prove strict containment in the first case but not in the latter. First we show that width one implies asymptotic invariance.

Remark 98. We do not prove that if CSP(Γ) is width one then $Pol(\Gamma)$ has an asymptotically invariant term. (A term is asymptotically invariant if the invariance of its iterates goes to zero.) This holds only in special cases.

Before we show the containment we prove that the reverse is not true. We give an example to an asymptotically invariant CSP that does not provide a set operation: [4]

Example 99. Let $\mathbf{A}$ be the following algebra on the 3-element set $D = \{-1, 0, +1\}$. $\mathbf{A}$ has one k-ary operation s_k for every positive integer k: the value of s_k depends only on the average of the coordinates (as real numbers): the value is 0 if the average is in $(-\frac{1}{3}; \frac{1}{3})$, it equals to $+1$ if the average is $\geq \frac{1}{3}$ and equals to -1 if the average is $\leq -\frac{1}{3}$.

[4]An earlier version of this paper stated the equivalence of the existence of symmetric operations, asymptotic invariance and width one. This was proved in [22], but the proof of this theorem turned out to be wrong as pointed out by Victor Dalmau.

Every operation s_k is symmetric and idempotent. Consider the ternary relations $R_+, R_- \subseteq A^3$:

$$(x_1, x_2, x_3) \in R_+ \iff x_1 + x_2 + x_3 \geq 1$$
$$(x_1, x_2, x_3) \in R_- \iff x_1 + x_2 + x_3 \leq -1$$

It is easy to check that R_+ and R_- are invariant under s_k.
Then CSP(Γ) defined by R_+ and R_- does not have width one: Suppose for a contradiction that t is a compatible set operation, we might assume that t is ternary. We know that $t(++-) = t(+-+) = t(-++) = t(+--) = t(-+-) = t(--+)$, denote this value by c. The three tuples $(++-), (+-+), (-++)$ are in R_+: if we apply t to these coordinatewise we get (ccc). On the other hand, we should get an element of R_+, hence $c = +1$. The same argument with R_- shows that $c = -1$, a contradiction.

We sketch the proof that CSP(Γ) is asymptotically invariant: If μ is a measure on $\{-1, 0, 1\}$ with $\mu(1) - \mu(-1) \notin \{-\frac{1}{3}, +\frac{1}{3}\}$ then by the law of large numbers for large values of k the operation s_k will almost always take the value -1 if $\mu(1) - \mu(-1) < -\frac{1}{3}$, 1 if $\mu(1) - \mu(-1) > \frac{1}{3}$ and 0 otherwise.

If $\mu(1) - \mu(-1) = \frac{1}{3}$ then consider the output distribution μ' of $s_2(\mu, \mu)$. A simple calculation shows that $\mu'(1) - \mu'(-1) = \frac{1}{3} + \frac{1}{3}\mu(0)$. Thus, if $\mu(0) \neq 0$ we can apply the previous construction and the desired asymptotically invariant sequence will be $s_k(s_2, \dots, s_2)$ as k tends to infinity. If $\mu(0) = 0$ then μ is the $\frac{1}{3}, 0, \frac{2}{3}$ distribution and $s_2(\mu, \mu)$ is $\frac{1}{9}, \frac{4}{9}, \frac{4}{9}$. Then we iterate again with s_2 and the desired asymptotically invariant expression will be the composition of s_k with $s_2(s_2, s_2)$. The argument is similar when $\mu(1) - \mu(-1) = -\frac{1}{3}$.

Now we return to proving that if CSP(Γ) is width one then $Pol(\Gamma)$ is asymptotically invariant, i.e. for every $\varepsilon > 0$ and measure μ on D, it has an operation with invariance $< \varepsilon$. Our proof will exploit the characterization of the width one class in Theorem 94.

Lemma 100. *Let μ be a measure on D with support $S \subseteq D$, and let f be an n-ary set operation. Then there is $c \in D$ such that f takes value c with probability at least $1 - |S|(1 - \min_{x \in S} \mu(x))^n$.*

Proof. For any $x \in S$, when $x_1, \dots, x_n$ are randomly and independently chosen according to μ, the probability that $x \notin \{x_1, \dots, x_n\}$ is at most $(1 - \min_{x \in S} \mu(x))^n$. Thus with probability at least $1 - |S|(1 - \min_{x \in S} \mu(x))^n$ all elements of S are present among $\{x_1, \dots, x_n\}$. Therefore the choice of c that f takes on inputs that do not miss any element from S satisfies the conditions of the lemma. ∎

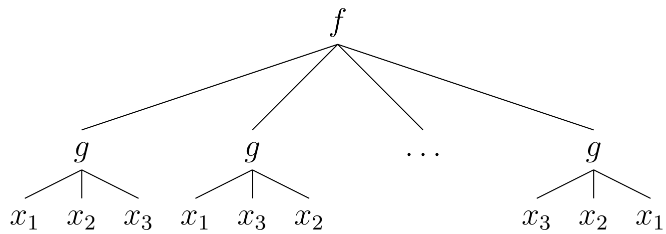


FIGURE 7. The symmetrizing product, $f \wr g$, of f with arity 6, and g with arity 3.

We now get that "width one $\rightarrow$ asymptotic invariance" as follows: By Corollary 95 if the width of CSP(Γ) is one, then $Pol(\Gamma)$ has set operations of arbitrary arity. We apply then Lemma 100 and the fact that $1 - |S|(1 - \min_{x \in S} \mu(x))^n \rightarrow 1$ as $n \rightarrow \infty$.

In the rest of the section we prove that asymptotic invariance implies the existence of symmetric operations of arbitrary arity. The latter is exactly the class of those CSP's that are solvable by the basic linear program [22]. It remains open if the converse of our theorem holds.

Lemma 101. *If an idempotent algebra $\mathbf{A} = \langle D, F \rangle$ is asymptotically invariant then it has term operations of arbitrary arity that are symmetric.*

Proof. The lemma is based on the following construction:

Definition 102. Let f be a term operation of arity $n!$, and g be a term operation of arity n. We define the *symmetrizing product* of f and g as the n-ary operation that we get by identifying variables of $f \odot g$ as follows:

(1) Write down all permutations of $\{1, \ldots, n\}$ in lexicographic order creating a sequence of length $n! \times n$, whose elements are from the set $\{1, \ldots, n\}$.
(2) Match these numbers with the $n! \times n$ variables of $f \odot g$ (in their natural order).
(3) For $1 \leq i \leq n$ replace a variable with x_i if that variable is matched with number i.

We denote the symmetrizing product of f and g by $f \wr g$ (see also Figure 7).

Definition 103. A multi-set is like a set, except that its elements may have multiple number of occurrences. The number of occurrences of an element is called its *frequency*. For instance, $\{\{1, 1, 2\}\}$ is a multi-set, where the frequency of 1 is two. $\{\{1, 2, 1\}\}$ denotes the same multi-set.

Definition 104. We say that an n-ary term f is symmetric with respect to the multi-set $\{\{c_1,\dots,c_n\}\}$ if $f(c_1,\dots,c_n)=f(c_{\pi(1)},\dots,c_{\pi(n)})$ for every permutation $\pi\in S_n$.

The following is a simple consequence of the idempotency of $\mathbf{A}$.

Lemma 105. *Let f be a term operation of arity $n!$, and g be a term operation of arity n for an idempotent algebra $\mathbf{A}$. Then if g is symmetric with respect to the multi-set $\{\{c_1,\dots,c_n\}\}$ then so is $f\wr g$.*

Note that an n-ary term is symmetric if it is symmetric with respect to all multi-sets of size n. The strategy of our proof is that we start from an arbitrary n-ary term g_0 and use an appropriate sequence (f_i) to create a sequence $g_1=f_0\wr g_0$; $g_2=f_1\wr g_1$; e.t.c. such that g_{i+1} is symmetric with respect to (at least) one more multi-set than g_i (while not loosing previous symmetries). Next we explain how to construct the sequence (f_i).

Let $H=\{\{c_1,\dots,c_n\}\}$ be a multi-set and g be an n-ary term. Define the multi-set of size (the sum of frequencies) $n!$:

$$H_g=\{\{g(c_{\pi(1)},\dots,c_{\pi(n)})\mid\ \pi\in S_n\}\}$$

The following is easy to show:

Lemma 106. *Let f be a term operation of arity $n!$, and g be a term operation of arity n for $\mathbf{A}$. If f is symmetric with respect to H_g then $f\wr g$ is symmetric with respect to H.*

The above lemma tells how to symmetrize g for a multi-set H : find a term operation f, which is symmetric w.r.t. H_g, then form the symmetrizing product. How do we find such an f? We create it from a sufficiently invariant function using the following lemma:

Lemma 107. *Let K be a multi-set of size n (the sum of the frequencies is n) and let μ_K be the measure on D, where for every $x\in D$ the measure of $\{x\}$ is the frequency of x divided by n. Assume that the operation f_0 is $\frac{1}{n!+1}$-invariant with respect to μ_K. Then there exists an identification of the variables of f_0 that the resulting f is symmetric with respect to K.*

Proof. Let $K=\{\{c_1,\dots,c_n\}\}$. We create $f(x_1,\dots,x_n)$ as follows: replace every variable of f_0 randomly, uniformly and independently by one of the variables $x_1,\dots,x_n$. We claim that this randomized construction works with positive probability.

Consider a permutation $\pi\in S_n$. The replacement $x_i\to c_{\pi(i)}$ $(1\le i\le n)$ of the variables of $f(x_1,\dots,x_n)$ is exactly a random assignment

Algorithm 1

```
(1) Select any g of arity n.
(2) While ∃ multi-set H such that g is not symmetric
    with respect to H:
    (a) Find an f which is symmetric with respect to H_g
        (using Lemma 107).
    (b) Construct f ≀ g.  This, by Lemma 106 is symmetric
        with respect to H, and by Lemma 105 keeps all
        symmetries of g.
    (c) Replace g by f ≀ g.
(3) Output g.
```

FIGURE 8. asymptotic invariance $\to$ $\forall$ n $\exists$ symmetric term operation of arity n.

to the variables of f_0, where each variable receives values independently according to μ_K.

If f_0 is $\frac{1}{n!+1}$-invariant, then the probability, that for *every* permutation $\pi \in S_n$ the $x_i \to c_{\pi(i)}$ $(1 \leq i \leq n)$ replacement results in output c, is at least $\frac{1}{n!+1}$. This follows from the union bound, when applied to the $n!$ choices of π. Hence f is symmetric with respect to K with non-zero probability, so the required f exists. □

Algorithm 1 in Figure 8 recaps how we obtain the symmetric function. □

References

[1] A. Atserias, *On Digraph Coloring Problems and Treewidth Duality*, in: Proceedings of the Twentieth Annual, IEEE Symp. on Logic in Computer Science, LICS 2005, 2005, pp. 106–115.

[2] L. Barto, M. Kozik, *Constraint Satisfaction Problems of Bounded Width*, in: Proceedings of the 50th Symposium on Foundations of Computer Science, FOCS'09, 595-603,

[3] L. Barto, M. Kozik, *Robust Satisfiability of Constraint Satisfaction Problems*, in: Proceedings of the 44th ACM Symposium on Theory of Computing, STOC'12, 931-940,

[4] L. Barto, M. Kozik, T. Niven, *CSP dichotomy holds for digraphs with no sink and source*, SIAM J. on Computing 38 (2009), Issue 5, 1782-1802,

[5] J. Berman, P. Idziak, P. Marković, R. McKenzie, M. Valeriote, R. Willard, *Tractability and learnability arising from algebras with few subpowers*, Proceedings of the 22nd IEEE Symposium on Logic in Computer Science (LICS 2007), (2007), 213–224.

[6] M. Bodirsky, J. Nešetřil: *Constraint Satisfaction with Countable Homogeneous Templates*. J. Log. Comput. 16(3): 359-373 (2006).

[7] A. Bulatov, *A dichotomy theorem for constraints on a three-element set*, Journal of the ACM, 53(1), 2006, 66-120.

[8] A. Bulatov, *Tractable conservative constraint satisfaction problems*, Proceedings of the 18th IEEE Symposium on Logic in Computer Science (LICS 2003), (2003), 321–330.

[9] A. Bulatov, *H-coloring dichotomy revisited*, Theoret. Sci Comp. Sci. 349,1 (2005), 31–39.

[10] A. Bulatov, *A graph of a relational structure and Constraint Satisfaction Problems*, In: Proceedings of the 19th IEEE Symposium on Logic in Computer Science, (LICS04), 2004.

[11] A. Bulatov, V.Dalmau *Mal'tsev constraints are tractable*, SIAM J. on Computing, 36(1), 2006, 16-27.

[12] A. Bulatov, V. Dalmau, *Towards a dichotomy theorem for the counting constraint satisfaction problem*, Information and Computation, 205(5), 2007, pp. 651-678.

[13] A. Bulatov, P. Jeavons, *Algebraic structures in combinatorial problems*, Tech. Rep. MATH-AL-4-2001, Technische Universitat Dresden, Dresden, Germany. available at http://web.comlab.ox.ac.uk/oucl/research/areas/constraints/publications/index.html.

[14] A. Bulatov, P. Jeavons, A. A. Krokhin, *Constraint satisfaction problems and finite algebras*, Automata, languages and programming (Geneva, 2001), Lecture notes in Comput. Sci., **1853**, Springer, Berlin, (2002), 272–282.

[15] H. Chen, *Quantified Constraint Satisfaction and the polynomially generated powers property*, 35th Colloquium on Automata, Languages and Programming, (ICALP), 2008.

[16] P. G. Jeavons, D. A. Cohen and M. Cooper, *Constraints consistency and closure*, Artificial Intelligence, **101** (1998) 251–265.

[17] P. G. Jeavons, D. A. Cohen and M. Gyssens, *Closure properties of constraints*, J. of the ACM **44**, 1997, 527–548.

[18] V. Dalmau, *Generalized majority-minority operations are tractable*, LICS05.

[19] V. Dalmau and A. Krokhin, *Robust satisfiability for CSPs: hardness and algorithmic results*, ACM Transactions on Computation Theory, 5(4), 2013, 15:1-15:25.

[20] V. Dalmau, P. G. Kolaitis, M. Vardi, *Conjunctive query-containment and constraint satisfaction*, Journal of Computer and System Sciences, **61**(2):302–332, 2000.

[21] V. Dalmau and J. Pearson, *Set Functions and Width 1 Problems*, Principles and Practice of Constraint Programming, CP'99.

[22] Ryan O'Donnell, Gábor Kun, Suguru Tamaki, Yuichi Yoshida and Yuan Zhou: *Linear programming, width-1 CSP's and robust satisfaction*, Innovations in Theoretical Computer Science 2012,

[23] Irit Dinur, *The PCP Theorem by gap amplification* JACM, Proc. of 38th STOC, pp. 241-250, 2006.

[24] Irit Dinur, Ehud Friedgut, Oded Regev, *Independent Sets in Graph Powers are Almost Contained in Juntas*, Geometric and Functional Analysis, accepted.

[25] Irit Dinur, Elchanan Mossel, Oded Regev, *Conditional Hardness for Approximate Coloring*, Proc. of 38th STOC, pp. 344-353, 2006.

[26] T. Feder, M. Y. Vardi, *The computational stucture of monotone monadic (SNP) and constraint satisfaction: A study through datalog and group theory*, SIAM Journal of Computing **28**, (1998), 57–104.

[27] J. Foniok, J. Nešetřil, C. Tardif, *Generalized dualities and maximal finite antichains in the homomorphism order of relational structures*, European Journal of Combinatorics 29(4), 2008, 881-899.

[28] J. Hastad, *Some optimal inapproximability results*, J. ACM 48(4): 798-859 (2001).

[29] P. Hell, J. Nešetřil, *On the complexity of H-coloring*, J. Combin. Theory Ser. B, **48**, (1990), 92–110.

[30] P. Hell, J. Nešetřil, *Colouring, Constraint Satisfaction and Complexity*, Computer Science Review 2(3), 2008, 143-163.

[31] P. Hell, J. Nešetřil, X. Zhu, *Duality and polynomial testing of tree homomorphisms*, Transactions of the American Mathematical Society, **348**(4):1281–1297, 1996.

[32] D. Hobby, R. McKenzie, *The structure of finite algebras*, Vol. 76 of Contemporary Mathematics, AMS, 1988.

[33] P. G. Jeavons, *On the algebraic structure of combinatorial problems*, Theoretical Computer Science, (1998), **200**, 185-204.

[34] E. W. Kiss, *An easy way to minimal algebras*, Int. J. of Alg. and Comp., **7**, 55–75, 1997.

[35] E. W. Kiss, *An Introduction to Tame Congruence Theory*, Proceedings of the 1996 NATO ASI Workshop on Algebraic Model Theory, 119–143, Kluwer, 1997.

[36] P. G. Kolaitis and M. Y. Vardi, *Conjunctive query containment and constraint satisfaction*, in: Proceedings of the seventeenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems, Seattle, Washington, 1998, pp. 205–213.

[37] P. G. Kolaitis, M. Vardi, *A game-theoretic approach to constraint satisfaction*, In Proceedings of the 17th National (US) Conference on Artificial Intelligence, AAAI'00, 175–181, 2000.

[38] G. Kun, *Constraints, MMSNP and expander relational structures*, Combinatorica 33(3), 335-347, 2013.

[39] G. Kun, J. Nešetřil: *Forbidden lifts (NP and CSP for combinatorists)*, European J. Comb. 29, (2008), 930-945.

[40] R. E. Ladner: On the structure of Polynomial Time Reducibility, Journal of the ACM, 22,1 (1975), 155–171.

[41] B. Larose, L. Zádori, *Bounded width problems and algebras*, Algebra Universalis, **56**, 439–466, 2007.

[42] B. Larose, L. Zádori, M. Valeriote, *Omitting types, bounded width and the ability to count*, manuscript, 2008.

[43] R. McKenzie, M. Maróti: *Existence theorems for weakly symmetric operations*, Algebra Universalis, 59(3-4), 2008, 463-489.

[44] J. Nešetřil, M. Siggers, L. Zádori: *A Combinatorial constraint satisfaction problem dichotomy classification conjecture*, European Journal of Combinatorics 31(1), 2010, 280-296.

[45] J. Nešetřil and C. Tardif, *Duality theorems for finite structures (characterising gaps and good characterizations)*, J. Combin. Theory B **80** (2000), 80–97.

[46] J. Nešetřil and X. Zhu, *On bounded treewidth dualitiy of graphs*, J. of Graph Theory **23**, 151–162, 1996.

[47] P. P. Pálfy, *Unary polynomials in algebras I.*, Algebra Univ. **18**, 262–273, 1984.

[48] Y. Rabinovich, A. Sinclair, A. Wigderson: *Quadratic Dynamical Systems (Preliminary Version)* FOCS 1992: 304-313.

[49] P. Raghavendra, *An Efficient Algorithm via Polymorphisms in the Value Oracle Model*, The Constraint Satisfaction Problem: Complexity and Approximability (2013): 16.

[50] B. Rossman, *Existential positive types and preservation under homomorphisms*, In 20th IEEE Symposium on Logic in Computer Science, 2005.

[51] T. J. Schaefer, *The complexity of satisfiability problems*, Proc. of the 10th STOC, pp. 216–226, 1978.

[52] Cs. Szabó, L. Zádori, *Idempotent totally symmetric operatins on finite posets*, Order **18**, 39–47, 2001 .

[53] M. Valeriote, *A subalgebra intersection property for congruence distributive varieties*, Canadian Journal of Mathematics, 61(2), 2009 451-464.

12. Appendix

12.1. **Unique Games.** For another example we compute $Pol(\Gamma)$ of (unrestricted) unique games. Let π be a permutation on D, and let

$$R_\pi = \{(x, \pi(x)) \mid \ x \in D\}.$$

Unique games with label set D are defined by the set of binary relations

$$\mathrm{UG}_D = \{R_\pi \mid \ \pi : D \to D \text{ is a permutation}\}.$$

What are the term operations that leave $\Gamma = \mathrm{UG}_D$ invariant?

Lemma 108. *$f(x_1, \ldots, x_n)$ leaves R_π invariant if and only if*

$$f(\pi(x_1), \ldots, \pi(x_n)) = \pi(f(x_1, \ldots, x_n)) \tag{11}$$

for all $x_1, \ldots, x_n \in D$.

Remark 109. Note that the invariance of f with respect to R_π only says that if all of the pairs $(x_1, \pi(x_1)), \ldots, (x_n, \pi(x_n))$ are in R_π, then

$$(f(x_1, \ldots, x_n), f(\pi(x_1), \ldots, \pi(x_n))) \in R_\pi,$$

But by the 1-1 property of permutations this immediately gives the lemma.

Corollary 110. *$f(x_1, \ldots, x_n) \in Pol(\mathrm{UG}_D)$ iff Equation (11) holds for every $\pi : D \to D$.*

To explicitly describe the elements of $Pol(\mathrm{UG}_D)$ we need some definitions. We define an equivalence relation $\Re$ on D^n: $(x_1, \ldots, x_n)$ belongs to the same equivalence class as $(x'_1, \ldots, x'_n)$ if and only if:

$$\forall 1 \le i, j \le n : \ x_i = x_j \ \leftrightarrow x'_i = x'_j. \tag{12}$$

If the above holds, we say that $(x_1, \ldots, x_n)$ and $(x'_1, \ldots, x'_n)$ have the same *equality structure.*

Lemma 111. *$(x_1, \ldots, x_n)$ and $(x'_1, \ldots, x'_n)$ have the same equality structure if and only if there is a permutation $\pi : D \to D$ such that $x'_i = \pi(x_i)$ for $1 \le i \le n$.*

We need yet another technical definition. To $x_1, \ldots, x_n \in D$ define

$$x_0 = \begin{cases} a & \text{if } D \setminus \{x_1, \ldots, x_n\} = \{a\}, \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Whether x_0 is defined or not depends only on the equality structure of $x_1, \ldots, x_n$.

Lemma 112. *If* $(x_1, \dots, x_n)$ *and* $(x'_1, \dots, x'_n)$ *have the same equality structure, and this equality structure allows for the definition of* x_0*, then there is a unique permutation* $\pi : D \to D$ *such that* $x'_i = \pi(x_i)$ *for* $0 \le i \le n$*.*

Theorem 113. $f \in Pol(\mathrm{UG}_D, n)$ *if and only if there exists* $\alpha : D^n \to \{0, 1, \dots, n\}$ *that is constant on the classes of* $\Re$ *such that*

$$f(x_1 \dots, x_n) = x_{\alpha(x_1 \dots, x_n)}. \tag{13}$$

Proof. First we show that every f with the structure as in the theorem, satisfies Equation (11) for all R_π. Proposition 110 will then imply that $f \in Pol(\mathrm{UG}_D, n)$. For any $\pi : D \to D$ the tuples $(x_1 \dots, x_n)$ and $(\pi(x_1) \dots, \pi(x_n))$ have the same equality structure. Hence

$$\alpha(x_1 \dots, x_n) = \alpha(\pi(x_1) \dots, \pi(x_n)) \stackrel{def}{=} \alpha.$$

Then $f(x_1 \dots, x_n) = x_\alpha$ and $f(\pi(x_1) \dots, \pi(x_n)) = \pi(x_\alpha)$, implying Equation (11). Note that in the $\alpha = 0$ case we used Lemma 112.

Next we show that if f is an n-ary term operation in $Pol(\mathrm{UG}_D)$, then it is necessarily of the form (13).

Lemma 114. *If* $f \in Pol(\mathrm{UG}_D)$*, then for every evaluation* $f(x_1, \dots, x_n) \in \{x_0, \dots, x_n\}$ *(where we omit* x_0 *from the r.h.s. if not defined).*

Proof. What we need to prove is that if $|\{x_1, \dots, x_n\}| \le |D| - 2$ then $f(x_1, \dots, x_n) \in \{x_1, \dots, x_n\}$. Consider π that fixes $x_1, \dots, x_n$, and un-fixes the remaining elements of D (e.g. cycles them). Then the assumption that $f(x_1, \dots, x_n)$ equals to any of the un-fixed elements, gives an immediate contradiction via Proposition 110 applied to π. □

Lemma 115. *If* $(x_1, \dots, x_n)$ *and* $(x'_1, \dots, x'_n)$ *have the same equality structure and* $f(x_1, \dots, x_n) = x_i$ *then* $f(x'_1, \dots, x'_n) = x'_i$

Proof. There exists a permutation π on D such that $\pi(x_i) = x'_i$ for $1 \le i \le n$ (and for $i = 0$, if applies). Applying Proposition 110 to π we get the lemma. □

□

12.2. **Convergence.** In this section we prove a lemma we need in Section 9.4.

Let $f : D^n \to D$ be a function. If we iterate f, and compute the output distribution, where the input variables for the expression tree are independently chosen from a fixed distribution μ on D, the resulting distributions do not necessary converge to a fixed distribution, even when f is idempotent. But we can still say something. Compactness of the space of distributions on D (with respect to the statistical difference, δ, as distance) implies that the set of limiting distributions $\mathfrak{L}(f,\mu)$ of the sequence $\mu_k = f^k(\mu^{n^k})$ is non-empty. Besides the non-emptyness we need yet another property of $\mathfrak{L}(f,\mu)$:

Lemma 116. *Let $f : D^n \to D$ be idempotent. There is a $\nu \in \mathcal{L}(f,\mu)$ such that $S = \mathrm{supp}(\nu)$ is invariant under f.*

Proof. We subdivide the simple proof into lemmas.

Lemma 117. *If ν is $\mathcal{L}(f,\mu)$, then so is $f(\nu^n)$.*

Proof. Let $\delta(f^k(\mu^{n^k}),\nu) \leq \epsilon$. Then $\delta(f^{k+1}(\mu^{n^{k+1}}), f(\nu^n)) \leq n\epsilon$ by Corollary 38. Since n is a fixed constant, $n\epsilon$ tends to zero when ϵ tends to zero. ☐

Definition 118. Let D be a finite set, $f : D^n \to D$ be a function, and $D' \subseteq D$. Define:

$$f(D') \stackrel{def}{=} \{x \mid \exists x_1, \ldots, x_n \in D' : x = f(x_1, \ldots, x_n)\}.$$

Remark 119. If f is idempotent, then $D' \subseteq f(D')$.

Lemma 120. *If $S \subseteq D$ occurs as a support of some $\nu \in \mathcal{L}(f,\mu)$, then so does $f(S)$.*

Proof. $\mathrm{supp}(f(\nu^n)) = f(\mathrm{supp}(\nu))$. Use Lemma 117. ☐

We are now ready to finish the proof of Lemma 116. Select $\nu \in \mathcal{L}(f,\mu)$ with maximal support. The idempotency of f implies $S \subseteq f(S)$. Lemma 120 and the maximality of S implies that $f(S) = S$. ☐