Dichotomies in Constraint Satisfaction: Canonical Functions and Numeric CSPs

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt

dem Bereich Mathematik und Naturwissenschaften
der Technischen Universität Dresden

von

MSc Antoine Mottet

geboren am 20. Oktober 1992 in Cherbourg

Eingereicht am 31. Mai 2018
Verteidigt am 14. August 2018

Die Dissertation wurde in der Zeit von September 2015 bis Mai 2018
im Institut für Algebra angefertigt.

Contents

1 Introduction **1**

2 Preliminaries **7**
2.1 Constraint Satisfaction Problems . . . 7
2.2 Primitive Positive Constructions . . . 9
2.3 Logic and Model Theory . . . 11
2.3.1 Countable categoricity . . . 11
2.3.2 Homogeneity . . . 13
2.3.3 Preservation . . . 14
2.4 Topology . . . 15
2.5 Universal Algebra . . . 16
2.5.1 Clones . . . 16
2.5.2 Identities . . . 18
2.6 Model-completeness, Cores . . . 19
2.7 The Infinite-Domain Tractability Conjecture . . . 21

I Lifting Techniques **23**

3 Equations and Tractability Conditions **25**
3.1 The Type Structure . . . 25
3.2 Canonical Functions . . . 29
3.3 New Abstract Tractability Conditions . . . 33

4 Mashups **37**
4.1 Mashups . . . 37
4.2 Disjoint Unions of Structures . . . 41
4.3 Reducts of Unary Structures . . . 43
4.3.1 The case of tame endomorphisms . . . 44
4.3.2 The general case . . . 48

5 MMSNP: Proof of the Algebraic Dichotomy Conjecture **53**
5.1 MMSNP . . . 55
5.1.1 Connected MMSNP . . . 55
5.1.2 Templates for connected MMSNP sentences . . . 56
5.1.3 Statement of the main result . . . 57
5.2 Normal Forms . . . 58

5.2.1 The normal form for MMSNP . . . 58
5.2.2 Templates for sentences in normal form . . . 61
5.2.3 The strong normal form . . . 67
5.3 Precoloured MMSNP . . . 70
5.3.1 Basic properties of precoloured MMSNP . . . 70
5.3.2 Adding inequality . . . 70
5.3.3 The standard precolouration . . . 73
5.3.4 Proof of the precolouring theorem . . . 73
5.4 An Algebraic Dichotomy for MMSNP . . . 77
5.4.1 The tractable case . . . 78
5.4.2 The hard case . . . 80
5.4.3 The dichotomy: conclusion . . . 87

II Numeric CSPs 91

6 The Complexity of Discrete Temporal CSPs 93
6.1 Discrete Temporal Constraint Satisfaction Problems . . . 94
6.2 Model-Theoretic Considerations . . . 95
6.3 Detailed Statement of the Results . . . 98
6.4 Definability of Successor and Order . . . 99
6.4.1 Degrees . . . 100
6.4.2 Compactness . . . 100
6.4.3 Finite-range Endomorphisms . . . 104
6.4.4 Petrus . . . 107
6.4.5 Boundedness and rank . . . 109
6.4.6 Defining succ and $<$ . . . 119
6.5 Tractable Classes . . . 120
6.5.1 The Horn case . . . 121
6.5.2 Modular minimum and modular maximum . . . 122
6.6 The Classification . . . 122
6.6.1 First-order expansions of $(\mathbb{Z}; \mathrm{succ}, <)$ . . . 123
6.6.2 Endomorphisms of and Definability in positive reducts . . . 123
6.6.3 The non-positive case . . . 125
6.6.4 The positive case . . . 129
6.6.5 Concluding the classification . . . 132
6.7 Conclusion . . . 133

7 Disjunctive Linear Diophantine Constraints 135
7.1 Preliminaries . . . 136
7.2 Cores . . . 137
7.3 Hardness . . . 138
7.3.1 The fully modular case . . . 139
7.3.2 The unary case . . . 139
7.3.3 Arbitrary arities . . . 142
7.4 Tractability . . . 145
7.5 Conclusion . . . 147

Bibliography **149**

Index **155**

List of Figures **155**

Ever tried. Ever failed. No matter.
Try again. Fail again. Fail better.
Samuel Beckett

Acknowledgements

My experience as a PhD student has been greatly influenced by the people surrounding me and the present dissertation is the result of many interactions. I would like to give my warmest thanks:

To Max, Pavlos, and Karin. Our friendship has meant a lot to me in the past couple years and each of you has allowed me in some way to finish this project.

To Jakub, for the many interesting discussions that we had at the office, at the bar, or while climbing.

To my collaborators and co-authors, Manuel Bodirsky, Florent Madelaine, Barnaby Martin, Marcello Mamino, and Karin Quaas. It has been a pleasure to talk and work with you and I benefited a lot from our collaboration.

To my advisor, Manuel Bodirsky, for nurturing me during the past five years and for giving me his trust and the freedom to work even on the most ludicrous ideas that I had.

To QuantLA, for financing me for the past three years and for providing an environment where different scientific communities can interact. I underestimated how beneficial being part of such a program can be.

Finally, this project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 681988, "CSP-Infinity").

Chapter 1

Introduction

Constraint satisfaction problems (CSPs) form a large class of decision problems that contains numerous classical problems like the satisfiability problem for propositional formulas and the graph colourability problem. Feder and Vardi [52] gave the following logical formalization of the class of CSPs: every finite relational structure $\mathcal{A}$, the *template*, gives rise to the decision problem of determining whether there exists a homomorphism from a finite input structure $\mathcal{B}$ to $\mathcal{A}$. In their seminal paper, Feder and Vardi recognised that CSPs had a particular status in the landscape of computational complexity: despite the generality of these problems, it seemed impossible to construct NP-intermediate problems *à la Ladner* [72] within this class. The authors thus conjectured that the class of CSPs satisfies a *complexity dichotomy*, i.e., that every CSP is solvable in polynomial time or is NP-complete. The Feder-Vardi dichotomy conjecture was the motivation of an intensive line of research over the last two decades. Some of the landmarks of this research are the confirmation of the conjecture for special classes of templates, e.g., for the class of undirected graphs [55], for the class of smooth digraphs [5], and for templates with at most three elements [43, 84]. Finally, after being open for 25 years, Bulatov [44] and Zhuk [87] independently proved that the conjecture of Feder and Vardi indeed holds.

The success of the research program on the Feder-Vardi conjecture is based on the connection between constraint satisfaction problems and universal algebra. In their seminal paper, Feder and Vardi described polynomial-time algorithms for CSPs whose template satisfies some closure properties. These closure properties are properties of the *polymorphism clone* of the template and similar properties were later used to provide tractability or hardness criteria [61, 62]. Shortly thereafter, Bulatov, Jeavons, and Krokhin [46] proved that the complexity of the CSP depends only on the equational properties of the polymorphism clone of the template. They proved that trivial equational properties imply hardness of the CSP, and conjectured that the CSP is solvable in polynomial time if the polymorphism clone of the template satisfies some nontrivial equation. It is this conjecture that Bulatov and Zhuk finally proved, relying on recent developments in universal algebra. As a by-product of the fact that the delineation between polynomial-time tractability and NP-hardness can be stated algebraically, we also obtain that the *meta-problem for finite-domain CSPs* is decidable. That is, there exists an algorithm that, given a finite relational structure $\mathcal{A}$ as input, decides the complexity of the CSP of $\mathcal{A}$.

From Finite to Infinite

Since the template of a CSP is not part of the input of the problem, it is natural to also consider the case that the relational structure $\mathcal{A}$ is infinite. The class of problems that we obtain this way, the *infinite-domain CSPs* (∞CSP), strictly contains the class of finite-domain CSPs. Several problems from combinatorial optimisation or verification can be expressed as infinite-domain CSPs: feasibility of linear programs over $\mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$, and the model-checking problem for Kozen's modal μ-calculus [69] are examples of such problems that cannot be expressed as finite-domain CSPs. This increase in expressive power comes at a cost.

First, infinite-domain CSPs are not necessarily in NP. Indeed, some infinite-domain CSPs are even undecidable: the celebrated result of Matiyasevich, Robinson, Davis, and Putnam [79] states that the satisfiability problem of arbitrary polynomial equations over the integers is undecidable. In fact, every computational problem is equivalent to an infinite-domain CSP, in the sense that every decision problem is equivalent under polynomial-time Turing reductions to a CSP [16].

Secondly, the universal-algebraic approach to constraint satisfaction, so powerful in the finite setting, no longer works for arbitrary infinite-domain CSPs. More precisely, the polymorphism clone of an infinite template is no longer an invariant of the computational complexity of the associated CSP. In other words, two templates can have the same polymorphism clone while the corresponding CSPs have different complexity. Despite this negative result, Bodirsky and Nešetřil [32] showed that the universal-algebraic approach can still be used in the case that the template satisfies a well-known model theoretic property called *ω-categoricity*. Many problems from qualitative reasoning, a subfield of artificial intelligence, can be modeled with ω-categorical templates. Examples of such problems are reasoning in Allen's Interval Algebra or reasoning in the region calculi RCC-5 [13] and RCC-8 [39]. A somewhat more advanced way to recover the universal-algebraic approach is to use the model-theoretical notion of *saturation*. In a recent work, Bodirsky, Hils, and Martin [17] described a complexity invariant in terms of the polymorphism clone of a highly saturated extension of the template.

In the past decade, several classes of infinite-domain CSPs have been studied with the goal of developing a general theory of infinite-domain constraint satisfaction. On the negative side, it was shown that there exist CSPs with an ω-categorical template that are in coNP but neither in P nor coNP-hard [16] (under the assumption that $\mathsf{P} \neq \mathsf{NP}$). Actually, the same phenomenon is already observed for countable homogeneous structures in a finite relational language, which form a subclass of the class of ω-categorical structures. On the positive side, much is now known about the CSPs of *finitely bounded homogeneous* structures, which are homogeneous structures whose class of finite substructures has a finite universal first-order axiomatisation. In a growing body of work [20, 21, 22, 29, 34, 40, 68], special cases have been investigated and the authors proved that the classes under study admit a complexity dichotomy. Moreover, they showed that the delineation between polynomial-time solvability and NP-hardness can be described using algebraic and topological methods. This led Bodirsky and Pinsker to conjecture a possible generalisation of the finite-domain complexity dichotomy for the class of all CSPs that can be formulated with a template that is a first-order reduct of a finitely bounded homogeneous structure.

Contributions

The focus of this dissertation is to develop methods to study the complexity of infinite-domain CSPs and to employ these methods to show that some restricted subclasses of infinite-domain CSPs exhibit a complexity dichotomy. In the first part (Chapters 3, 4, and 5), I focus on ω-categorical templates, and more precisely on templates that fall into the class of first-order reducts of finitely bounded homogeneous structures. Bodirsky and Pinsker conjectured that this class of templates does have a P/NP-complete complexity dichotomy. In Chapter 3, I present a complexity-theoretic reduction from a large class of infinite-domain CSPs to finite-domain CSPs. This reduction preserves important algebraic properties of the templates and is used to *lift* algebraic conditions that are known for finite-domain CSPs. In particular, we are able to lift the algebraic conditions for membership in P and membership in the class Datalog. This reduction is quite general and provides a unified algorithm for many problems that were studied in the literature. For example, the tractable Graph Satisfiability problems [34], the tractable fragments of RCC-5 [63], and the tractable equivalence CSPs [40] can all be solved using the algorithm given in Chapter 3 in combination with any algorithm for finite-domain CSPs.

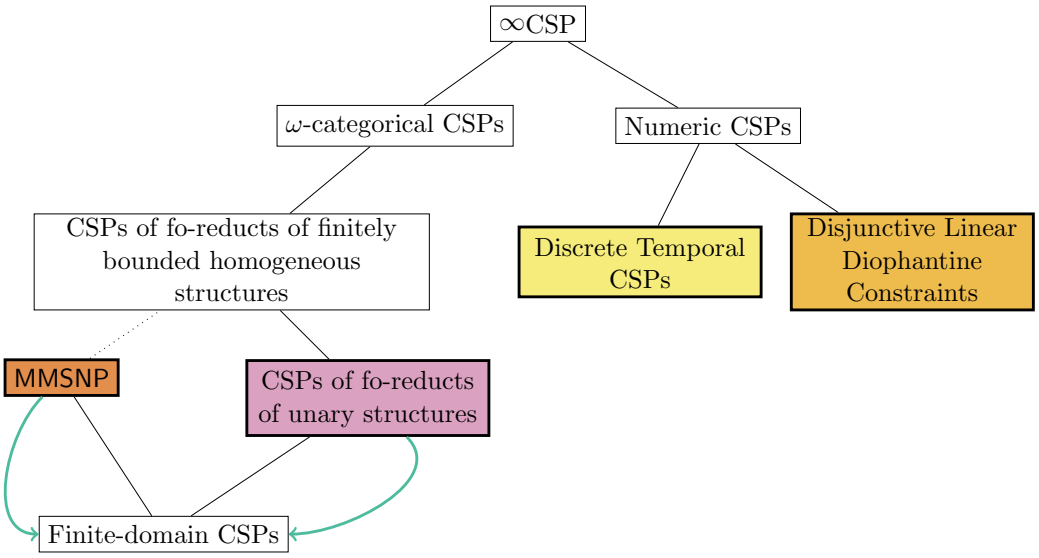

Figure 1.1: Overview of the different classes of problems involved in this work, partially ordered by inclusion. In colour are the classes for which we establish a complexity dichotomy and describe the border between polynomial-time tractability and NP-completeness using algebraic methods.

In Chapter 4, we develop another technique that we use, informally, to lift NP-hardness from the finite. This technique is based on the notion of *mashups* that was first defined to study first-order reducts of equality with finitely many constants [30]. The combination of Chapter 3 and Chapter 4 gives a powerful tool to approach the tractability conjecture by Bodirsky and Pinsker by exploiting the dichotomy from finite-domain CSPs. In particular, we apply these techniques to study the structures that are first-order reducts of any structure with a *unary* signature. This class contains all the CSPs with finite domain as well as all the equality CSPs, which were studied by Bodirsky and

Kára [21].

Chapter 5 is about CSPs that can be described in the fragment of second-order logic called MMSNP (Monotone Monadic Strict NP). This fragment was studied by Feder and Vardi [52], whose goal was to find large fragments of existential second-order logic that do not have the ability to encode NP-intermediate problems. Feder and Vardi showed that MMSNP and finite-domain CSPs are equivalent under randomised polynomial-time reductions, and this reduction was later derandomised by Kun [71]. Combining this result with the complexity dichotomy for finite-domain CSPs [44, 87], this implies that the logic MMSNP indeed has a dichotomy. A result of Bodirsky and Dalmau [14] implies that the problems in MMSNP can be seen as CSPs of infinite-domain structures that fall into the class of Bodirsky and Pinsker. In this chapter, we prove that the Bodirsky-Pinsker conjecture holds for CSPs that can be described in MMSNP. The polynomial-time tractable fragments of MMSNP are obtained with the reduction from Chapter 3, while the hardness proof uses recent Ramsey-theoretic results [58]. As a corollary, we obtain a new proof of the complexity dichotomy for MMSNP.

In the second part of this dissertation, I focus on so-called *numeric CSPs* [25]. These are CSPs whose template is a structure over $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, or $\mathbb{C}$, and whose relations are definable using first-order logic and arithmetic. Such templates are typically not ω-categorical. In Chapter 6, the structures under consideration are first-order reducts of the linear order $(\mathbb{Z}; <)$. The CSPs of these structures include all the CSPs of first-order reducts of $(\mathbb{Q}; <)$, which were studied by Bodirsky and Kára [23], and this class also includes the Max-Atom problem where the numeric inputs are given in unary [8]. This class is also a large generalisation of *locally finite distance CSPs* [15]. For this work, we use the aforementioned approach by Bodirsky, Hils, and Martin, and study the countably saturated extensions of such reducts. With this analysis, we are able to derive a P/NP-complete dichotomy for the CSPs of these structures. Finally, in Chapter 7, we study CSPs with *disjunctive linear Diophantine constraints*. A disjunctive linear Diophantine constraint is a relation that can be defined in first-order logic over the structure $(\mathbb{Z}; +, 1)$; such a constraint can be seen as a disjunction of systems of linear and modular equations over $\mathbb{Z}$, whence the name. In this final chapter, we consider the CSPs of templates containing $+$ in their signature and whose relations are definable by disjunctive linear Diophantine constraints. We show that each such CSP is in P or NP-complete.

Figure 1.1 contains a schematic description of the results in this dissertation as well as a general context in which these results naturally fit. Most of the results in this dissertation have appeared in the following articles:

[24] Manuel Bodirsky, Florent Madelaine, and Antoine Mottet. "A universal-algebraic proof of the dichotomy for Monotone Monadic SNP". In: *Proceedings of the Annual Symposium on Logic in Computer Science (LICS)*. 2018, pp. 105–114. DOI: 10.1145/3209108.3209156.

[26] Manuel Bodirsky, Barnaby Martin, Marcello Mamino, and Antoine Mottet. "The complexity of disjunctive linear Diophantine constraints". In: *Proceedings of the International Symposium on Mathematical Foundations of Computer Science (MFCS)*. 2018.

[27] Manuel Bodirsky, Barnaby Martin, and Antoine Mottet. "Constraint Satisfaction Problems over the Integers with Successor". In: *Proceedings of the International Colloquium on Automata, Languages and Programming (ICALP)*. 2015, pp. 256–267. DOI: 10.1007/978-3-662-47672-7_21.

[28] Manuel Bodirsky, Barnaby Martin, and Antoine Mottet. "Discrete Temporal Constraint Satisfaction Problems". In: *Journal of the ACM* 65.2 (2018), 9:1–9:41. DOI: 10.1145/3154832.

[30] Manuel Bodirsky and Antoine Mottet. "Reducts of finitely bounded homogeneous structures, and lifting tractability from finite-domain constraint satisfaction". In: *Proceedings of the Annual Symposium on Logic in Computer Science (LICS)*. 2016, pp. 623–632. DOI: 10.1145/2933575.2934515.

[31] Manuel Bodirsky and Antoine Mottet. "A Dichotomy for First-Order Reducts of Unary Structures". In: *Logical Methods in Computer Science* 14.2 (2018). DOI: 10.23638/LMCS-14(2:13)2018.

Chapter 2

Preliminaries

We present in this chapter the mathematical background that is necessary for an understanding of the results in the following chapters. We try to keep the presentation succinct and goal-oriented; the interested reader is invited to consult the classical textbooks [47, 56, 77, 86] as well as [6, 11] for the more recent material.

2.1 Constraint Satisfaction Problems

A *signature* τ is a set of function symbols and relation symbols, where with every symbol comes a positive natural number, called the *arity* of the symbol. Unless specified otherwise, all signatures are assumed to be finite. In the following, the Greek letters τ, σ, and ρ denote a signature unless mentioned otherwise. A τ*-structure* is a tuple $\mathcal{A} = (A; (Z^{\mathcal{A}})_{Z \in \tau})$ where A is a set, called the *domain* of $\mathcal{A}$ and:

- if $Z \in \tau$ is a function symbol of arity $k \in \mathbb{N}$, then $Z^{\mathcal{A}}$ is an operation $A^k \to A$,
- if $Z \in \tau$ is a relation symbol of arity $k \in \mathbb{N}$, then $Z^{\mathcal{A}}$ is a subset of A^k.

Structures will be denoted by calligraphic letters $\mathcal{A}$, and their domain by the corresponding Roman letter A. When a signature only contains relation symbols, we call it *relational*. A *relational structure* is a structure whose signature is relational. A *functional* signature is a signature that only contains function symbols. An *algebra* $\mathbf{A}$ is a structure whose signature σ is functional, and functions of the form $f^{\mathbf{A}}$ for $f \in \sigma$ are called the *fundamental operations* of $\mathbf{A}$.

Let $\mathcal{A}, \mathcal{B}$ be two structures with the same signature τ. A *homomorphism* $h \colon \mathcal{A} \to \mathcal{B}$ is a function $A \to B$ such that for every relation symbol $R \in \tau$ of arity k, we have

$$\forall (a_1, \ldots, a_k) \in A^k, \big((a_1, \ldots, a_k) \in R^{\mathcal{A}} \implies (h(a_1), \ldots, h(a_k)) \in R^{\mathcal{B}}\big), \tag{2.1}$$

and such that for every function symbol $f \in \tau$ of arity k we have $h(f^{\mathcal{A}}(a_1, \ldots, a_k)) = f^{\mathcal{B}}(h(a_1), \ldots, h(a_k))$. We also write $h(\overline{a})$ for the tuple $(h(a_1), \ldots, h(a_k))$. The notation $\mathcal{A} \to \mathcal{B}$ will be used to denote that there exists a homomorphism $h \colon \mathcal{A} \to \mathcal{B}$.

Definition 2.1. Let $\mathcal{A}$ be a relational structure with signature τ. The *constraint satisfaction problem* of $\mathcal{A}$, denoted by $\mathrm{CSP}(\mathcal{A})$, is the following computational problem:

Input: a finite structure $\mathcal{B}$ with signature τ,
Output: does there exist a homomorphism $\mathcal{B} \to \mathcal{A}$?

The structure $\mathcal{A}$ is also called the *constraint language* of CSP($\mathcal{A}$). Note that for the problem to be well-defined the signature of $\mathcal{A}$ needs to be finite, otherwise different encodings of the input structure $\mathcal{B}$ can give computational problems with different complexities.

Another way to define the CSP of a structure is as follows. A *primitive positive τ-formula* (pp-formula, for short) is a first-order formula that is built only with existential quantifications and conjunctions of positive atoms from τ (that is, universal quantifications, disjunctions, and negations are not allowed). A *sentence* is a formula without free variables.

Definition 2.2. Let $\mathcal{A}$ be a relational structure with signature τ. Let CSP$'(\mathcal{A})$ be the following computational problem:

Input: a primitive positive τ-sentence ϕ,
Output: is ϕ true in $\mathcal{A}$?

It is a folklore result that the problems CSP$'(\mathcal{A})$ and CSP($\mathcal{A}$) are equivalent up to logspace reductions. We give the proof here as it underlines the connection between pp-formulas and finite structures via the notions of canonical databases and canonical queries.

Proposition 2.1. *Let $\mathcal{A}$ be a relational structure. The problems* CSP($\mathcal{A}$) *and* CSP$'(\mathcal{A})$ *are equivalent up to logspace reductions.*

Proof. Let $\mathcal{B}$ be an input of CSP($\mathcal{A}$). Let $\phi_{\mathcal{B}}(b_1, \ldots, b_k)$ be the *canonical query* of $\mathcal{B}$. That is, $\phi_{\mathcal{B}}$ is a conjunction of the atomic formulas $R(c_1, \ldots, c_r)$ where $R \in \tau$ and $c_1, \ldots, c_r \in \{b_1, \ldots, b_k\}$ are such that $(c_1, \ldots, c_r) \in R^{\mathcal{B}}$. Thus, the variables of $\phi_{\mathcal{B}}$ are the elements of $\mathcal{B}$. We prove that $\mathcal{B} \to \mathcal{A}$ if, and only if, $\mathcal{A} \models \exists b_1, \ldots, b_k.\phi_{\mathcal{B}}(b_1, \ldots, b_k)$. Let h be a homomorphism from $\mathcal{B}$ to $\mathcal{A}$. Then it is clear that $\phi_{\mathcal{B}}(h(b_1), \ldots, h(b_k))$ holds in $\mathcal{A}$, so that $\mathcal{A} \models \exists b_1, \ldots, b_k.\phi_{\mathcal{B}}(b_1, \ldots, b_k)$ is true. Conversely, let $c_1, \ldots, c_k \in A$ be such that $\mathcal{A} \models \phi_{\mathcal{B}}(a_1, \ldots, a_k)$. It is clear that the map $h\colon b_i \mapsto c_i$ is a homomorphism from $\mathcal{B}$ to $\mathcal{A}$.

Conversely, given an input $\exists x_1, \ldots, x_k.\phi(x_1, \ldots, x_k)$ of CSP$'(\mathcal{A})$, one builds a structure $\mathcal{B}$ with domain $\{x_1, \ldots, x_k\}$ and whose relations are built from ϕ in the obvious manner. This structure is called the *canonical database* of ϕ. □

The structure $\mathcal{A}$ is called the *template* of CSP($\mathcal{A}$). When a decision problem P is said to be a CSP, we mean that there exists a structure $\mathcal{A}$ such that P and CSP($\mathcal{A}$) are the same problems, i.e., they have exactly the same set of "yes" instances. We note that if $\mathcal{A}$ has a finite domain, then CSP($\mathcal{A}$) belongs to NP.

Example 1. Let τ_{Graph} be the signature of graphs, i.e., $\tau_{Graph} = \{E\}$ where E is a binary relation symbol. Let $\mathcal{K}_r$ be the complete graph on r vertices. Given a finite graph $\mathcal{G}$, a function $h\colon \mathcal{G} \to \mathcal{K}_r$ is simply a labeling of the vertices of $\mathcal{G}$ with r colours. The function h is a homomorphism if, and only if, whenever $(a, b) \in E^{\mathcal{G}}$, we have $(h(a), h(b)) \in E^{\mathcal{K}_r}$, i.e., $h(a) \neq h(b)$. Thus, CSP($\mathcal{K}_r$) is r-Colourability, the problem of deciding whether the vertices of a given input graph can be coloured with r colours in a way that two adjacent vertices do not have the same colour. This problem is NP-complete for $r \geq 3$ and solvable in logarithmic space if $r \leq 2$.

Example 2. Consider the τ_{Graph}-structure $\mathcal{Q} = (\mathbb{Q}; E^{\mathcal{Q}})$ where $(a, b) \in E^{\mathcal{Q}}$ if, and only if, $a < b$. A directed graph $\mathcal{B}$ has a homomorphism to $\mathcal{Q}$ if, and only if, $\mathcal{B}$ does not contain any directed cycle. Thus, CSP($\mathcal{Q}$) is the Digraph-Acyclicity problem and is NL-complete.

Note that in the previous example, one could have chosen $\mathbb{N}$ or $\mathbb{Z}$ as a base set without changing the associated decision problem. Thus, it is possible for two different structures to have the same CSP.

Example 3. Let $\tau_{Arith} = \{+, \times, 1\}$ be the relational signature where $+$ and $\times$ are ternary relation symbols and 1 is a unary relation symbol. Let $\mathcal{Z} = (\mathbb{Z}; +^{\mathcal{Z}}, \times^{\mathcal{Z}}, 1^{\mathcal{Z}})$ be the following τ_{Arith}-structure:

- $(a, b, c) \in +^{\mathcal{Z}}$ if, and only if, $a + b = c$,
- $(a, b, c) \in \times^{\mathcal{Z}}$ if, and only if, $a \times b = c$,
- $a \in 1^{\mathcal{Z}}$ if, and only if, $a = 1$.

The problem $\mathrm{CSP}'(\mathcal{Z})$ is then easily seen to be equivalent to the problem of deciding whether a given system of polynomial equations has an integer solution. This problem is known as *Hilbert's tenth problem* and is undecidable [79].

2.2 Primitive Positive Constructions

The computational complexity of $\mathrm{CSP}(\mathcal{A})$ is a function of $\mathcal{A}$. In the following, we present tools that are used to investigate this function. There are several approaches for this investigation, namely the approach via *primitive positive interpretations* [46] and the approach via *primitive positive constructions* [6]. We present here the latter, for the reasons that it makes some statements simpler and that it is in general more powerful than the former.

The central notion of this section is the notion of *definability* of a relation.

Definition 2.3. Let $\mathcal{A}$ be a structure and let $R \subseteq A^k$ be a k-ary relation on A. We say that R is *first-order definable* (fo-definable) over $\mathcal{A}$ if there exists a first-order formula $\phi(x_1, \ldots, x_k)$ such that the equivalence

$$(a_1, \ldots, a_k) \in R \Longleftrightarrow \mathcal{A} \models \phi(a_1, \ldots, a_k)$$

holds for all $a_1, \ldots, a_k \in A$. The formula ϕ is said to be a *definition* of R over $\mathcal{A}$.

We say that R is *primitively positively definable* (pp-definable) over $\mathcal{A}$ if in the definition above one can take ϕ to be a pp-formula.

For an arbitrary relation $R \subseteq A^k$ over the same domain as the τ-structure $\mathcal{A}$, we write $(\mathcal{A}, R)$ for the *expansion* of $\mathcal{A}$ by R, that is, the structure with signature $\tau \cup \{\star\}$, where $\star$ is a fresh k-ary relation symbol, and such that $\star^{(\mathcal{A},R)} = R$ and $S^{(\mathcal{A},R)} = S^{\mathcal{A}}$ for every symbol $S \in \tau$. For ease of notation, we often assume that R itself is the fresh relation symbol denoting R.

Lemma 2.2 ([46])**.** *Let $\mathcal{A}$ be a relational structure and let $R \subseteq A^k$ be a relation that has a pp-definition over $\mathcal{A}$. Then* $\mathrm{CSP}(\mathcal{A}, R)$ *reduces in logarithmic space to* $\mathrm{CSP}(\mathcal{A})$*. In particular, if $\mathcal{B}$ is a relational structure over the same domain as $\mathcal{A}$ and such that every relation of $\mathcal{B}$ has a pp-definition over $\mathcal{A}$, then there is a logspace reduction from* $\mathrm{CSP}(\mathcal{B})$ *to* $\mathrm{CSP}(\mathcal{A})$*.*

We now give an example of how to use Lemma 2.2 to prove hardness of a CSP.

Example 4. Let $\mathcal{C}_5$ be the undirected 5-cycle, seen as a τ_{Graph}-structure. Let $\phi(x, y)$ be the formula

$$\exists z_1, z_2\, (E(x, z_1) \wedge E(z_1, z_2) \wedge E(z_2, y))$$

and let R be the binary relation defined by ϕ in $\mathcal{C}_5$. First, note that $E \subseteq R$, since if $(a, b) \in E$ then $E(a, b) \wedge E(b, a) \wedge E(a, b)$ holds, so that by taking $z_1 := b$ and $z_2 := a$ we obtain that $\phi(a, b)$ holds. Moreover, if a, b are at distance 2 in $\mathcal{C}_5$, it is clear that $\phi(a, b)$ holds. Finally, $\phi(a, a)$ does not hold for any $a \in C_5$. It follows that R is the full irreflexive relation on C_5. In other words, the structure (C_5, R) is the complete graph $\mathcal{K}_5$ on five vertices. We saw in Example 1 that CSP($\mathcal{K}_5$) is NP-complete, so that by Lemma 2.2 also CSP($\mathcal{C}_5$) is NP-complete.

A *primitive positive power* (pp-power) of $\mathcal{A}$ is a structure $\mathcal{B}$ with domain A^d, for $d \in \mathbb{N}$, whose k-ary relations are pp-definable when viewed as dk-ary relations over $\mathcal{A}$. Formally, for every relation $R^{\mathcal{B}}$ of $\mathcal{B}$ of arity k, there exists a pp-formula $\phi(x_1^1, \ldots, x_d^1, \ldots, x_1^k, \ldots, x_d^k)$ in the signature of $\mathcal{A}$ such that the equivalence

$$(\overline{a}^1, \ldots, \overline{a}^k) \in R^{\mathcal{B}} \Leftrightarrow \mathcal{A} \models \phi(a_1^1, \ldots, a_d^1, \ldots, a_1^k, \ldots, a_d^k)$$

holds for all $\overline{a}^1, \ldots, \overline{a}^k \in A^d$. Lemma 2.2 trivially generalises to the case that $\mathcal{B}$ is a pp-power of $\mathcal{A}$ and gives a reduction from CSP($\mathcal{B}$) to CSP($\mathcal{A}$).

Another way to transfer complexity results is with the notion of *homomorphic equivalence*. Two structures $\mathcal{A}$ and $\mathcal{B}$ with the same signature are homomorphically equivalent if there is a homomorphism from $\mathcal{A}$ to $\mathcal{B}$ and vice versa. It is clear from the definition of the CSP that if $\mathcal{A}$ and $\mathcal{B}$ are homomorphically equivalent, then CSP($\mathcal{A}$) and CSP($\mathcal{B}$) are the same computational problem.

A structure $\mathcal{B}$ is said to have a *primitive positive construction* (pp-construction) over $\mathcal{A}$ if it is homomorphically equivalent to a pp-power of $\mathcal{A}$, that is, if there exists a pp-power $\mathcal{C}$ of $\mathcal{A}$ such that $\mathcal{C}$ and $\mathcal{B}$ are homomorphically equivalent. It follows from the previous remarks that if $\mathcal{B}$ has a pp-construction over $\mathcal{A}$, then there is a logspace reduction from CSP($\mathcal{B}$) to CSP($\mathcal{A}$).

Lemma 2.3 ([6])**.** *Let $\mathcal{A}$ be a relational structure and let $\mathcal{B}$ have a pp-construction over $\mathcal{A}$. Then there is a logspace reduction from* CSP($\mathcal{B}$) *to* CSP($\mathcal{A}$)*.*

Let 1-IN-3-SAT be the structure with domain $\{0, 1\}$ and with one ternary relation $R = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. The problem CSP(1-IN-3-SAT) is known to be NP-complete [84], so that we have the following sufficient condition for NP-hardness.

Corollary 2.4. *Let $\mathcal{A}$ be a relational structure. If* 1-IN-3-SAT *has a pp-construction over $\mathcal{A}$, then* CSP($\mathcal{A}$) *is NP-hard.*

Bulatov, Jeavons, and Krokhin [46] noticed that on every example of a finite structure $\mathcal{A}$ such that CSP($\mathcal{A}$) is NP-complete, the hardness of CSP($\mathcal{A}$) could be explained by Corollary 2.4. They conjectured that if $\mathcal{A}$ is finite and cannot pp-construct 1-IN-3-SAT, then CSP($\mathcal{A}$) must be decidable in polynomial time. This conjecture was finally proven recently and independently by Bulatov [44] and Zhuk [87].

Theorem 2.5 (Finite-domain tractability theorem)**.** *Let $\mathcal{A}$ be a relational structure with a finite domain. Then exactly one of the following items holds:*

- 1-IN-3-SAT *has no pp-construction over* $\mathcal{A}$ *and* $\mathrm{CSP}(\mathcal{A})$ *is in* P*,*
- 1-IN-3-SAT *has a pp-construction over* $\mathcal{A}$ *and* $\mathrm{CSP}(\mathcal{A})$ *is* NP*-complete.*

We stress out the fact that Lemma 2.3 holds for *arbitrary* relational structures with finite signature. Therefore, understanding the complexity of $\mathrm{CSP}(\mathcal{A})$ for an arbitrary relational structure $\mathcal{A}$ amounts to understanding the complexity of $\mathrm{CSP}(\mathcal{B})$, for every $\mathcal{B}$ that has a pp-construction over $\mathcal{A}$. Indeed, for any complexity class C that is closed under logspace reductions, we obtain that $\mathrm{CSP}(\mathcal{A})$ is in C if, and only if, $\mathrm{CSP}(\mathcal{B})$ is in C for every $\mathcal{B}$ with a pp-construction over $\mathcal{A}$. This observation is the keystone that holds the edifice of constraint satisfaction and is the motivation of the model-theoretic, universal-algebraic, and topological approaches to constraint satisfaction that we present in the next sections[1].

2.3 Logic and Model Theory

An *embedding* $h\colon \mathcal{A} \hookrightarrow \mathcal{B}$ is an injective homomorphism $\mathcal{A} \hookrightarrow \mathcal{B}$ such that the statement in (2.1) is an equivalence. An *isomorphism* $h\colon \mathcal{A} \xrightarrow{\sim} \mathcal{B}$ is a surjective embedding. An *endomorphism* is a homomorphism $\mathcal{A} \to \mathcal{A}$, a *self-embedding* is an embedding $\mathcal{A} \hookrightarrow \mathcal{A}$, and an *automorphism* is an isomorphism $\mathcal{A} \xrightarrow{\sim} \mathcal{A}$. We write $\mathrm{End}(\mathcal{A})$ and $\mathrm{Aut}(\mathcal{A})$ for the sets of endomorphisms and automorphisms of a relational structure $\mathcal{A}$. Note that when equipped with composition, $\mathrm{End}(\mathcal{A})$ forms a monoid and $\mathrm{Aut}(\mathcal{A})$ forms a group.

A τ-structure $\mathcal{B}$ is a *substructure* of the τ-structure $\mathcal{A}$ (written $\mathcal{B} \subseteq \mathcal{A}$) if:

- $B \subseteq A$,
- for every relation symbol $R \in \tau$ of arity k and $b_1, \ldots, b_k \in B$, we have
$$(b_1, \ldots, b_k) \in R^{\mathcal{B}} \Leftrightarrow (b_1, \ldots, b_k) \in R^{\mathcal{A}}$$
- for every function symbol $f \in \tau$ of arity k and $b_1, \ldots, b_k$, we have
$$f^{\mathcal{B}}(b_1, \ldots, b_k) = f^{\mathcal{A}}(b_1, \ldots, b_k).$$

If $S \subseteq A$, we write $\mathcal{A}[S]$ for the structure *generated by* S *in* $\mathcal{A}$, i.e., the smallest substructure $\mathcal{B}$ of $\mathcal{A}$ and whose domain contains S. If $\mathcal{A}$ is a relational structure, then the domain of $\mathcal{B}$ is exactly S and we call $\mathcal{A}[S]$ the structure *induced by* S.

2.3.1 Countable categoricity

A countable structure $\mathcal{A}$ is ω*-categorical* if for every countable structure $\mathcal{B}$ satisfying the same first-order sentences, the structures $\mathcal{A}$ and $\mathcal{B}$ are isomorphic (i.e., there exists an isomorphism $\mathcal{A} \xrightarrow{\sim} \mathcal{B}$). An algebraic characterization of ω-categorical structures was given by Ryll-Nardzewski, Engeler, and Svenonius. Let Γ be a subgroup of the full symmetric group on A. The *orbit* of a tuple $(a_1, \ldots, a_n) \in A^n$ under Γ is the set containing the tuples

[1] We are guilty of a slight anachronism here, as pp-constructions were discovered rather recently and after the mentioned approaches. The true motivations rely on related concepts that are not introduced here and that were subsumed by pp-constructions. We justify this anachronism on the grounds of storytelling.

$(\alpha(a_1), \ldots, \alpha(a_n))$ for $\alpha \in \Gamma$. For every $n \geq 1$, we obtain an equivalence relation on n-tuples from A: $(a_1, \ldots, a_n) \sim_\Gamma (b_1, \ldots, b_n)$ if, and only if, the orbit of $(a_1, \ldots, a_n)$ under Γ is equal to the orbit of $(b_1, \ldots, b_n)$ under Γ. We say that Γ is an *oligomorphic permutation group* if for every $n \geq 1$ this equivalence relation has only finitely many equivalence classes.

Theorem 2.6 (Ryll-Nardzewski, Engeler, Svenonius, see Theorem 6.3.1 in [56])**.** *Let $\mathcal{A}$ be a countable relational structure with a countable signature. Then the following are equivalent:*

1. *$\mathcal{A}$ is ω-categorical,*
2. *$\mathrm{Aut}(\mathcal{A})$ is oligomorphic.*

Example 5. Let $\mathcal{Q} = (\mathbb{Q}; <)$ and let $\Gamma = \mathrm{Aut}(\mathcal{Q})$. By definition, Γ contains exactly the bijective maps $\alpha\colon \mathbb{Q} \to \mathbb{Q}$ that are increasing. There is only one orbit of elements (i.e., 1-tuples) under Γ: for arbitrary $a, b \in \mathbb{Q}$, the map $x \mapsto x + (b - a)$ is an increasing bijection that maps a to b and therefore $a \sim_\Gamma b$; we say that Γ is *transitive*. There are three orbits of pairs under Γ, namely the sets $\{(a, b) \in \mathbb{Q}^2 \mid a < b\}$, $\{(a, b) \in \mathbb{Q}^2 \mid a = b\}$, and $\{(a, b) \in \mathbb{Q}^2 \mid a > b\}$. Generalizing further, it can be seen that the orbit of an n-tuple $(a_1, \ldots, a_n)$ is completely determined by the weak linear order induced by the elements $a_1, \ldots, a_n$ in $\mathbb{Q}$. In particular, the number of orbits of n-tuples under Γ is finite and is called the nth Fubini number. As a consequence, Theorem 2.6 implies that $\mathcal{Q}$ is ω-categorical.

Example 6. Let $\mathcal{S} = (\mathbb{Z}; \mathrm{succ})$, where $\mathrm{succ} = \{(a, b) \in \mathbb{Z}^2 \mid b = a + 1\}$. As above, $\mathrm{Aut}(\mathcal{S})$ is transitive, but there are infinitely many orbits of pairs under $\mathrm{Aut}(\mathcal{S})$. Indeed, one sees that (a', b') is in the orbit of (a, b) under $\mathrm{Aut}(\mathcal{S})$ if and only if $b - a = b' - a'$. Therefore, $\mathcal{S}$ is not ω-categorical.

In the light of Theorem 2.6, the class of ω-categorical structures is a natural extension of the class of finite structures (as $\mathrm{Aut}(\mathcal{A})$ is trivially oligomorphic when $\mathcal{A}$ is finite). This pseudo-finiteness allows us to use some forms of "local-to-global" arguments when working with ω-categorical structures. We give here one example of such arguments.

Proposition 2.7. *Let $\mathcal{A}, \mathcal{B}$ be relational structures with a countable signature and suppose that $\mathcal{A}$ is ω-categorical. There exists a homomorphism $\mathcal{B} \to \mathcal{A}$ if, and only if, there exists a homomorphism $\mathcal{C} \to \mathcal{A}$ for every finite substructure $\mathcal{C}$ of $\mathcal{B}$.*

Proof. The left-to-right direction is trivial.

For the converse direction, we build an infinite finitely branching forest $\mathcal{T}$ as follows. Given a finite subset C of B and two maps $h, h'\colon C \to A$, define $h \sim h'$ if there exists $\alpha \in \mathrm{Aut}(\mathcal{A})$ such that $h = \alpha \circ h'$. Fix an enumeration $(b_i)_{i \in \mathbb{N}}$ of B. By assumption, for every i, there exists a homomorphism $\mathcal{B}[b_0, \ldots, b_i] \to \mathcal{A}$. Let T be the set of equivalence classes $[h]$ of homomorphisms $h\colon \mathcal{B}[b_0, \ldots, b_i] \to \mathcal{A}$ where $i \in \mathbb{N}$. Declare that the set $\{[h], [g]\}$ is an edge in $\mathcal{T}$ if $h\colon \mathcal{B}[b_0, \ldots, b_i] \to \mathcal{A}$, $g\colon \mathcal{B}[b_0, \ldots, b_{i+1}] \to \mathcal{A}$, and there exists $\alpha \in \mathrm{Aut}(\mathcal{A})$ such that $g|_{\{b_0, \ldots, b_i\}} = \alpha \circ h$. Note that this definition does not depend on the choice of the representatives of $[h]$ and $[g]$. Moreover, every class $[h]$ belongs to finitely many edges in $\mathcal{T}$, by Theorem 2.6 and the assumption that $\mathcal{A}$ is ω-categorical. It follows from König's tree lemma that $\mathcal{T}$ contains an infinite branch that we write $([h_i])_{i \in \mathbb{N}}$. Now, we define a map g from B to A such that for all $i \in \mathbb{N}$, we have $g|_{\{b_0, \ldots, b_i\}} \sim h_i$. We

proceed by induction on i as follows. For $i = 0$, define $g(b_0) := h_0(b_0)$. It is clear that $g|_{\{b_0\}} \sim h_0$. Suppose now that $g(b_j)$ has been defined for all $j \leq i$ and that $g|_{\{b_1,\ldots,b_i\}} \sim h_i$. Since $[h_i]$ is adjacent to $[h_{i+1}]$, there exists $\alpha \in \mathrm{Aut}(\mathcal{A})$ such that $h_{i+1}|_{\{b_0,\ldots,b_i\}} = \alpha \circ h_i$. Moreover, since h_i is equivalent to g, there exists $\beta \in \mathrm{Aut}(\mathcal{A})$ such that $h_i = \beta \circ g|_{\{b_0,\ldots,b_i\}}$. We obtain that $(\alpha\beta)^{-1} \circ h_{i+1}|_{\{b_0,\ldots,b_i\}} = g|_{\{b_0,\ldots,b_i\}}$. Define $g(b_{i+1}) := (\alpha\beta)^{-1} \circ h_{i+1}(b_{i+1})$. By definition, we have that $g|_{\{b_0,\ldots,b_{i+1}\}} \sim h_{i+1}$.

It remains to argue that the map defined above is a homomorphism $\mathcal{B} \to \mathcal{A}$. Let R be a relation symbol of arity k from the signature of $\mathcal{A}$ and let $(c_1, \ldots, c_k) \in R^{\mathcal{B}}$. There exists an $i \in \mathbb{N}$ such that $c_1, \ldots, c_k \in \{b_0, \ldots, b_i\}$. The map h_i is a homomorphism, so that $(h_i(c_1), \ldots, h_i(c_k)) \in R^{\mathcal{A}}$. Since $g|_{\{b_0,\ldots,b_i\}} \sim h_i$, there exists $\alpha \in \mathrm{Aut}(\mathcal{A})$ such that $g|_{\{b_0,\ldots,b_i\}} = \alpha \circ h_i$. Since α is an automorphism of $\mathcal{A}$, we obtain that $(g(c_1), \ldots, g(c_k)) \in R^{\mathcal{A}}$. □

The structure of the previous proof will appear numerous times in this dissertation. For this to work, the necessary properties of the object that we want to construct are:

- locality (e.g., being a homomorphism is a local property),
- invariance under automorphisms (e.g., if $h\colon \mathcal{B} \to \mathcal{A}$ is a homomorphism and $\alpha \in \mathrm{Aut}(\mathcal{A})$, then $\alpha \circ h\colon \mathcal{B} \to \mathcal{A}$ is a homomorphism).

Since the condition of being an embedding is also local and invariant under automorphisms, one gets that Proposition 2.7 also holds with homomorphisms being replaced by embeddings.

Orbits are related to the model-theoretic notion of *types*. The *first-order type* of a tuple $(a_1, \ldots, a_n) \in A^n$ in $\mathcal{A}$ is the set $\mathrm{tp}^{\mathcal{A}}(\overline{a}) := \{\phi(x_1, \ldots, x_n) \mid \mathcal{A} \models \phi(a_1, \ldots, a_n)\}$ of all first-order formulas that are satisfied by $(a_1, \ldots, a_n)$. Note that the type of a tuple is necessarily an infinite set of formulas, as it contains every tautological formula such as $x_1 = x_1, x_1 = x_1 \vee x_1 = x_1, \ldots$ as well as every first-order sentence that is true in $\mathcal{A}$. Moreover, if two tuples are in the same orbit under $\mathrm{Aut}(\mathcal{A})$, then they have the same type in $\mathcal{A}$ (this follows from the fact that automorphisms preserve first-order formulas, as we will see in Section 2.3.3). The converse is true in some structures and in particular it is true in ω-categorical structures, as we see below.

We say that $\mathrm{tp}^{\mathcal{A}}(\overline{a})$ is *isolated* by a formula ϕ if $\mathcal{A} \models \phi(\overline{a})$ holds and every $\overline{b} \in A^n$ such that $\mathcal{A} \models \phi(\overline{b})$ satisfies $\mathrm{tp}^{\mathcal{A}}(\overline{a}) = \mathrm{tp}^{\mathcal{A}}(\overline{b})$.

Theorem 2.8 (Theorem 6.3.1 and Corollary 6.3.3 in [56])**.** *Let $\mathcal{A}$ be an ω-categorical structure. Let $\overline{a}, \overline{b} \in A^n$. Then $\overline{a}$ and $\overline{b}$ are in the same orbit under* $\mathrm{Aut}(\mathcal{A})$ *if, and only if, they have the same type in $\mathcal{A}$. Moreover, every type is isolated by a first-order formula.*

2.3.2 Homogeneity

In Example 5, we saw that not only the orbits of pairs in $\mathcal{Q}$ are definable (which also follows from Theorem 2.8), but they are even definable by *quantifier-free* first-order formulas. This comes from the fact that for every first-order formula $\phi(x_1, \ldots, x_n)$, there exists a quantifier-free formula $\psi(x_1, \ldots, x_n)$ such that $\mathcal{Q} \models \forall x_1, \ldots, x_n(\phi(x_1, \ldots, x_n) \Leftrightarrow \psi(x_1, \ldots, x_n))$. We say that $\mathcal{Q}$ admits *quantifier elimination*. We describe in this section an important class of ω-categorical structures that admit quantifier elimination, and a generic way of constructing structures in this class.

Definition 2.4. Let $\mathcal{A}$ be a structure in a relational language. We say that $\mathcal{A}$ is *homogeneous* if for all finite isomorphic substructures $\mathcal{B}, \mathcal{C}$ of $\mathcal{A}$ (i.e., substructures for which there exists $f\colon \mathcal{B} \xrightarrow{\sim} \mathcal{C}$), there exists an automorphism $\alpha \in \mathrm{Aut}(\mathcal{A})$ that extends f.

Proposition 2.9. *Let $\mathcal{A}$ be a relational structure. Then $\mathcal{A}$ is ω-categorical and admits quantifier elimination if, and only if, $\mathcal{A}$ is homogeneous.*

The *age* of a structure $\mathcal{A}$ is the class $\mathrm{Age}(\mathcal{A})$ of all finite structures $\mathcal{B}$ that embed into $\mathcal{A}$, i.e., for which there exists an embedding $\mathcal{B} \hookrightarrow \mathcal{A}$. If $\mathcal{A}$ is homogeneous, one sees that $\mathrm{Age}(\mathcal{A})$ is closed under isomorphisms and has the following properties:

Hereditary Property For every $\mathcal{B}$ in $\mathrm{Age}(\mathcal{A})$ and every substructure $\mathcal{C}$ of $\mathcal{B}$, one has $\mathcal{C} \in \mathrm{Age}(\mathcal{A})$.

Amalgamation Property For all $\mathcal{B}_1, \mathcal{B}_2 \in \mathrm{Age}(\mathcal{A})$ and all $e_i\colon \mathcal{C} \hookrightarrow \mathcal{B}_i$ ($i \in \{1,2\}$) embeddings, there exist a structure $\mathcal{D} \in \mathrm{Age}(\mathcal{A})$ and embeddings $f_i\colon \mathcal{B}_i \hookrightarrow \mathcal{D}$ such that $f_1 \circ e_1 = f_2 \circ e_2$.

A class $\mathfrak{C}$ of finite structures is called an *amalgamation class* if it is closed under isomorphisms and satisfies the two properties above.

Theorem 2.10 (Fraïssé [53], see also Theorem 6.1.2 in [56])**.** *Let $\mathfrak{C}$ be a class of finite structures. Then $\mathfrak{C}$ is an amalgamation class if, and only if, there exists a countable homogeneous structure $\mathcal{A}$ such that $\mathfrak{C} = \mathrm{Age}(\mathcal{A})$. Moreover, $\mathcal{A}$ is unique up to isomorphism among the class of countable homogeneous structures.*

The structure $\mathcal{A}$ in the statement above is called the *Fraïssé limit of* $\mathfrak{C}$.

Example 7. Let $\mathfrak{LO}$ be the class of all finite linear orders. One sees that $\mathfrak{LO}$ is an amalgamation class, which means that $\mathfrak{LO}$ has a Fraïssé limit. It is clear that $\mathfrak{LO} = \mathrm{Age}(\mathcal{Q})$ and $\mathcal{Q}$ is a countable homogeneous structure, thus $\mathcal{Q}$ is the Fraïssé limit of $\mathfrak{LO}$. Note that $\mathrm{Age}(\mathbb{Z}; <)$ is also $\mathfrak{LO}$, and that $(\mathbb{Z}; <)$ and $(\mathbb{Q}; <)$ are *not* isomorphic.

2.3.3 Preservation

Let $f\colon A^n \to A$ be an operation on A and let $R \subseteq A^k$ be a relation. We say that f *preserves* R, or that R is invariant under f, if for every $\bar{t}^1, \ldots, \bar{t}^n \in R$ the k-tuple

$$f(\bar{t}^1, \ldots, \bar{t}^n) := (f(t_1^1, \ldots, t_1^n), \ldots, f(t_k^1, \ldots, t_k^n))$$

obtained by applying f to $\bar{t}^1, \ldots, \bar{t}^n$ componentwise is in R. Given a set $\mathscr{F}$ of functions on A, we write $\mathrm{Inv}(\mathscr{F})$ for the set of relations on A that are invariant under every function in $\mathscr{F}$.

Definition 2.5. Let $\mathcal{A}$ be a relational structure and let f be an operation on $\mathcal{A}$. We say that f is a *polymorphism* of $\mathcal{A}$ if f preserves every relation of $\mathcal{A}$. We write $\mathrm{Pol}(\mathcal{A})$ for the set of all polymorphisms of $\mathcal{A}$.

Note that the unary polymorphisms of a structure are exactly its endomorphisms.

Define the *existential positive* (ep) fragment of first-order logic to be the set of first-order formulas that are built using conjunctions, disjunctions, and existential quantifications. A relation is said to be *existentially positively definable* (ep-definable) in $\mathcal{A}$ if it is defined by an existential positive formula.

Lemma 2.11. *Let $\mathcal{A}$ be a relational structure and let $R \subseteq A^k$ be a relation on the domain of $\mathcal{A}$.*

1. *If R is first-order definable over $\mathcal{A}$, then R is invariant under all automorphisms of $\mathcal{A}$, i.e., $R \in \mathrm{Inv}(\mathrm{Aut}(\mathcal{A}))$;*

2. *If R is existentially positively definable over $\mathcal{A}$, then R is invariant under all endomorphisms of $\mathcal{A}$;*

3. *If R is primitively positively definable over $\mathcal{A}$, then R is invariant under all polymorphisms of $\mathcal{A}$.*

For an arbitrary structure $\mathcal{A}$, it is not true that if R is preserved by all automorphisms of $\mathcal{A}$ (resp. endomorphisms, polymorphisms), then R is fo-definable (resp. ep-, pp-) over $\mathcal{A}$. However, this implication holds for the class of ω-categorical structures. We write $\langle\mathcal{A}\rangle_{\mathrm{fo}}$ for the set of relations that are fo-definable over $\mathcal{A}$ and define similarly $\langle\mathcal{A}\rangle_{\mathrm{ep}}$ and $\langle\mathcal{A}\rangle_{\mathrm{pp}}$.

Theorem 2.12 (Bodirsky, Nešetřil [32])**.** *Let $\mathcal{A}$ be an ω-categorical structure and let $R \subseteq A^k$ be a relation on the domain of $\mathcal{A}$. Then R is first-order definable over $\mathcal{A}$ if, and only if, R is invariant under all automorphisms of $\mathcal{A}$, i.e., $\langle\mathcal{A}\rangle_{\mathrm{fo}} = \mathrm{Inv}(\mathrm{Aut}(\mathcal{A}))$. Similarly, the equalities $\langle\mathcal{A}\rangle_{\mathrm{ep}} = \mathrm{Inv}(\mathrm{End}(\mathcal{A}))$ and $\langle\mathcal{A}\rangle_{\mathrm{pp}} = \mathrm{Inv}(\mathrm{Pol}(\mathcal{A}))$ hold.*

A consequence of the third equality in Theorem 2.12 is that if $\mathrm{Pol}(\mathcal{A}) \subseteq \mathrm{Pol}(\mathcal{B})$ and $\mathcal{A}$ is ω-categorical, then all the relations of $\mathcal{B}$ are pp-definable over $\mathcal{A}$. The converse is true (even for arbitrary structures $\mathcal{A}$), thus giving an algebraic characterization of when Lemma 2.2 can be applied.

Corollary 2.13. *Let $\mathcal{A}$ and $\mathcal{B}$ be relational structures, and assume that $\mathcal{A}$ is ω-categorical. Then $\mathrm{Pol}(\mathcal{A}) \subseteq \mathrm{Pol}(\mathcal{B})$ if, and only if, all the relations of $\mathcal{B}$ are pp-definable over $\mathcal{A}$. In particular, if $\mathrm{Pol}(\mathcal{A}) \subseteq \mathrm{Pol}(\mathcal{B})$ then $\mathrm{CSP}(\mathcal{B})$ reduces to $\mathrm{CSP}(\mathcal{A})$.*

A similar characterisation of when Lemma 2.3 can be applied exists (under some conditions) and necessitates notions from universal algebra and topology.

2.4 Topology

Let B be a countable set, and let $k \geq 1$. The set $\mathcal{O}_B^{(k)}$ of all maps from B^k to B can be endowed with a natural metric, that is, a map $d\colon \mathcal{O}_B^{(k)} \times \mathcal{O}_B^{(k)} \to \mathbb{R}$ satisfying the following properties:

- $d(f, f) = 0$ for all $f \in \mathcal{O}_B^{(k)}$,

- $d(f, g) = d(g, f)$ for all $f, g \in \mathcal{O}_B^{(k)}$,

- $d(f, g) \leq d(f, h) + d(g, h)$ for all $f, g, h \in \mathcal{O}_B^{(k)}$.

This metric can be defined as follows. Fix an enumeration $(\bar{b}_n)_{n\in\mathbb{N}}$ of the k-tuples of B. Define $d(f,g)$ to be 0 if $f=g$, and otherwise to be $\frac{1}{2^n}$, where n is the smallest index such that $f(\bar{b}_n) \neq g(\bar{b}_n)$.

While the definition of the metric depends on the enumeration of B^k that we chose, the topology that is induced on $\mathcal{O}_B^{(k)}$ by any two such metrics is the same. We call this topology the *topology of pointwise convergence*. A basic open set in this topology is a set of the form $\mathcal{U}_{\bar{a},b} := \{f \in \mathcal{O}_B^{(k)} \mid f(\bar{a}) = b\}$, where $\bar{a} \in B^k$ and $b \in B$. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions in $\mathcal{O}_B^{(k)}$, and let $g \in \mathcal{O}_B^{(k)}$. We say that f_n *converges* to g (or that g is a limit of $(f_n)_{n\in\mathbb{N}}$) if for every finite set $S \subseteq B$, there exists $N \in \mathbb{N}$ such that for every $n \geq N$, we have $f_n|_S = g|_S$. Given a set $\mathscr{F}$ of functions in $\mathcal{O}_B^{(k)}$, we write $\overline{\mathscr{F}}$ for the set of limits of sequences from $\mathscr{F}$.

We endow $\mathcal{O}_B := \cup_{k\geq 1}\mathcal{O}_B^{(k)}$ with the disjoint union topology. Let $\mathscr{B} \subseteq \mathcal{O}_B$ and $\mathscr{C} \subseteq \mathcal{O}_C$.

If $\mathscr{B}$ and $\mathscr{C}$ are endowed with the corresponding subspace topology, an arity-preserving function $\xi\colon \mathscr{B} \to \mathscr{C}$ is *continuous* if for every finite subset $C' \subset C$ of C and every $f \in \mathscr{B}$, there exists a finite subset $B' \subset B$ of B such that for every $g \in \mathscr{B}$ with the same arity as f, if $f|_{B'} = g|_{B'}$ then $\xi(f)|_{C'} = \xi(g)|_{C'}$. We say that $\xi\colon \mathscr{B} \to \mathscr{C}$ is *uniformly continuous* if the set B' only depends on C' and not on f. Of particular interest for us will be the case when C is itself a finite set. Then, ξ is uniformly continuous if, and only if, for every $k \geq 1$ there exists $B' \subset B$ such that $f|_{B'} = g|_{B'}$ implies $\xi(f) = \xi(g)$ for all $f, g \in \mathscr{B}$ of arity k. Note that if B is finite, every arity-preserving map $\xi\colon \mathscr{B} \to \mathscr{C}$ is uniformly continuous.

2.5 Universal Algebra

2.5.1 Clones

So far, we have only seen $\mathrm{Pol}(\mathcal{A})$ as a set of functions. It is a simple observation that if $f \in \mathrm{Pol}(\mathcal{A})$ is an n-ary polymorphism of $\mathcal{A}$ and $g_1, \ldots, g_n$ are m-ary polymorphisms of $\mathcal{A}$, then

$$f \circ (g_1, \ldots, g_n)\colon (a_1, \ldots, a_m) \mapsto f(g_1(\bar{a}), \ldots, g_n(\bar{a}))$$

is an m-ary polymorphism of $\mathcal{A}$. Moreover, let $\pi_i^m\colon A^m \to A$ denote the ith m-ary projection on A. Then for all $m \geq 1$ and $i \in \{1, \ldots, m\}$, the function π_i^m is in $\mathrm{Pol}(\mathcal{A})$.

Definition 2.6. Let $\mathscr{A}$ be a set of operations on A. We say that $\mathscr{A}$ is a *clone* if it contains the projections on A and is closed under compositions.

Thus, $\mathrm{Pol}(\mathcal{A})$ is a clone for every relational structure $\mathcal{A}$. Moreover, the composition operation on $\mathrm{Pol}(\mathcal{A})$ is continuous. We say that $\mathrm{Pol}(\mathcal{A})$ is a *topological clone*. It is an easy exercise to show that $\mathrm{Pol}(\mathcal{A})$ is a *closed* subset of $\mathcal{O}_A$, that is, that the limit of a converging sequence of polymorphisms of $\mathcal{A}$ is again a polymorphism of $\mathcal{A}$. Given two operations f, g on A and a set of unary functions $\mathscr{U}$ on A, we say that f is *interpolated* by g modulo $\mathscr{U}$ if $f \in \overline{\{ug(v_1, \ldots, v_k) \mid u, v_1, \ldots, v_k \in \mathscr{U}\}}$. We also write $\mathscr{U}g\mathscr{U}$ for $\{ug(v_1, \ldots, v_k) \mid u, v_1, \ldots, v_k \in \mathscr{U}\}$. We say that f is *generated* by g modulo $\mathscr{U}$ if f is contained in the smallest closed clone containing g and $\mathscr{U}$. Since $\mathrm{Pol}(\mathcal{A})$ is a closed clone for every relational structure $\mathcal{A}$, if $g \in \mathrm{Pol}(\mathcal{A})$ and $\mathscr{U} \subseteq \mathrm{End}(\mathcal{A})$, we obtain that $f \in \mathrm{Pol}(\mathcal{A})$ if f is interpolated or generated by g modulo $\mathscr{U}$.

Definition 2.7. Let $\mathscr{A}$ and $\mathscr{B}$ be two clones on the sets A and B. A *clone homomorphism* is a map $\xi\colon \mathscr{A} \to \mathscr{B}$ such that:

- if $f \in \mathscr{A}$ has arity n, then $\xi(f)$ has arity n,
- for every $n \geq 1$ and $i \in \{1, \dots, n\}$, $\xi(\pi_i^n) = \pi_i^n$,
- for every n-ary $f \in \mathscr{A}$ and m-ary $g_1, \dots, g_n \in \mathscr{B}$, one has

$$\xi(f \circ (g_1, \dots, c_n)) = \xi(f) \circ (\xi(g_1), \dots, \xi(g_n)).$$

One sees from the definition that clone homomorphisms preserve identities. Identities will only be defined formally in the next subsection, so we rather give here an example. Consider the case that $\mathscr{A}$ contains a binary function f that is associative, that is, satisfies $f(a, f(b, c)) = f(f(a, b), c)$ for all $a, b, c \in A$. One can rewrite this as $f(\pi_1^3, f(\pi_2^3, \pi_3^3)) = f(f(\pi_1^3, \pi_2^3), \pi_3^3)$. If $\xi\colon \mathscr{A} \to \mathscr{B}$ is a clone homomorphism, then $\xi(f)(\pi_1^3, \xi(f)(\pi_2^3, \pi_3^3)) = \xi(f)(\xi(f)(\pi_1^3, \pi_2^3), \pi_3^3)$ holds, so that the function $\xi(f) \in \mathscr{B}$ is also associative. Barto, Opršal, and Pinsker [6] gave a weaker notion of homomorphism between clones and showed its tight connection with pp-constructions.

Definition 2.8. Let $\mathscr{A}$ and $\mathscr{B}$ be two clones on the sets A and B. A *clonoid homomorphism* is a map $\xi\colon \mathscr{A} \to \mathscr{B}$ such that:

- if $f \in \mathscr{A}$ has arity n, then $\xi(f)$ has arity n,
- for every n-ary $f \in \mathscr{A}$ and m-ary *projections* $g_1, \dots, g_n$, one has

$$\xi(f \circ (g_1, \dots, g_n)) = \xi(f) \circ (g_1, \dots, g_n).$$

Note that a clonoid homomorphism need not map π_i^n to π_i^n.

Theorem 2.14 ([6])**.** *Let $\mathcal{A}$ be an ω-categorical structure and let $\mathcal{B}$ be finite. Let $\mathscr{A} = \mathrm{Pol}(\mathcal{A})$ and $\mathscr{B} = \mathrm{Pol}(\mathcal{B})$. The following are equivalent:*

1. *$\mathcal{B}$ has a pp-construction over $\mathcal{A}$,*
2. *there exists a uniformly continuous clonoid homomorphism $\mathscr{A} \xrightarrow{u.c.c.h.} \mathscr{B}$.*

The implication (1.) $\Rightarrow$ (2.) *holds for arbitrary structures $\mathcal{A}$ and $\mathcal{B}$.*

This gives the desired algebraic characterisation of when Lemma 2.3 can be applied, at least when $\mathcal{B}$ is finite. In particular, we obtain the following pendant to Corollary 2.4. Let $\mathscr{P} = \mathrm{Pol}(\text{1-IN-3-SAT})$.

Corollary 2.15. *Let $\mathcal{A}$ be an ω-categorical structure. If there is a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{A}) \xrightarrow{u.c.c.h.} \mathscr{P}$, then $\mathrm{CSP}(\mathcal{A})$ is NP-hard.*

2.5.2 Identities

Another way of studying the question of when $\mathcal{B}$ is pp-constructible in $\mathcal{A}$ is to study the *identities* that are true in $\mathrm{Pol}(\mathcal{A})$ and $\mathrm{Pol}(\mathcal{B})$. Let X be a set and let σ be a functional signature. The set of *σ-terms* over X is the smallest set of expressions $t(x_1, \ldots, x_n)$ (with $x_1, \ldots, x_n \in X$) such that:

- $x_i(x_1, \ldots, x_n)$ is a term (often denoted by x_i when $x_1, \ldots, x_n$ are clear from the context),
- if $t_1(x_1, \ldots, x_n), \ldots, t_k(x_1, \ldots, x_n)$ are terms and $f \in \sigma$ is a symbol of arity k, then $(f(t_1, \ldots, t_k))(x_1, \ldots, x_n)$ is a term.

We say that a term has *height 1* if it is of the form $f(x_1, \ldots, x_n)$ for $f \in \sigma$.

A *σ-identity* is formally a pair $(s(x_1, \ldots, x_n), t(x_1, \ldots, x_n))$ of σ-terms. We write identities as equations $s(x_1, \ldots, x_n) \approx t(x_1, \ldots, x_n)$. We say that a set Σ of σ-identities is satisfiable in a clone $\mathscr{A}$ if there exists a map $\xi \colon \sigma \to \mathscr{A}$ such that for all $(s, t) \in \Sigma$ the statement

$$\forall a_1, \ldots, a_n \in A,\ \xi(s)(a_1, \ldots, a_n) = \xi(t)(a_1, \ldots, a_n)$$

holds (where ξ is extended from σ to the set of σ-terms in the natural way). An identity $s \approx t$ is said to have height 1 if both s and t have height 1.

The following proposition follows directly from the definitions.

Proposition 2.16. *Let $\mathscr{A}, \mathscr{B}$ be two clones. The following are equivalent:*

1. *every set of height 1 identities that is satisfiable in $\mathscr{A}$ is satisfiable in $\mathscr{B}$,*
2. *there exists a (not necessarily uniformly continuous) clonoid homomorphism $\mathscr{A} \to \mathscr{B}$.*

The previous statement has a very important corollary. If $\mathcal{A}$ and $\mathcal{B}$ are finite structures such that $\mathrm{Pol}(\mathcal{A})$ and $\mathrm{Pol}(\mathcal{B})$ satisfy the same height 1 identities, then there are clonoid homomorphisms $\mathrm{Pol}(\mathcal{A}) \to \mathrm{Pol}(\mathcal{B})$ and $\mathrm{Pol}(\mathcal{B}) \to \mathrm{Pol}(\mathcal{A})$. Since the structures are finite these homomorphisms are automatically uniformly continuous, so that by Theorem 2.14 and Lemma 2.3 we obtain that $\mathrm{CSP}(\mathcal{A})$ and $\mathrm{CSP}(\mathcal{B})$ have the same complexity up to logspace reductions. In other words, the complexity of finite-domain CSPs is completely encoded into height 1 identities.

Consider again the case of 1-in-3-Sat and its clone of polymorphisms. It can be seen that $\mathscr{P}$ consists only of projection operations. Therefore, $\mathscr{P} \subseteq \mathscr{A}$ for every clone $\mathscr{A}$, so that the only identities that are satisfiable in $\mathscr{P}$ are the identities that are satisfiable in every clone $\mathscr{A}$. We call these identities *trivial*. We give now some examples of important non-trivial identities. A *cyclic* operation is a function $f \colon A^n \to A$ (for $n \geq 2$) such that $\forall a_1, \ldots, a_n \in A,\ f(a_1, \ldots, a_n) = f(a_2, \ldots, a_n, a_1)$. A *weak near-unanimity* operation is a function $f \colon A^n \to A$ (for $n \geq 3$) such that for all $a, b \in A,\ f(a, b, \ldots, b) = f(b, a, b, \ldots, b) = \cdots = f(b, \ldots, b, a)$. A *Siggers* operation is a function $f \colon A^6 \to A$ such that for all $a, b, c \in A,\ f(a, b, a, c, b, c) = f(b, a, c, a, c, b)$. It is easy to see that no projection can be cyclic, a weak near-unanimity, or a Siggers operation, so that these are indeed examples of non-trivial identities. These identities are in fact "minimally non-trivial" for clones of operations on a finite set, in the sense of the following theorem.

Theorem 2.17 ([4, 6, 54, 78])**.** *Let $\mathscr{A}$ be a clone on a finite set. The following are equivalent:*

1. *There is no clonoid homomorphism $\mathscr{A} \xrightarrow{c.h.} \mathscr{P}$,*
2. *$\mathscr{A}$ satisfies a non-trivial identity,*
3. *$\mathscr{A}$ contains a weak near-unanimity operation,*
4. *$\mathscr{A}$ contains a cyclic operation,*
5. *$\mathscr{A}$ contains a Siggers operation.*

Note that by the finite-domain dichotomy theorem (Theorem 2.5), if $\mathcal{A}$ is a finite structure such that $\mathrm{Pol}(\mathcal{A})$ contains one of the operations in the previous statement, then $\mathrm{CSP}(\mathcal{A})$ is in P!

The role of identities for infinite-domain constraint satisfaction is less clear than in the finite case for two reasons. First, Theorem 2.14 does not hold in general if $\mathcal{A}$ is not ω-categorical; thus, it is a possibility that $\mathrm{Pol}(\mathcal{A})$ does not satisfy any nontrivial identities while $\mathrm{CSP}(\mathcal{A})$ is in P. Conversely, it is possible to construct structures with a binary cyclic polymorphism and an undecidable CSP. Secondly, even for an ω-categorical structure $\mathcal{A}$, Theorem 2.14 is about algebraic and topological properties of $\mathrm{Pol}(\mathcal{A})$, while Proposition 2.16 is purely algebraic. Therefore, knowing that $\mathrm{Pol}(\mathcal{A})$ does not satisfy any nontrivial height 1 identities gives a clonoid homomorphism $\mathrm{Pol}(\mathcal{A}) \to \mathscr{P}$ that is not necessarily uniformly continuous and is *a priori* not enough to prove that $\mathrm{CSP}(\mathcal{A})$ is NP-hard.

2.6 Model-completeness, Cores

We have seen with Example 2 that several templates can have the same CSP. For ω-categorical structures, there is however a canonical template.

A finite or countably infinite ω-categorical structure $\mathcal{A}$ is called a *core* if all endomorphisms of $\mathcal{A}$ are embeddings, and it is called *model-complete* if all embeddings of $\mathcal{A}$ into $\mathcal{A}$ preserve all first-order formulas.

Theorem 2.18 (Theorem 16 in [10])**.** *Every ω-categorical structure $\mathcal{A}$ is homomorphically equivalent to an ω-categorical model-complete core, which is up to isomorphism unique and embeds into $\mathcal{A}$.*

Note that since $\mathcal{A}$ and its model-complete core are homomorphically equivalent, they have pp-constructions in one another and their CSP has the same complexity.

Proposition 2.19 (Theorem 18 in [10])**.** *For a countable ω-categorical structure $\mathcal{A}$, the following are equivalent.*

- *$\mathcal{A}$ is a model-complete core;*
- *the orbits of tuples under $\mathrm{Aut}(\mathcal{A})$ are pp-definable in $\mathcal{A}$;*
- $\mathrm{End}(\mathcal{A}) = \overline{\mathrm{Aut}(\mathcal{A})}$*, that is, every endomorphism of $\mathcal{A}$ is a limit of automorphisms.*

$$\begin{tikzcd}
\text{every height 1 identity in } \mathrm{Pol}(\mathcal{A}) \text{ is satisfiable in } \mathrm{Pol}(\mathcal{B}) \arrow[r, leftrightarrow, "\text{Prop. 2.16}"] & \mathrm{Pol}(\mathcal{A}) \arrow[r, "c.h."] & \mathrm{Pol}(\mathcal{B}) \\
\mathcal{B} \text{ has a pp-construction in } \mathcal{A} \arrow[u] \arrow[r, leftrightarrow, "\text{Theorem 2.14}"'] & \mathrm{Pol}(\mathcal{A}) \arrow[u] \arrow[r, "u.c.c.h."] & \mathrm{Pol}(\mathcal{B}) \arrow[u]
\end{tikzcd}$$

Figure 2.1: $\mathcal{A}$ is an ω-categorical structure and $\mathcal{B}$ is a finite structure. The conditions in the bottom row imply that $\mathrm{CSP}(\mathcal{B})$ reduces to $\mathrm{CSP}(\mathcal{A})$.

By the previous proposition, a model-complete core $\mathcal{A}$ has the property that expanding $\mathcal{A}$ by a relation symbol for every orbit of tuples under $\mathrm{Aut}(\mathcal{A})$ does not change the complexity of the CSP, since these relations are pp-definable in $\mathcal{A}$. It is also possible to name individual elements of $\mathcal{A}$ by unary singleton relations without changing the complexity of the CSP:

Proposition 2.20 (Bodirsky [10], Barto, Opršal, Pinsker [6])**.** *Let $\mathcal{A}$ be an ω-categorical model-complete core, and let $a \in A$. The structure $(\mathcal{A}, a)$ has a pp-construction in $\mathcal{A}$, and in particular* $\mathrm{CSP}(\mathcal{A}, \{a\})$ *and* $\mathrm{CSP}(\mathcal{A})$ *have the same complexity.*

Note that if $\mathcal{A}$ is finite and is a core (model-completeness is a vacuous condition for finite structures), then one can add a unary relation for each element of the domain of $\mathcal{A}$. Let $\mathcal{B}$ be the resulting structure, that is, $\mathcal{B}$ is $(\mathcal{A}, a_1, \ldots, a_n)$ where $A = \{a_1, \ldots, a_n\}$. The polymorphisms of $\mathcal{B}$ are *idempotent*, that is, every polymorphism $f \in \mathrm{Pol}(\mathcal{B})$ satisfies $f(a, \ldots, a) = a$ for all $a \in A$. For finite idempotent clones, another equivalent condition can be added to Theorem 2.17. We will give this equivalent condition in Chapter 4, and it will be at the heart of our lifting techniques in Chapters 4 and 5.

For ω-categorical model-complete cores, Barto and Pinsker [7] recently proved a statement analogous to Theorem 2.17. Let $\mathscr{U}$ be a set of unary functions on A. A *pseudo-Siggers* operation modulo $\mathscr{U}$ is a function $f\colon A^6 \to A$ such that there exists $e_1, e_2 \in \mathscr{U}$ such that $e_1 f(a, b, a, c, b, c) = e_2 f(b, a, c, a, c, b)$ holds for all $a, b, c \in A$.

Theorem 2.21 ([2, 7])**.** *Let $\mathcal{A}$ be an ω-categorical structure and let $\mathcal{B}$ be its model-complete core. The following are equivalent:*

1. *there is no uniformly continuous clonoid homomorphism* $\mathrm{Pol}(\mathcal{A}) \xrightarrow{u.c.c.h.} \mathscr{P}$,

2. $\mathrm{Pol}(\mathcal{B})$ *contains a pseudo-Siggers operation modulo* $\mathrm{End}(\mathcal{B})$.

Note that in the previous statement, the first item is a statement about the algebraic and topological structure of $\mathrm{Pol}(\mathcal{A})$ and $\mathrm{Pol}(\mathcal{B})$. The second item, however, is a purely algebraic statement about $\mathrm{Pol}(\mathcal{B})$. We give a summary of the main statements in this section in Figure 2.1 and Figure 2.2.

$$\begin{tikzcd}
\begin{array}{c}\mathrm{Pol}(\mathcal{A}) \text{ does not satisfy any} \\ \text{non-trivial height 1 identity}\end{array} & \mathrm{Pol}(\mathcal{A}) \arrow[l, leftrightarrow] \arrow[r, "c.h."] & \mathscr{P} \\
\text{1-IN-3-SAT has a pp-construction in } \mathcal{A} \arrow[u] & \mathrm{Pol}(\mathcal{A}) \arrow[l, leftrightarrow] \arrow[u] \arrow[r, "u.c.c.h."] & \mathscr{P}
\end{tikzcd}$$

Figure 2.2: $\mathcal{A}$ is an ω-categorical structure. The conditions in the bottom row imply that $\mathrm{CSP}(\mathcal{A})$ is NP-hard.

2.7 The Infinite-Domain Tractability Conjecture

Finally, we present here the infinite-domain tractability conjecture by Bodirsky and Pinsker, or rather an equivalent form thereof. The algebraic and topological notions involved in the statement have been presented above; it remains now to define the scope of the conjecture. We have seen that an algebraic and topological approach to complexity is available for CSPs of ω-categorical structures, and it would therefore be natural to assume that the complexity of said CSPs is well-behaved. However, Bodirsky and Grohe [16] proved that there are ω-categorical (and even homogeneous) structures whose CSP is in coNP but is neither in P nor coNP-complete (if $\mathsf{P} \neq \mathsf{NP}$). Their constructions rely on structures that are not finitely bounded, in the following sense.

For a set $\mathfrak{F}$ of τ-structures, let $\mathrm{Forb}^{\mathrm{ind}}(\mathfrak{F})$ be the set of finite τ-structures $\mathcal{A}$ such that for every $\mathcal{F} \in \mathfrak{F}$, there is no embedding of $\mathcal{F}$ into $\mathcal{A}$. A relational structure $\mathcal{B}$ is called *finitely bounded* if there exists a finite set of finite structures $\mathfrak{F}$ (the *bounds*) such that $\mathrm{Age}(\mathcal{B}) = \mathrm{Forb}^{\mathrm{ind}}(\mathfrak{F})$.

Example 8. Consider the Henson graph $\mathcal{H}_3$, which is uniquely defined up to isomorphism by the properties that it is homogeneous and is such that a finite graph $\mathcal{G}$ is an induced subgraph of $\mathcal{H}_3$ if, and only if, $\mathcal{G}$ does not contain any triangle. Let $\mathfrak{F} = \{\mathcal{K}_3\}$. By definition, we have $\mathrm{Age}(\mathcal{H}_3) = \mathrm{Forb}^{\mathrm{ind}}(\mathfrak{F})$, so that $\mathcal{H}_3$ is an example of a finitely bounded structure.

Let $\mathcal{A}$ and $\mathcal{B}$ be two relational structures. We say that $\mathcal{A}$ is a *first-order reduct* of $\mathcal{B}$ if $\mathcal{A}$ and $\mathcal{B}$ have the same domain and every relation of $\mathcal{A}$ is definable in $\mathcal{B}$ by a first-order formula. Note that Theorem 2.6 and Lemma 2.11 imply that if $\mathcal{B}$ is ω-categorical then $\mathcal{A}$ is as well.

Conjecture 1 (Infinite-domain tractability conjecture; see e.g. [2, 6, 37])**.** *Let $\mathcal{A}$ be a first-order reduct of a finitely bounded homogeneous structure. If there is no uniformly continuous clonoid homomorphism* $\mathrm{Pol}(\mathcal{A}) \xrightarrow{u.c.c.h.} \mathscr{P}$*, then* $\mathrm{CSP}(\mathcal{A})$ *is solvable in polynomial time.*

To this date, Conjecture 1 has been confirmed in a number of special cases. In particular, the conjecture holds in the case that $\mathcal{A}$ is a first-order reduct of:

- $(A, =)$, the naked structure;

- $(\mathbb{Q}, <)$;
- any countable homogeneous undirected graph.

In the next chapters, we prove the conjecture in the following cases:

- For every finitely bounded homogeneous structure $\mathcal{B}$, we prove that there exists a relation $R_{\mathcal{B}}$ such that for every first-order reduct $\mathcal{A}$ of $\mathcal{B}$, the structure $(\mathcal{A}, R_{\mathcal{B}})$ satisfies Conjecture 1.
- We prove that every first-order reduct of a unary structure satisfies the dichotomy.
- Finally, we prove that the class of CSPs expressible in the logic MMSNP satisfies the dichotomy.

Part I

Lifting Techniques

Chapter 3

Equations and Tractability Conditions

Finite-domain CSPs are much better understood than their infinite counterpart, mainly because the theory of finite algebras is much better understood than the theory of infinite algebras. Thus, a possible approach for the study of infinite-domain CSPs is to reduce them to finite-domain CSPs in a controlled way. In this chapter, we prove that this approach is indeed possible and we present a polynomial-time reduction from infinite-domain constraint satisfaction problems to finite-domain constraint satisfaction problems. This reduction applies to the CSP of every structure $\mathcal{A}$ that is a quantifier-free reduct of a finitely bounded structure. Moreover, the reduction preserves the algebraic properties of $\mathcal{A}$, and we leverage this to obtain new abstract algebraic tractability conditions. This chapter contains published results from [30].

3.1 The Type Structure

Write $[n]$ for the set $\{1, \dots, n\}$. The quantifier-free (qf-) type of a tuple $(b_1, \dots, b_m)$ in $\mathcal{B}$ is the set of all quantifier-free formulas $\phi(z_1, \dots, z_m)$ such that $\mathcal{B} \models \phi(b_1, \dots, b_m)$. If $\mathcal{B}$ has a finite relational signature then there are only finitely many quantifier-free m-types in $\mathcal{B}$.

Let m be a positive integer and let $\mathcal{A}$ be a structure whose relations are definable in $\mathcal{B}$ by a quantifier-free formula (i.e., $\mathcal{A}$ is a quantifier-free reduct of $\mathcal{B}$). We define $T_{\mathcal{B},m}(\mathcal{A})$ to be the relational structure whose domain is the set of quantifier-free m-types of $\mathcal{B}$ and whose relations are as follows.

- For each symbol R of $\mathcal{A}$ of arity r, let $\chi(z_1, \dots, z_r)$ be a definition of R in $\mathcal{B}$. For $i\colon [r] \to [m]$ we write $\langle \chi(z_{i(1)}, \dots, z_{i(r)}) \rangle$ for the unary relation that consists of all the types that contain $\chi(z_{i(1)}, \dots, z_{i(r)})$, and add all such relations to $T_{\mathcal{B},m}(\mathcal{A})$[1].

- For each $r \in [m]$ and $i, j\colon [r] \to [m]$, define $\mathrm{Comp}_{i,j}$ to be the binary relation that contains the pairs (p, q) of m-types such that for every quantifier-free formula $\chi(z_1, \dots, z_s)$ of $\mathcal{B}$ and $t\colon [s] \to [r]$, the formula $\chi(z_{it(1)}, \dots, z_{it(s)})$ is in p iff $\chi(z_{jt(1)}, \dots, z_{jt(s)})$ is in q.

[1] In the following, we use functions to index tuples. This notation allows us to avoid double-subscripting and to conveniently talk about subtuples.

Note that if $(a_1, \ldots, a_m)$ is of type p and $(b_1, \ldots, b_m)$ of type q, then $\mathrm{Comp}_{i,j}(p, q)$ holds if and only if $(a_{i(1)}, \ldots, a_{i(r)})$ and $(b_{j(1)}, \ldots, b_{j(r)})$ have the same type in $\mathcal{B}$. Also note that if $i\colon [m] \to [m]$ is the identity map, then $\mathrm{Comp}_{i,i}$ denotes the equality relation on the domain of $T_{\mathcal{B},m}(\mathcal{A})$.

The next theorem holds for arbitrary finitely bounded structures $\mathcal{B}$.

Theorem 3.1. *Let $\mathcal{A}$ be a quantifier-free reduct of a finitely bounded structure $\mathcal{B}$, and suppose that $\mathcal{A}$ has a finite signature. Let m_a be the maximal arity of a relation in $\mathcal{A}$ or $\mathcal{B}$, and m_b be the maximal size of a bound for $\mathcal{B}$. Let m be at least* $\max(m_a + 1, m_b, 3)$. *Then there is a polynomial-time reduction from* $\mathrm{CSP}(\mathcal{A})$ *to* $\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$.

We give in the next section a sufficient condition for the existence of a polynomial-time reduction in the other direction, from $\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$ to $\mathrm{CSP}(\mathcal{A})$.

Proof of Theorem 3.1. Let Ψ be an instance of $\mathrm{CSP}(\mathcal{A})$, and let $V = \{x_1, \ldots, x_n\}$ be the variables of Ψ. Assume without loss of generality that $n \geq m$. We build an instance Φ of $\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$ as follows.

- The variable set of Φ is the set I of increasing functions[2] from $[m]$ to V (where the variables are endowed with an arbitrary linear order). The idea of the reduction is that the variable $v \in I$ of Φ represents the qf-type of $(h(v(1)), \ldots, h(v(m)))$ in a satisfying assignment h for Ψ.
- For each conjunct ψ of Ψ we add unary constraints to Φ as follows. The formula ψ must be of the form $R(j(1), \ldots, j(r))$ where R is a relation of $\mathcal{A}$ and $j\colon [r] \to V$. By assumption, R has a qf-definition $\chi(z_1, \ldots, z_r)$ over $\mathcal{B}$. Let $v \in I$ be such that $\mathrm{im}(j) \subseteq \mathrm{im}(v)$. Let U be the relation symbol of $T_{\mathcal{B},m}(\mathcal{A})$ that denotes the unary relation $\langle \chi(z_{v^{-1}j(1)}, \ldots, z_{v^{-1}j(r)}) \rangle$. We then add $U(v)$ to Φ.
- Finally, for all $u, v \in I$ let $k\colon [r] \to \mathrm{im}(u) \cap \mathrm{im}(v)$ be a bijection. We then add the constraint $\mathrm{Comp}_{u^{-1}k, v^{-1}k}(u, v)$.

Before proving that the given reduction indeed works, we give an illustrating example.

Example 9. Let $\mathcal{A}$ be $(\mathbb{N}; =, \neq)$. We illustrate the reduction with the concrete instance

$$x_1 = x_2 \wedge x_2 = x_3 \wedge x_3 = x_4 \wedge x_1 \neq x_4 \,.$$

of $\mathrm{CSP}(\mathcal{A})$. The structure $(\mathbb{N}; =, \neq)$ is a reduct of the homogeneous structure with domain $\mathbb{N}$ and the empty signature, which has no bounds. We have in this example $m = 3$.

The structure $T_{\mathcal{B},3}(\mathcal{A})$ has a domain of size five, where each element corresponds to a partition of $\{z_1, z_2, z_3\}$. The structure has a unary relation U_1 for $\langle z_2 = z_3 \rangle$, containing all partitions in which z_2 and z_3 belong to the same part. Similarly, the structure has a relation U_2 for $\langle z_1 = z_3 \rangle$, U_3 for $\langle z_1 = z_2 \rangle$, V_1 for $\langle z_2 \neq z_3 \rangle$, V_2 for $\langle z_1 \neq z_3 \rangle$, and V_3 for $\langle z_1 \neq z_2 \rangle$. The instance Φ of $\mathrm{CSP}(T_{\mathcal{B},3}(\mathcal{A}))$ that our reduction creates has four variables, for the four order-preserving injections from $[3] \to \{x_1, x_2, x_3, x_4\}$ (where we order $x_1, \ldots, x_4$ according to their index). Call v_1, v_2, v_3, v_4 these variables, where $\mathrm{im}(v_j) = \{x_1, \ldots, x_4\} \setminus \{x_j\}$. We then have the following constraints in Φ:

[2] One could take I to be the set of all functions $[m] \to V$ without any change to the reduction. We choose here to only take increasing functions so that the presentation of the example below is more concise.

- $U_3(v_3)$ and $U_3(v_4)$ for the constraint $x_1 = x_2$ in Ψ;
- $U_1(v_4)$ and $U_3(v_1)$ for the constraint $x_2 = x_3$ in Ψ;
- $U_1(v_2)$ and $U_1(v_1)$ for the constraint $x_3 = x_4$ in Ψ;
- $V_2(v_2)$ and $V_2(v_3)$ for the constraint $x_1 \neq x_4$ in Ψ.

For the compatibility constraints we only give an example. Let $k, k' \colon [2] \to [4]$ be such that $k(1,2) = (1,3)$ and $k'(1,2) = (1,2)$. Then $\mathrm{Comp}_{k,k'}(v_4, v_2)$ and $\mathrm{Comp}_{k',k'}(v_4, v_3)$ are in Φ. $\square$

We now prove that the reduction is correct. Let $h \colon V \to B$ be an assignment of the variables to the domain of $\mathcal{B}$. Let $\chi(z_1, \ldots, z_r)$ be a qf-formula in the language of $\mathcal{B}$, let $j \colon [r] \to V$, and let v in I be such that $\mathrm{im}(j) \subseteq \mathrm{im}(v)$. We first note the following property:

$$\begin{aligned} &\mathcal{B} \models \chi(h(j(1)), \ldots, h(j(r))) \\ &\text{iff } (h(v(1)), \ldots, h(v(m))) \text{ satisfies } \chi(z_{v^{-1}j(1)}, \ldots, z_{v^{-1}j(r)}) \text{ in } \mathcal{B}. \end{aligned} \tag{‡}$$

The property (‡) holds since in the type of the tuple $(h(v(1)), \ldots, h(v(m)))$, the variable z_i represents the element $h(v(i))$, and therefore $z_{v^{-1}j(i)}$ represents $h(j(i))$.

(Ψ satisfiable implies Φ satisfiable.) Suppose that $h \colon V \to B$ satisfies Ψ in $\mathcal{A}$. To show that Φ is satisfiable in $T_{\mathcal{B},m}(\mathcal{A})$ define $g \colon I \to T_{\mathcal{B},m}(\mathcal{A})$ by setting $g(v)$ to be the type of $(h(v(1)), \ldots, h(v(m)))$ in $\mathcal{B}$, for every $v \in I$. To see that all the constraints of Φ are satisfied by g, let $U(v)$ be a constraint in Φ that has been introduced for a conjunct of the form $R(j(1), \ldots, j(r))$ in Ψ, where $j \colon [r] \to V$. Let $\chi(z_1, \ldots, z_r)$ be a qf-formula that defines R in $\mathcal{B}$. Then

$$\begin{aligned} & \mathcal{A} \models R(h(j(1)), \ldots, h(j(r))) \\ \Rightarrow\ & \mathcal{B} \models \chi(h(j(1)), \ldots, h(j(r))) \\ \Rightarrow\ & \chi(z_{v^{-1}j(1)}, \ldots, z_{v^{-1}j(m)}) \in g(v) && \text{(because of (‡))} \\ \Rightarrow\ & T_{\mathcal{B},m}(\mathcal{A}) \models U(g(v)). \end{aligned}$$

Next, consider a constraint of the form $\mathrm{Comp}_{u^{-1}k, v^{-1}k}(u, v)$ in Φ, and let $r := |\mathrm{im}(k)|$. Let $\chi(z_1, \ldots, z_s)$ be a qf-formula in the language of $\mathcal{B}$ and let $t \colon [s] \to [r]$. Suppose that $\chi(z_{u^{-1}kt(1)}, \ldots, z_{u^{-1}kt(s)})$ is in $g(u)$. From (‡) we obtain that $\mathcal{B} \models \chi(h(kt(1)), \ldots, h(kt(s)))$. Again by (‡) we get that $\chi(z_{v^{-1}kt(1)}, \ldots, z_{v^{-1}kt(s)})$ is in $g(v)$. Hence,

$$T_{\mathcal{B},m}(\mathcal{A}) \models \mathrm{Comp}_{u^{-1}k, v^{-1}k}(g(u), g(v))$$

holds.

(Φ satisfiable implies Ψ satisfiable.) Conversely, suppose that Φ is satisfiable in $T_{\mathcal{B},m}(\mathcal{A})$. That is, there exists a map h from I to the m-types in $\mathcal{B}$ that satisfies all conjuncts of Φ. We show how to obtain an assignment $\{x_1, \ldots, x_n\} \to A$ that satisfies Ψ in $\mathcal{A}$. Define an equivalence relation $\sim$ on V as follows. Let $x, y \in V$. Let $u \in I$ be such that there are $p, q \in [m]$ such that $u(p) = x$ and $u(q) = y$. We define $x \sim y$ if, and only if, $h(u)$ contains the formula $z_p = z_q$. Note that the choice of u is not important: if u', p', q' are such that $u'(p') = x$ and $u'(q') = y$, the intersection of $\mathrm{im}(u)$ and $\mathrm{im}(u')$

contains $\{x, y\}$. Let $k\colon [r] \to \mathrm{im}(u) \cap \mathrm{im}(u')$ be a bijection. By construction, the constraint $\mathrm{Comp}_{u^{-1}k, u'^{-1}k}(u, u')$ is satisfied by h, which by definition of the relation means that $h(u)$ contains $z_p = z_q$ iff $h(u')$ contains $z_{p'} = z_{q'}$.

We prove that $\sim$ is an equivalence relation. Reflexivity and symmetry are clear from the definition. Assume that $x \sim y$ and $y \sim z$. Let $w \in I, p, q, r$ be such that $w(p) = x, w(q) = y$, and $w(r) = z$, which is possible since $m \geq 3$. Since $x \sim y$, the previous paragraph implies that $h(w)$ contains the formula $z_p = z_q$. Similarly, since $y \sim z$, the formula $z_q = z_r$ is in $h(w)$. Since $h(w)$ is a type, transitivity of equality implies that $z_p = z_r$ is in $h(w)$, so that $x \sim z$.

Define a structure $\mathcal{C}$ on $V/_\sim$ as follows. For every k-ary relation symbol R of $\mathcal{B}$ and k elements $[y_1], \ldots, [y_k]$ of $V/_\sim$, let $w \in I, p_1, \ldots, p_k \in [m]$ be such that $w(p_i) = y_i$ (such a w exists since $m \geq k$). Add the tuple $([y_1], \ldots, [y_k])$ to $R^{\mathcal{C}}$ if and only if $h(w)$ contains the formula $R(z_{p_1}, \ldots, z_{p_k})$. As in the paragraph above, this definition does not depend on the choice of the representatives $y_1, \ldots, y_k$ or on the choice of w. Proving that the definition does not depend on w is straightforward. Suppose now that $y_1 \sim y_1'$, and let $w \in I$ be such that $(w(p_1), \ldots, w(p_k)) = (y_1, \ldots, y_k)$ and such that $h(w)$ contains $R(z_{p_1}, \ldots, z_{p_k})$. Let $w' \in I$ be such that $(w'(q), w'(p_1'), \ldots, w'(p_k')) = (y_1', y_1, y_2, \ldots, y_k)$, which is possible since $m \geq k + 1$. We prove that $h(w')$ contains $R(z_q, z_{p_2'}, \ldots, z_{p_k'})$. Since $y \sim y'$, we have that $h(w')$ contains $z_q = z_{p_1'}$. Moreover, the images of w' and w intersect on $y_1, \ldots, y_k$, and since h satisfies the Comp constraints, we obtain that $h(w')$ contains $R(z_{p_1'}, \ldots, z_{p_k'})$. It follows that $h(w')$ contains $R(z_q, z_{p_2'}, \ldots, z_{p_k'})$. Therefore, the definition of R in $\mathcal{C}$ does not depend on the choice of the representative for the first entry of the tuple. By iterating this argument for each coordinate, we obtain that $R^{\mathcal{C}}$ is well-defined.

We claim that $\mathcal{C}$ embeds into $\mathcal{B}$. Otherwise, there would exist a bound $\mathcal{D}$ of size $k \leq m$ for $\mathcal{B}$ such that $\mathcal{D}$ embeds into $\mathcal{C}$. Let $[y_1], \ldots, [y_k]$ be the elements of the image of $\mathcal{D}$ under this embedding. Since $k \leq m$, there exist $w \in I, p_1, \ldots, p_k$ such that $(w(p_1), \ldots, w(p_k)) = (y_1, \ldots, y_k)$. The quantifier-free type of $([y_1], \ldots, [y_k])$ in $\mathcal{C}$ is in $h(w)$, by the previous paragraph. It follows that if $(a_1, \ldots, a_m) \in B^m$ is a tuple whose quantifier-free type is $h(w)$, there is an embedding of $\mathcal{D}$ into the substructure of $\mathcal{B}$ induced by $\{a_1, \ldots, a_m\}$. This contradicts the fact that $\mathcal{D}$ does not embed into $\mathcal{B}$.

Let e be an embedding $\mathcal{C} \hookrightarrow \mathcal{B}$. For $x \in V$ define $f(x) := e([x])$. We claim that $f\colon \{x_1, \ldots, x_n\} \to A$ is a valid assignment for Ψ. Let $R(j(1), \ldots, j(r))$ be a constraint from Ψ, where $j\colon [r] \to V$. Let $v \in I$ be such that $\mathrm{im}(j) \subseteq \mathrm{im}(v)$, and such that the constraint $\langle \chi(z_{v^{-1}j(1)}, \ldots, z_{v^{-1}j(r)}) \rangle(v)$ is in Φ. Since h satisfies this constraint, $h(v)$ contains $\chi(z_{v^{-1}j(1)}, \ldots, z_{v^{-1}j(r)})$. It follows that $\mathcal{C} \models \chi([j(1)], \ldots, [j(r)])$. Since e embeds $\mathcal{C}$ into $\mathcal{B}$, we obtain $\mathcal{B} \models \chi(f(j(1)), \ldots, f(j(r)))$, whence $\mathcal{A} \models R(f(j(1)), \ldots, f(j(r)))$, as required.

The given reduction can be performed in polynomial time: the number of variables in the new instance is in $O(n^m)$, and if c is the number of constraints in Ψ, then the number of constraints in Φ is in $O(cn^m + n^{2m})$. Each of the new constraints can be constructed in constant time. $\square$

We mention that the reduction is in fact a *first-order reduction*. Indeed, given an input $\mathcal{C}$ to $\mathrm{CSP}(\mathcal{A})$, the corresponding input $\mathcal{D}$ of $\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$ that we build is definable by first-order formulas (the domain of $\mathcal{D}$ is the set of m-tuples from $\mathcal{C}$, and each constraint in $\mathcal{D}$ is definable in terms of the constraints of $\mathcal{C}$ (see [1] for a formal definition of first-order reductions). Since first-order reductions have low complexity [60], we obtain that if

$\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$ is solvable by Datalog or is in L, then so is $\mathrm{CSP}(\mathcal{A})$.

We also note that Theorem 3.1 applies to all CSPs that can be described by an SNP sentence (SNP is defined in Chapter 5; for SNP in connection to CSPs see, e.g., [52]).

3.2 Canonical Functions

Since $\mathrm{CSP}(\mathcal{A})$ reduces to $\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$ when $\mathcal{A}$ is a quantifier-free reduct of the finitely bounded structure $\mathcal{B}$, we can derive upper bounds on the complexity of $\mathrm{CSP}(\mathcal{A})$ from upper bounds on the complexity of $\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$. We know from Section 2.5.2 that the complexity of $\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$ is encoded into the identities that are satisfiable in $\mathrm{Pol}(T_{\mathcal{B},m}(\mathcal{A}))$, since $T_{\mathcal{B},m}(\mathcal{A})$ is a finite structure. In this section, our goal is then to investigate the structure of $\mathrm{Pol}(T_{\mathcal{B},m}(\mathcal{A}))$, especially in the case that $\mathcal{B}$ is homogeneous.

Remember that the domain of $T_{\mathcal{B},m}(\mathcal{A})$ consists of the quantifier-free types in $\mathcal{B}$. When $\mathcal{B}$ is homogeneous, we have by Proposition 2.9 and Theorem 2.8 that the quantifier-free types correspond to orbits of tuples under $\mathrm{Aut}(\mathcal{B})$. A polymorphism of $T_{\mathcal{B},m}(\mathcal{A})$ is therefore an operation on the set of orbits of m-tuples under $\mathrm{Aut}(\mathcal{B})$. This motivates the following definition. Let $\mathcal{A}^m\big/\mathrm{Aut}(\mathcal{A})$ be the set of orbits of m-tuples under the natural action of $\mathrm{Aut}(\mathcal{A})$ on A^m.

Definition 3.1. Let $\mathcal{A}, \mathcal{B}$ be two structures. A function $f\colon A^k \to B$ is *canonical* from $\mathcal{A}$ to $\mathcal{B}$ if for all $m \geq 1$, all m-tuples $\bar{t}_1, \ldots, \bar{t}_k \in A^m$ and all $\alpha_1, \ldots, \alpha_k \in \mathrm{Aut}(\mathcal{A})$, there exists $\beta \in \mathrm{Aut}(\mathcal{B})$ such that $\beta f(\bar{t}_1, \ldots, \bar{t}_k) = f(\alpha_1(\bar{t}_1), \ldots, \alpha_k(\bar{t}_k))$.

One can rephrase Definition 3.1 as follows. A function $f\colon A^k \to B$ is canonical from $\mathcal{A}$ to $\mathcal{B}$ if for all $m \geq 1$ it induces a function $f_m^{\mathrm{typ}}\colon \left(\mathcal{A}^m\big/\mathrm{Aut}(\mathcal{A})\right)^k \to \mathcal{B}^m\big/\mathrm{Aut}(\mathcal{B})$ as follows: on input $O_1, \ldots, O_k \in \mathcal{A}^m\big/\mathrm{Aut}(\mathcal{A})$, let $\bar{a}^1, \ldots, \bar{a}^k$ be m-tuples of elements of $\mathcal{A}$ such that the orbit of $\bar{a}^i$ under $\mathrm{Aut}(\mathcal{A})$ is O_i. Let O' be the orbit of the m-tuple $f(\bar{a}^1, \ldots, \bar{a}^k)$ under $\mathrm{Aut}(\mathcal{A})$, and define $f_m^{\mathrm{typ}}(O_1, \ldots, O_k) := O'$. It follows from the definition of canonicity that this definition does not depend on the choice of the tuples $\bar{a}^1, \ldots, \bar{a}^k$. The function f_m^{typ} is called the *behaviour* of f on m-tuples. Note that if $\mathcal{A}$ is a homogeneous structure in a finite language whose relations have arity at most m, then $f\colon A^k \to A$ is canonical if and only if it induces a function $f_m^{\mathrm{typ}}\colon \left(\mathcal{A}^m\big/\mathrm{Aut}(\mathcal{A})\right)^k \to \mathcal{B}^m\big/\mathrm{Aut}(\mathcal{B})$ for this particular m. Moreover, if $\mathcal{A}$ and $\mathcal{B}$ are ω-categorical structures, there are only finitely many behaviours on m-tuples as both $\mathcal{A}^m\big/\mathrm{Aut}(\mathcal{A})$ and $\mathcal{B}^m\big/\mathrm{Aut}(\mathcal{B})$ are finite. We say that f is canonical with respect to $\mathcal{A}$ if it is canonical from $\mathcal{A}$ to $\mathcal{A}$.

Example 10. We illustrate the notion of canonicity for the structure $\mathcal{Q} = (\mathbb{Q}; <)$. Recall from Example 5 that $\mathcal{Q}$ is ω-categorical, and is moreover homogeneous in a finite relational language whose relations have arity at most 2. Thus, unary functions canonical with respect to $\mathcal{Q}$ correspond to behaviours $\mathcal{Q}^2\big/\mathrm{Aut}(\mathcal{Q}) \to \mathcal{Q}^2\big/\mathrm{Aut}(\mathcal{Q})$ (but not every behaviour is realisable by an actual function on $\mathbb{Q}$). For example, the identity behaviour

$$\{(x,y) \in \mathbb{Q}^2 \mid x < y\} \mapsto \{(x,y) \in \mathbb{Q}^2 \mid x < y\}$$
$$\{(x,y) \in \mathbb{Q}^2 \mid x = y\} \mapsto \{(x,y) \in \mathbb{Q}^2 \mid x = y\}$$
$$\{(x,y) \in \mathbb{Q}^2 \mid x > y\} \mapsto \{(x,y) \in \mathbb{Q}^2 \mid x > y\}$$

is realisable by any increasing function, and thus any increasing function is canonical with respect to $\mathcal{Q}$.

The behaviour

$$\{(x,y) \in \mathbb{Q}^2 \mid x < y\} \mapsto \{(x,y) \in \mathbb{Q}^2 \mid x > y\}$$
$$\{(x,y) \in \mathbb{Q}^2 \mid x = y\} \mapsto \{(x,y) \in \mathbb{Q}^2 \mid x = y\}$$
$$\{(x,y) \in \mathbb{Q}^2 \mid x > y\} \mapsto \{(x,y) \in \mathbb{Q}^2 \mid x < y\}$$

is realisable by any decreasing function. Another example is the behaviour of a constant function. It can be checked that any other behaviour is *not* realisable as a function $\mathbb{Q} \to \mathbb{Q}$, and there are therefore essentially only three unary canonical functions with respect to $\mathcal{Q}$.

It is easy to check that the set of functions that are canonical with respect to $\mathcal{A}$ forms a clone; similarly, for every clone $\mathscr{A}$ on A, the set of functions in $\mathscr{A}$ that are canonical with respect to $\mathcal{A}$ forms a subclone of $\mathscr{A}$.

Canonical functions were discovered by Bodirsky and Pinsker [33], where they were used to classify all minimal closed clones containing the automorphism group of the random graph. Since then, their usefulness has been demonstrated in many contexts: they were used to classify the first-order reducts of the random ordered graph [38], to classify the complexity of the CSP of first-order reducts of $(\mathbb{Q}, <)$ [23] and the random graph [34], or to give decidability results for variants of CSPs with *infinite* input structures [66, 67]. For us, the interest in canonical functions comes from the following lemmas where we relate $\mathrm{Pol}(T_{\mathcal{B},m}(\mathcal{A}))$ with the clone of polymorphisms of $\mathcal{A}$ that are canonical with respect to $\mathcal{B}$.

Let $\mathcal{A}$ be a reduct of a homogeneous relational structure $\mathcal{B}$, and let $\mathscr{C}$ be the subclone of $\mathrm{Pol}(\mathcal{A})$ that consists of the functions that are canonical with respect to $\mathcal{B}$. Let $\mathscr{C}_m^{\mathrm{typ}} = \{f_m^{\mathrm{typ}} \mid f \in \mathscr{C}\}$. It is easy to check that $\mathscr{C}_m^{\mathrm{typ}}$ is a clone of functions on $\mathcal{B}^m \big/ \mathrm{Aut}(\mathcal{B})$, and that the map $\xi_m^{\mathrm{typ}} \colon f \mapsto f_m^{\mathrm{typ}}$ is a continuous clone homomorphism from $\mathscr{C}$ to $\mathscr{C}_m^{\mathrm{typ}}$.

Lemma 3.2. *Let $\mathcal{A}$ be a reduct of a homogeneous relational structure $\mathcal{B}$ and let $\mathscr{C}$ be the clone of polymorphisms of $\mathcal{A}$ that are canonical with respect to $\mathcal{B}$. For all $m \geq 1$, we have $\mathscr{C}_m^{\mathrm{typ}} \subseteq \mathrm{Pol}(T_{\mathcal{B},m}(\mathcal{A}))$.*

Proof. First, note that the inclusion in the statement makes sense: since $\mathcal{B}$ is homogeneous, it is ω-categorical and admits quantifier-elimination by Proposition 2.9. Thus, orbits under $\mathrm{Aut}(\mathcal{B})$ and quantifier-free types in $\mathcal{B}$ are in one-to-one correspondence, and $\mathscr{C}_m^{\mathrm{typ}}$ can be seen as a clone on the domain of $T_{\mathcal{B},m}(\mathcal{A})$.

We have to show that $\xi_m^{\mathrm{typ}}(f) \in \mathrm{Pol}(T_{\mathcal{B},m}(\mathcal{A}))$ for every $f \in \mathscr{C}$. Let k be the arity of f. Let $\chi(z_1, \ldots, z_r)$ be a qf-definition of a relation of $\mathcal{A}$. Let $i \colon [r] \to [m]$ and let $p_1, \ldots, p_k$ be types in the relation $\langle \chi(z_{i(1)}, \ldots, z_{i(r)}) \rangle$ of $T_{\mathcal{B},m}(\mathcal{A})$. Let $\overline{a}^1, \ldots, \overline{a}^k$ be m-tuples whose types are $p_1, \ldots, p_k$ respectively. Since $\mathcal{B}$ is homogeneous the orbits of m-tuples under $\mathrm{Aut}(\mathcal{B})$ and the qf-types of $\mathcal{B}$ are in one-to-one correspondence and $\xi_m^{\mathrm{typ}}(f)$ can be seen as an operation on m-types. We have that $\xi_m^{\mathrm{typ}}(f)(p_1, \ldots, p_k)$ is the type of $f(\overline{a}^1, \ldots, \overline{a}^k)$ in $\mathcal{B}$. Since f preserves the relation defined by $\chi(z_{i(1)}, \ldots, z_{i(r)})$, it follows that $f(\overline{a}^1, \ldots, \overline{a}^k)$ satisfies $\chi(z_{i(1)}, \ldots, z_{i(r)})$, which means that $\chi(z_{i(1)}, \ldots, z_{i(r)})$ is contained in the type of this tuple. Therefore, $\xi_m^{\mathrm{typ}}(f)$ preserves the relations of $T_{\mathcal{B},m}(\mathcal{A})$ of the first kind.

We now prove that $\xi_m^{\mathrm{typ}}(f)$ preserves the relations of the second kind in $T_{\mathcal{B},m}(\mathcal{A})$. Indeed, let $(p_1, q_1), \ldots, (p_k, q_k)$ be pairs of types in $\mathrm{Comp}_{i,j}$. Let $(\overline{a}^1, \overline{b}^1), \ldots, (\overline{a}^k, \overline{b}^k)$ be pairs

of m-tuples such that $\mathrm{tp}(\overline{a}^l) = p_l$ and $\mathrm{tp}(\overline{b}^l) = q_l$ for all $l \in [k]$. As noted above, the definition of $\mathrm{Comp}_{i,j}$ implies that the tuples $(a^l_{i(1)}, \dots, a^l_{i(r)})$ and $(b^l_{j(1)}, \dots, b^l_{j(r)})$ have the same type in $\mathcal{B}$ for all $l \in [k]$. Since f is canonical, we have that $(f(a^1_{i(1)}, \dots, a^k_{i(1)}), \dots, f(a^1_{i(r)}, \dots, a^k_{i(r)}))$ has the same type as $(f(b^1_{j(1)}, \dots, b^k_{j(1)}), \dots, f(b^1_{j(r)}, \dots, b^k_{j(r)}))$ in $\mathcal{B}$. This implies that

$$\mathrm{Comp}_{i,j}\big(\xi^{\mathrm{typ}}_m(f)(p_1, \dots, p_k), \xi^{\mathrm{typ}}_m(f)(q_1, \dots, q_k)\big)$$

holds in $T_{\mathcal{B},m}(\mathcal{A})$, which concludes the proof. □

Suppose that $\mathcal{B}$ is homogeneous in a finite relational language, and that $\mathcal{A}$ is a reduct of $\mathcal{B}$. Suppose moreover that every polymorphism of $\mathcal{A}$ is canonical with respect to $\mathrm{Aut}(\mathcal{B})$. The lemma above implies that ξ^{typ}_m is a continuous homomorphism from $\mathrm{Pol}(\mathcal{A})$ to $\mathrm{Pol}(T_{\mathcal{B},m}(\mathcal{A}))$, if m is greater than the arity of the language of $\mathcal{B}$. This in turn implies that there is a polynomial-time reduction from $\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$ to $\mathrm{CSP}(\mathcal{A})$ [35]. This proves the following corollary.

Corollary 3.3. *Let $\mathcal{B}$ be a finitely bounded homogeneous structure in a finite relational language, and let $\mathcal{A}$ be a first-order reduct of $\mathcal{B}$. Let m be defined as in Theorem 3.1. Suppose that all the polymorphisms of $\mathcal{A}$ are canonical with respect to $\mathcal{B}$. Then* $\mathrm{CSP}(\mathcal{A})$ *and* $\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$ *are polynomial-time equivalent. In particular, Conjecture 1 holds for all such structures $\mathcal{A}$.*

Remark 1. Let $\mathcal{B}$ be a homogeneous structure in a language of maximal arity m. Let $R_{\mathcal{B}}$ be the relation of arity $2m$ such that for $\overline{a}, \overline{b} \in B^m$, the $2m$-tuple $(\overline{a}, \overline{b})$ is in $R_{\mathcal{B}}$ if, and only if, $\overline{a}$ and $\overline{b}$ are in the same orbit under $\mathrm{Aut}(\mathcal{B})$. Note that a function $f\colon B^k \to B$ is canonical with respect to $\mathcal{B}$ if, and only if, f preserves $R_{\mathcal{B}}$. Thus, a consequence of Corollary 3.3 is that Conjecture 1 holds for all reducts $\mathcal{A}$ of $\mathcal{B}$ such that $R_{\mathcal{B}}$ is pp-definable over $\mathcal{A}$.

If $\mathcal{A}$ is a reduct of a finitely bounded homogeneous structure $\mathcal{B}$, then the inclusion in Lemma 3.2 becomes an equality, for m large enough. This fact is only mentioned for completeness and not used later, so we only sketch the proof.

Lemma 3.4. *Let $\mathcal{A}$ be a reduct of a finitely bounded homogeneous structure $\mathcal{B}$ and let $\mathscr{C}$ be the polymorphisms of $\mathcal{A}$ that are canonical with respect to* $\mathrm{Aut}(\mathcal{B})$*. Let m be larger than each bound of $\mathcal{B}$ and strictly larger than the maximal arity of $\mathcal{A}$ and $\mathcal{B}$. Then $\mathscr{C}^{\mathrm{typ}}_m =$* $\mathrm{Pol}(T_{\mathcal{B},m}(\mathcal{A}))$.

Proof sketch. The inclusion $\mathscr{C}^{\mathrm{typ}}_m \subseteq \mathrm{Pol}(T_{\mathcal{B},m}(\mathcal{A}))$ has been shown in Lemma 3.2. For the reverse inclusion, we prove that for every $g \in \mathrm{Pol}(T_{\mathcal{B},m}(\mathcal{A}))$ there exists an $f \in \mathscr{C}$ such that $\xi^{\mathrm{typ}}_m(f) = g$. Let k be the arity of g. We prove that for every subset F of A there exists a function h from $F^k \to A$ such that for all $\bar{a}^1, \dots, \bar{a}^k \in F^m$ whose types are $p_1, \dots, p_k$, respectively, $h(\bar{a}^1, \dots, \bar{a}^k)$ has type $g(p_1, \dots, p_k)$. A standard compactness argument then shows the existence of a function $f\colon A^k \to A$ such that for all $\bar{a}^1, \dots, \bar{a}^k \in A^m$ whose types are $p_1, \dots, p_k$, respectively, $f(\bar{a}^1, \dots, \bar{a}^k)$ has type $g(p_1, \dots, p_k)$, and such a function must satisfy $\xi^{\mathrm{typ}}_m(f) = g$.

Note that we can assume without loss of generality that $\mathcal{B}$ has for each relation symbol R also a relation symbol for the complement of $R^{\mathcal{B}}$. This does not change $\mathscr{C}^{\mathrm{typ}}_m$ or $T_{\mathcal{B},m}(\mathcal{A})$. The existence of a function h with the properties as stated above can then be expressed as

an instance Ψ of CSP($\mathcal{B}$) where the variable set is F^k and where we impose constraints from $\mathcal{B}$ on $\bar{a}^1, \dots, \bar{a}^k$ to enforce that in any solution h to this instance the tuple $h(\bar{a}^1, \dots, \bar{a}^k)$ satisfies $g(p_1, \dots, p_k)$. Let Φ be the instance of CSP($T_{\mathcal{B},m}(\mathcal{B})$) obtained from Ψ under the reduction from CSP($\mathcal{B}$) to CSP($T_{\mathcal{B},m}(\mathcal{B})$) described in the proof of Theorem 3.1. The variables of Φ are the order-preserving injections from $[m]$ to F^k. For $v\colon [m] \to F^k$ and $i \leq k$, let p_i be the type of $(v(1).i, \dots, v(m).i)$ in $\mathcal{B}$. Then the mapping h that sends v to $g(p_1, \dots, p_k)$, for all variables v of Φ, is a solution to Φ:

- the constraints of Φ of the form $\langle \chi(\dots) \rangle(v)$ have been introduced to translate constraints of Ψ, and it is easy to see that they are satisfied by the choice of these constraints of Ψ and by the choice of h.
- The other constraints of Φ are of the form $\mathrm{Comp}_{i,j}(u, v)$ where u, v are order-preserving injections from $[m]$ to F^k. Since g is a k-ary polymorphism of $\mathrm{Pol}(T_{\mathcal{B},m}(\mathcal{B}))$ and hence preserves the relations $\mathrm{Comp}_{i,j}$ of $T_{\mathcal{B},m}(\mathcal{B})$, it follows that h satisfies these constraints, too.

By Theorem 3.1, the instance Ψ of CSP($\mathcal{B}$) is satisfiable, too. □

It is also possible to relate the identities that are satisfiable in $\mathscr{C}_m^{\mathrm{typ}}$ with the identities that are satisfiable in $\mathscr{C}$. Clearly, since ξ_m^{typ} is a clone homomorphism from $\mathscr{C}$ to $\mathscr{C}_m^{\mathrm{typ}}$, every identity satisfiable in $\mathscr{C}$ is satisfiable in $\mathscr{C}_m^{\mathrm{typ}}$. The converse is not necessarily true, but we have the following lemma and proposition.

Lemma 3.5 (Bodirsky, Pinsker, Pongrácz [37]). *Let $\mathcal{B}$ be an ω-categorical structure and let $f, g\colon B^k \to B$. Let $s(x_1, \dots, x_k) \approx t(x_1, \dots, x_k)$ be an equation. Suppose that for every finite subset A of B, there exists an automorphism α of $\mathcal{B}$ such that $\alpha f(a_1, \dots, a_k) = g(a_1, \dots, a_k)$ holds for all $a_1, \dots, a_k \in A$. Then there exists $e_1, e_2 \in \overline{\mathrm{Aut}(\mathcal{B})}$ such that $e_1 f(a_1, \dots, a_k) = e_2 g(a_1, \dots, a_k)$ holds for all $a_1, \dots, a_k \in B$.*

Proposition 3.6 (Bodirsky, Pinsker, Pongrácz [37]). *Let $m \geq 1$, and let $\mathcal{B}$ be a homogeneous structure whose relations have arity at most m. Let $\mathscr{C}$ be a closed clone of functions on B that are canonical with respect to $\mathcal{B}$. Let $f(x_1, \dots, x_n) \approx g(x_1, \dots, x_n)$ be an equation that is satisfiable in $\mathscr{C}_m^{\mathrm{typ}}$. Then the equation $e_1 f(x_1, \dots, x_n) \approx e_2 g(x_1, \dots, x_n)$ is satisfiable with $f, g \in \mathscr{C}$ and $e_1, e_2 \in \overline{\mathrm{Aut}(\mathcal{B})}$.*

In particular, there is a statement similar to Theorem 2.17 for clones of canonical functions. Let $f\colon A^n \to A$ and let $\mathscr{U}$ be a set of unary functions on A. We say that f is *pseudo-cyclic* modulo $\mathscr{U}$ if there exist $e_1, e_2 \in \mathscr{U}$ such that

$$\forall a_1, \dots, a_n \in A,\ e_1 f(a_1, \dots, a_n) = e_2 f(a_2, \dots, a_n, a_1)$$

holds. Similarly, f is *pseudo weak near-unanimity* (pseudo-WNU) modulo $\mathscr{U}$ if there exist $e_1, \dots, e_n \in \mathscr{U}$ such that

$$\forall a, b \in A,\ e_1 f(a, b, \dots, b) = e_2 f(b, a, b, \dots, b) = \dots = e_n f(b, \dots, b, a)$$

holds.

Corollary 3.7. *Let $m \geq 1$, and let $\mathcal{B}$ be a homogeneous structure whose relations have arity at most m. Let $\mathscr{C}$ be a closed clone of functions on B that are canonical with respect to $\mathcal{B}$. Suppose that $\mathscr{C}_m^{\mathrm{typ}}$ is idempotent. The following are equivalent:*

1. *There is no continuous clone homomorphism $\mathscr{C} \to \mathscr{P}$,*
2. *$\mathscr{C}$ satisfies a non-trivial identity,*
3. *$\mathscr{C}$ contains a pseudo-WNU operation modulo $\overline{\mathrm{Aut}(\mathcal{B})}$,*
4. *$\mathscr{C}$ contains a pseudo-cyclic operation modulo $\overline{\mathrm{Aut}(\mathcal{B})}$,*
5. *$\mathscr{C}$ contains a pseudo-Siggers operation modulo $\overline{\mathrm{Aut}(\mathcal{B})}$.*

Note that item 1. in the previous statement is about clone homomorphisms (as opposed to clonoid homomorphisms), due to the fact that the identities from items 3., 4., and 5. do not have height 1. It is unknown whether the items in the previous statement are equivalent to the absence of a uniformly continuous clonoid homomorphism from $\mathscr{C}$ to $\mathscr{P}$.

3.3 New Abstract Tractability Conditions

Most of the known conditions that imply that $\mathrm{CSP}(\mathcal{A})$ is in P are *concrete* conditions, i.e., they are conditions on the abstract algebraic structure of $\mathrm{Pol}(\mathcal{A})$ (such as identities) as well as conditions on the underlying structure $\mathcal{A}$. One notable exception is tractability from *quasi near-unanimity polymorphisms*, that is, polymorphisms that satisfy the identity

$$\begin{aligned} f(y,x,\dots,x) &= f(x,y,x,\dots,x) = \cdots \\ &= f(x,\dots,x,y) = f(x,\dots,x). \end{aligned}$$

If $\mathcal{A}$ has a quasi near-unanimity polymorphism then $\mathrm{CSP}(\mathcal{A})$ is in P [14]: this is an *abstract* tractability condition. Using Lemma 3.2, one can derive new tractability conditions for reducts of finitely bounded homogeneous structures. The tractability conditions that we are able to lift this way from the finite are all of the abstract type.

Theorem 3.8. *Let $\mathcal{A}$ be a finite-signature reduct of a finitely bounded homogeneous structure $\mathcal{B}$. Suppose that $\mathcal{A}$ has a four-ary polymorphism f and a ternary polymorphism g that are canonical with respect to $\mathcal{B}$, that are weak near-unanimity operations modulo $\overline{\mathrm{Aut}(\mathcal{B})}$, and such that there are operations e_1, e_2 in $\overline{\mathrm{Aut}(\mathcal{B})}$ with $e_1(f(y,x,x,x)) = e_2(g(y,x,x))$ for all x, y. Then $\mathrm{CSP}(\mathcal{A})$ is in P.*

Proof. Let m be as in the statement of Theorem 3.1. By Lemma 3.2, $f' := \xi_m^{\mathrm{typ}}(f)$ and $g' := \xi_m^{\mathrm{typ}}(g)$ are polymorphisms of $T_{\mathcal{B},m}(\mathbb{A})$. Moreover, f' and g' must be weak near-unanimity operations, and they satisfy $f'(y,x,x,x) = g'(y,x,x)$. It follows from [70] in combination with [3] that $T_{\mathcal{B},m}(\mathbb{A})$ is in P (it can be solved by a Datalog program). Theorem 3.1 then implies that $\mathrm{CSP}(\mathcal{A})$ is in P, too. □

Note that since the reduction from $\mathrm{CSP}(\mathcal{A})$ to $\mathrm{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$ presented in Section 3.1 is a first-order reduction, it is computable in Datalog. In particular, the hypotheses of Theorem 3.8 imply that $\mathrm{CSP}(\mathcal{A})$ is in Datalog. This result generalises many tractability results from the literature, for instance

- the polynomial-time tractable fragments of RCC-5 [63];
- the two polynomial-time algorithms for partially-ordered time from [41];

- polynomial-time tractable equality constraints [21];
- all polynomial-time tractable equivalence CSPs [40].

In all cases, the respective structures $\mathcal{A}$ have a polymorphism f such that $\xi_2^{\text{typ}}(f)$ is a semilattice operation [62]. Finite structures with a semilattice polymorphism also have weak near-unanimity polymorphisms f' and g' that satisfy $f'(y,x,x,x) = g'(y,x,x)$ (see [70]), and hence $\mathcal{A}$ satisfies the assumptions of Theorem 3.8.

Using the same idea as in Theorem 3.1, one obtains a series of new abstract tractability conditions: for every known abstract tractability condition for finite domain CSPs, we obtain an abstract tractability condition for reducts of finitely bounded homogeneous structures $\mathcal{B}$. To show this, we first observe that the functions on A that are canonical with respect to $\text{Aut}(\mathcal{A})$ can be characterised algebraically.

Proposition 3.9. *Let $\mathcal{B}$ be a homogeneous model-complete core with a finite relational language. Then $f\colon B^n \to B$ is canonical with respect to $\mathcal{B}$ if and only if for all $a_1,\dots,a_n \in \text{End}(\mathcal{B})$ there exist $e_1, e_2 \in \text{End}(\mathcal{B})$ such that*

$$e_1 \circ f \circ (a_1,\dots,a_n) = e_2 \circ f\,.$$

Proof. The "if" direction is clear. In the other direction, the assumption that f is canonical gives that for every finite subset of B, the equation $f \circ (a_1,\dots,a_n) \approx f$ is satisfiable modulo $\text{Aut}(\mathcal{B})$. By Lemma 3.5, there exist $e_1, e_2 \in \overline{\text{Aut}(\mathcal{B})} = \text{End}(\mathcal{B})$ such that $e_1 \circ f \circ (a_1,\dots,a_n) = e_2 \circ f$. □

Proposition 3.9 shows that the following close relative to Theorem 3.8 is an abstract tractability condition.

Theorem 3.10. *Let $\mathcal{A}$ be a finitely bounded homogeneous model-complete core. Suppose that $\mathcal{A}$ has a four-ary polymorphism f and a ternary polymorphism g that are canonical with respect to $\mathcal{A}$, that are weak near-unanimity operations modulo $\text{End}(\mathcal{A})$ and such that there are operations $e_1, e_2 \in \text{End}(\mathcal{A})$ with $e_1(f(y,x,x,x)) = e_2(g(y,x,x))$ for all x, y. Then $\text{CSP}(\mathcal{A})$ is in P.*

In the same way as in Theorem 3.10 every abstract tractability result for finite-domain CSPs can be lifted to an abstract tractability condition for ω-categorical CSPs. Note that the polynomial-time tractable cases in the classification for Graph-SAT problems [34] can also be explained with the help of Corollary 3.11 below, using the recent solution to the finite-domain tractability conjecture.

Corollary 3.11. *Let $\mathcal{A}$ be a finite-signature reduct of a finitely bounded homogeneous structure $\mathcal{B}$, and suppose that $\mathcal{A}$ has a Siggers (or weak near-unanimity, or cyclic) polymorphism f modulo operations from $\overline{\text{Aut}(\mathcal{B})}$ such that f is canonical with respect to $\mathcal{B}$. Then $\text{CSP}(\mathcal{A})$ is in P.*

Proof. Let m be as in the statement of Theorem 3.1. By Lemma 3.2, $\xi_m^{\text{typ}}(f)$ is a polymorphism of $T_{\mathcal{B},m}(\mathcal{A})$. Since $\xi_m^{\text{typ}}(f)$ is a Siggers operation, Theorem 2.5 implies that $\text{CSP}(T_{\mathcal{B},m}(\mathcal{A}))$ is in P and Theorem 3.1 implies that $\text{CSP}(\mathcal{A})$ is in P. □

Finally, we mention that the non-trivial polynomial-time tractable cases for reducts of $\mathcal{Q} = (\mathbb{Q}; <)$ provide examples that cannot be lifted from finite-domain tractability results this way, since the respective languages do not have non-trivial *canonical* polymorphisms. As an example, $\mathrm{CSP}(T_{\mathcal{Q},m}(\mathcal{Q}))$ is NP-complete for $m \geq 3$, while $\mathrm{CSP}(\mathcal{Q})$ and $\mathrm{CSP}(T_{\mathcal{Q},2}(\mathcal{Q}))$ are solvable in polynomial time.

Chapter 4

Mashups

Theorem 3.10 and Corollary 3.11 in the previous chapter are of the form "if the canonical polymorphisms of $\mathcal{A}$ satisfy some given nontrivial equation, then $\mathrm{CSP}(\mathcal{A})$ is in P." In this chapter, we investigate what can be said when the canonical polymorphisms of $\mathcal{A}$ do not satisfy nontrivial equations. We can reformulate this as follows. Let $\mathscr{C} \subseteq \mathrm{Pol}(\mathcal{A})$ be the clone of polymorphisms of $\mathcal{A}$ that are canonical (with respect to some homogeneous finitely bounded structure $\mathcal{B}$). Corollary 3.11 is then equivalent (by Corollary 3.7) to the statement: if there is no uniformly continuous clone homomorphism $\mathscr{C} \to \mathscr{P}$, then $\mathrm{CSP}(\mathcal{A})$ is in P. On the other hand, we know that the existence of a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{A}) \xrightarrow{u.c.c.h.} \mathscr{P}$ implies that $\mathrm{CSP}(\mathcal{A})$ is NP-complete (Corollary 2.15). This naturally raises the question as to when the existence of a uniformly continuous clone homomorphism $\mathscr{C} \to \mathscr{P}$ implies the existence of a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{A}) \xrightarrow{u.c.c.h.} \mathscr{P}$.

We focus on the case where every operation $f \in \mathrm{Pol}(\mathcal{A})$ interpolates modulo $\mathrm{Aut}(\mathcal{B})$ an operation that is canonical with respect to $\mathrm{Aut}(\mathcal{B})$. We call this the *canonisation property*. In this setting, there is a natural candidate for extending a clone homomorphism $\xi\colon \mathscr{C} \to \mathscr{P}$ to $\phi\colon \mathrm{Pol}(\mathcal{A}) \to \mathscr{P}$. Indeed, for every $f \in \mathrm{Pol}(\mathcal{A})$, define $\phi(f)$ to be $\xi(g)$, where g is any canonical function in $\overline{\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})}$. We prove that when this definition does not depend on the choice of g, then ϕ is indeed a uniformly continuous clonoid homomorphism. We also give a property of $\mathrm{Pol}(\mathcal{A})$ – that we call the *mashup property* – that implies that the extension ϕ is well-defined. This chapter contains published results from [30, 31].

4.1 Mashups

We start with the central definition of this chapter.

Definition 4.1. Let $g, h\colon B^k \to B$, let $1 \leq \ell \leq k$, and let $r, s \in B$. An operation ω is an *ℓ-mashup of g and h over $\{r, s\}$* if the following equations hold:

$$\omega(r, \dots, r, s, r, \dots, r) = g(r, \dots, r, s, r, \dots, r),$$
$$\omega(s, \dots, s, r, s, \dots, s) = h(s, \dots, s, r, s, \dots, s),$$

where the different entry in the arguments above is the ℓ-th entry. In case $\ell = 1$, we simply write *mashup*.

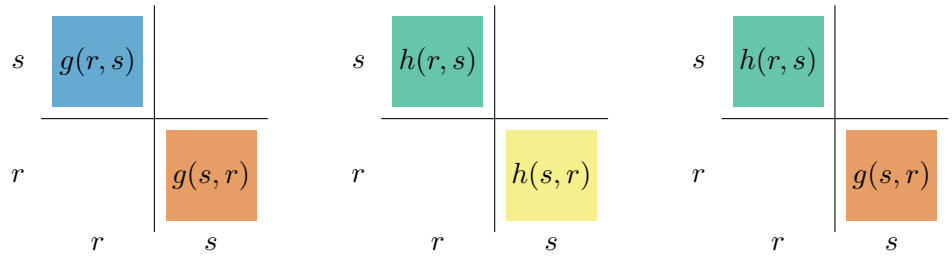

Figure 4.1: A mashup (right) of two binary functions (left and center).

In the following, we encourage the reader to work with the case $k = 2$ in mind. In this case, an example of a mashup of two binary functions over $\{r, s\}$ is shown in Figure 4.1.

The motivation for this definition is the following. If g and h are assumed to be projections, and we know that for all ℓ there exists an ℓ-mashup of g and h, then g and h must be the same projection.

Let $\mathbf{B}$ be an algebra. The class of all algebras $\mathbf{A}$ such that $\mathbf{A}$ is the homomorphic image of a subalgebra of $\mathbf{B}$ is denoted by $\mathsf{HS}(\mathbf{B})$. We call $\mathbf{A}$ a *subfactor* of $\mathbf{B}$. A *trivial algebra* is an algebra over at least 2 elements whose operations are all projections. Remember that an idempotent algebra $\mathbf{B}$ is an algebra whose every operation $f^{\mathbf{B}}$ satisfies $f^{\mathbf{B}}(b, \dots, b) = b$ for all $b \in B$.

Proposition 4.1. *Let $\mathbf{B}$ be a finite idempotent algebra, and let $\mathscr{B}$ be the clone generated by the fundamental operations of $\mathbf{B}$. The following are equivalent:*

1. *there is a clonoid homomorphism $\mathscr{B} \to \mathscr{P}$,*
2. *there is a clone homomorphism $\mathscr{B} \to \mathscr{P}$,*
3. *$\mathsf{HS}(\mathbf{B})$ contains a trivial algebra.*

Proof. The first two items are equivalent by Theorem 1.4 in [6]. The last two items were proved to be equivalent by Bulatov and Jeavons [45]. □

Lemma 4.2. *Let $\mathbf{B}$ be an algebra, and let $g^{\mathbf{B}}, h^{\mathbf{B}}$ be operations of $\mathbf{B}$ of arity k. Suppose that for all $\ell \in \{1, \dots, k\}$ and all distinct elements $r, s \in B$, there exists an operation of $\mathbf{B}$ that is an ℓ-mashup of $g^{\mathbf{B}}$ and $h^{\mathbf{B}}$ over $\{r, s\}$. Then for every trivial algebra $\mathbf{T}$ in $\mathsf{HS}(\mathbf{B})$, we have $g^{\mathbf{T}} = h^{\mathbf{T}}$.*

Proof. Let $\mathbf{S}$ be a subalgebra of $\mathbf{B}$ and μ be a homomorphism $\mathbf{S} \to \mathbf{T}$. Suppose that $g^{\mathbf{T}}$ is the ℓ-th projection. Let r, s be two elements of $\mathbf{S}$ which are mapped by μ to different elements of $\mathbf{T}$, and let $\omega^{\mathbf{B}}$ be an ℓ-mashup of $g^{\mathbf{B}}$ and $h^{\mathbf{B}}$ over $\{r, s\}$. By the assumption that $g^{\mathbf{T}}$ is the ℓ-th projection, we have

$$g^{\mathbf{T}}(\mu(r), \dots, \mu(r), \mu(s), \mu(r), \dots, \mu(r)) = \mu(s),$$

so that

$$\begin{aligned}
\omega^{\mathbf{T}}(\mu(r), \dots, \mu(r), \mu(s), \mu(r), \dots, \mu(r)) &= \mu(\omega^{\mathbf{B}}(r, \dots, r, s, r, \dots, r)) \\
&= \mu(g^{\mathbf{B}}(r, \dots, r, s, r, \dots, r)) \\
&= g^{\mathbf{T}}(\mu(r), \dots, \mu(r), \mu(s), \mu(r), \dots, \mu(r)) \\
&= \mu(s)
\end{aligned}$$

which implies that $\omega^{\mathbf{T}}$ is the ℓ-th projection, by the fact that $\mu(s) \neq \mu(r)$. Whence, $\omega^{\mathbf{T}}(\mu(s), \ldots, \mu(s), \mu(r), \mu(s), \ldots, \mu(s)) = \mu(r)$, and since $\omega^{\mathbf{B}}$ is a mashup of $g^{\mathbf{B}}$ and $h^{\mathbf{B}}$ over $\{r, s\}$, we obtain

$$\begin{aligned} h^{\mathbf{T}}(\mu(s), \ldots, \mu(s), \mu(r), \mu(s), \ldots, \mu(s)) &= \mu(h^{\mathbf{B}}(s, \ldots, s, r, s, \ldots, s)) \\ &= \mu(\omega^{\mathbf{B}}(s, \ldots, s, r, s, \ldots, s)) \\ &= \omega^{\mathbf{T}}(\mu(s), \ldots, \mu(s), \mu(r), \mu(s), \ldots, \mu(s)) \\ &= \mu(r) \end{aligned}$$

which implies that $h^{\mathbf{T}}$ is the ℓ-th projection and that $g^{\mathbf{T}} = h^{\mathbf{T}}$ holds. □

Let $\mathcal{A}, \mathcal{B}$ be two structures. In the following, the algebra $\mathbf{B}$ we consider is the algebra on $\mathcal{B}/\mathrm{Aut}(\mathcal{B})$ whose operations are of the form $\xi_1^{\mathrm{typ}}(f)$, where f is a polymorphism of $\mathcal{A}$ that is canonical with respect to $\mathcal{B}$. In order to prove that the mashup of two operations $\xi_1^{\mathrm{typ}}(g), \xi_1^{\mathrm{typ}}(h)$ exists in this algebra, we therefore need to prove that there exists a canonical function $\omega \in \mathrm{Pol}(\mathcal{A})$ that induces this mashup in $\mathbf{B}$. This motivates Definitions 4.2 and 4.3 below.

Definition 4.2. Let $\mathcal{A}, \mathcal{B}$ be two relational structures. We say that $\mathcal{B}$ has the *mashup property* relative to $\mathcal{A}$ if the following condition holds: for all $f \in \mathrm{Pol}(\mathcal{A})$, all functions g, h in $\overline{\mathrm{Aut}(\mathcal{B}) f \, \mathrm{Aut}(\mathcal{B})}$ that are canonical with respect to $\mathcal{B}$, and all orbits $O_1, O_2 \in \mathcal{B}/\mathrm{Aut}(\mathcal{B})$, we have that $\mathrm{Pol}(\mathcal{A})$ contains a function ω that is canonical with respect to $\mathcal{B}$ and such that $\xi_1^{\mathrm{typ}}(\omega)$ is a 1-mashup of $\xi_1^{\mathrm{typ}}(g)$ and $\xi_1^{\mathrm{typ}}(h)$ over $\{O_1, O_2\}$.

Note that in the definition above, it is equivalent to ask for the existence of a 1-mashup or for ℓ-mashups for all ℓ.

Definition 4.3. Let $\mathcal{A}, \mathcal{B}$ be two relational structures. We say that $\mathcal{B}$ has the *canonisation property* relative to $\mathcal{A}$ if for every operation $f \in \mathrm{Pol}(\mathcal{A})$, there exists in $\overline{\mathrm{Aut}(\mathcal{B}) f \, \mathrm{Aut}(\mathcal{B})}$ an operation g that is canonical with respect to $\mathcal{B}$.

We say that $\mathcal{B}$ has the canonisation property if it has the canonisation property relative to all its first-order reducts, and similarly for the mashup property. It is known that if $\mathrm{Aut}(\mathcal{B})$ is an oligomorphic extremely amenable group, then $\mathcal{B}$ has the canonisation property (this is a reformulation of Theorem 1 in [36]). In the next section, we give examples of structures $\mathcal{B}$ such that $\mathrm{Aut}(\mathcal{B})$ is not extremely amenable but such that $\mathcal{B}$ has the canonisation property.

Theorem 4.3 (Mashup theorem)**.** *Let $\mathcal{A}, \mathcal{B}$ be ω-categorical structures such that $\mathcal{A}$ is a first-order reduct of $\mathcal{B}$. Let $\mathscr{C}$ be the clone of polymorphisms of $\mathcal{A}$ that are canonical with respect to $\mathcal{B}$. Suppose that $\mathcal{B}$ has the mashup property and the canonisation property relative to $\mathcal{A}$, and that $\mathscr{C}_1^{\mathrm{typ}}$ is idempotent. If there exists a clonoid homomorphism from $\mathscr{C}_1^{\mathrm{typ}}$ to $\mathscr{P}$, then there exists a clonoid homomorphism ϕ from $\mathrm{Pol}(\mathcal{A})$ to $\mathscr{P}$. Moreover, ϕ is constant on sets of the form $\overline{\mathrm{Aut}(\mathcal{B}) f \, \mathrm{Aut}(\mathcal{B})}$ for $f \in \mathrm{Pol}(\mathcal{A})$.*

Proof. Let $\mathbf{B}$ be an algebra on $\mathcal{B}/\mathrm{Aut}(\mathcal{B})$ whose operations are exactly the operations in $\mathscr{C}_1^{\mathrm{typ}}$. The algebra $\mathbf{B}$ is idempotent by hypothesis on $\mathscr{C}_1^{\mathrm{typ}}$, and since $\mathcal{B}$ is ω-categorical,

$\mathbf{B}$ is finite. It follows from Proposition 4.1 that there exists a subalgebra $\mathbf{S}$ of $\mathbf{B}$ and a homomorphism $\mu\colon \mathbf{S} \to \mathbf{T}$ where $\mathbf{T}$ is a trivial algebra with two elements $\mu(r), \mu(s)$. Let $\xi'\colon \mathscr{C}_1^{\text{typ}} \to \mathscr{P}$ be the clone homomorphism that maps an operation $g^{\mathbf{B}}$ to $g^{\mathbf{T}}$. This homomorphism has the property that for two operations $g^{\mathbf{B}}, h^{\mathbf{B}}$ in $\mathscr{D}_1^{\text{typ}}$ of arity k, if for all $\ell \in \{1, \dots, k\}$ there is an operation in $\mathscr{C}_1^{\text{typ}}$ which is an ℓ-mashup of $g^{\mathbf{B}}$ and $h^{\mathbf{B}}$ over $\{r, s\}$, then $\xi'(g) = \xi'(h)$. By composition of ξ_1^{typ} with ξ', we obtain a clone homomorphism $\xi\colon \mathscr{C} \to \mathscr{P}$. Define the extension ϕ of ξ to the whole clone $\mathrm{Pol}(\mathcal{A})$ by setting $\phi(f) := \xi(g)$, where g is any function in $\overline{\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})}$ that is canonical with respect to $\mathcal{B}$ – such a function exists by the canonisation property, and is in $\mathrm{Pol}(\mathcal{A})$ since $\mathcal{A}$ is a first-order reduct of $\mathcal{B}$. We claim that ϕ is well-defined, and that it is a uniformly continuous clonoid homomorphism.

ϕ is well defined: let g, h be canonical and in $\overline{\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})}$. By the mashup property, we obtain for each $\ell \in \{1, \dots, k\}$ an operation $\omega \in \mathrm{Pol}(\mathcal{A})$ which is canonical and such that $\xi_1^{\text{typ}}(w)$ is an ℓ-mashup of $\xi_1^{\text{typ}}(g)$ and $\xi_1^{\text{typ}}(h)$ over $\{r, s\}$. Since this holds for all $\ell \in \{1, \dots, k\}$, we have by Lemma 4.2 that $\xi'(\xi_1^{\text{typ}}(g)) = \xi'(\xi_1^{\text{typ}}(h))$, i.e., $\xi(g) = \xi(h)$ and ϕ is well defined.

ϕ is constant on $\overline{\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})}$, for $f \in \mathrm{Pol}(\mathcal{A})$: let f' be in $\overline{\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})}$. Let g be canonical and interpolated by f' modulo $\mathrm{Aut}(\mathcal{B})$. Note that g is also interpolated by f modulo $\mathrm{Aut}(\mathcal{B})$, so that $\phi(f) = \xi(g) = \phi(f')$. It follows that ϕ is constant on $\overline{\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})}$.

ϕ is a clonoid homomorphism: we need to prove that

$$\phi(f \circ (\pi_{i_1}^m, \dots, \pi_{i_k}^m)) = \phi(f) \circ (\pi_{i_1}^m, \dots, \pi_{i_k}^m)$$

for every $f \in \mathrm{Pol}(\mathcal{A})$ of arity $k \geq 1$, every $m \geq 1$, and every $i_1, \dots, i_k \in \{1, \dots, m\}$. Let $g\colon B^k \to B$ be canonical and interpolated by f modulo $\mathrm{Aut}(\mathcal{B})$. Then $g \circ (\pi_{i_1}^m, \dots, \pi_{i_k}^m)$ is interpolated modulo $\mathrm{Aut}(\mathcal{B})$ by $f \circ (\pi_{i_1}^m, \dots, \pi_{i_k}^m)$. So

$$\begin{aligned}
\phi(f \circ (\pi_{i_1}^m, \dots, \pi_{i_k}^m)) &= \xi(g \circ (\pi_{i_1}^m, \dots, \pi_{i_k}^m)) && (4.1)\\
&= \xi(g) \circ (\pi_{i_1}^m, \dots, \pi_{i_k}^m) && (4.2)\\
&= \phi(f) \circ (\pi_{i_1}^m, \dots, \pi_{i_k}^m) && (4.3)
\end{aligned}$$

where (4.1) and (4.3) hold by definition of ϕ, and (4.2) holds since ξ is a clone homomorphism. □

The clonoid homomorphism that we built in the previous theorem might not be continuous; for this, we would need that for every sequence $(f_n)_{n \in \mathbb{N}}$ of polymorphisms of $\mathcal{A}$ converging to f, we have $\phi(f_n) = \phi(f)$ for every large enough n. This problem is easily solved in the case that $\overline{\mathrm{Aut}(\mathcal{B})} = \mathrm{End}(\mathcal{A})$.

Corollary 4.4. *Let $\mathcal{A}, \mathcal{B}$ be ω-categorical structures such that $\overline{\mathrm{Aut}(\mathcal{B})} = \mathrm{End}(\mathcal{A})$. Let $\mathscr{C}$ be the clone of polymorphisms of $\mathcal{A}$ that are canonical with respect to $\mathcal{B}$. Suppose that $\mathcal{B}$ has the mashup property and the canonisation property relative to $\mathcal{A}$. If there exists a clonoid homomorphism from $\mathscr{C}_1^{\text{typ}}$ to $\mathscr{P}$, then there exists a uniformly continuous clonoid homomorphism ϕ from $\mathrm{Pol}(\mathcal{A})$ to $\mathscr{P}$.*

Proof. If $\overline{\mathrm{Aut}(\mathcal{B})} = \mathrm{End}(\mathcal{A})$, then $\mathcal{A}$ is a first-order reduct of $\mathcal{B}$ and $\mathscr{C}_1^{\text{typ}}$ is idempotent so we can apply the previous theorem and obtain a clonoid homomorphism $\phi\colon \mathrm{Pol}(\mathcal{A}) \to \mathscr{P}$

that is constant on sets of the form $\overline{\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})}$ for $f \in \mathrm{Pol}(\mathcal{A})$. In particular, if $f \in \mathrm{Pol}(\mathcal{A})$ and $e \in \overline{\mathrm{Aut}(\mathcal{B})}$, we have $\phi(e \circ f) = \phi(f)$. Theorem 2.21 together with Corollary I.8 in [2] give that either there is a uniformly continuous clonoid homomorphism as in the statement, or there is a Siggers operation w in $\mathrm{Pol}(\mathcal{A})$ modulo $\overline{\mathrm{Aut}(\mathcal{B})}$. In the second case, $\phi(w)$ would be a Siggers operation in $\mathscr{P}$: indeed, suppose that $e_1, e_2 \in \overline{\mathrm{Aut}(\mathcal{B})}$ are such that

$$\forall x, y, z \in B,\ e_1 w(x, y, x, z, y, z) = e_2 w(y, x, z, x, z, y).$$

This property can also be written in the language of clones as

$$e_1 \circ w \circ (\pi_1^3, \pi_2^3, \pi_1^3, \pi_3^3, \pi_2^3, \pi_3^3) = e_2 \circ w \circ (\pi_2^3, \pi_1^3, \pi_3^3, \pi_1^3, \pi_3^3, \pi_2^3).$$

Applying ϕ on both sides of the equation, we obtain

$$\phi(w) \circ (\pi_1^3, \pi_2^3, \pi_1^3, \pi_3^3, \pi_2^3, \pi_3^3) = \phi(w) \circ (\pi_2^3, \pi_1^3, \pi_3^3, \pi_1^3, \pi_3^3, \pi_2^3),$$

as ϕ is a clonoid homomorphism such that $\phi(e_1 \circ w) = \phi(w) = \phi(e_2 \circ w)$. Whence, $\phi(w)$ is a Siggers operation in $\mathscr{P}$, a contradiction. So the first case must apply, i.e., there is a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{A}) \xrightarrow{u.c.c.h.} \mathscr{P}$. □

4.2 Disjoint Unions of Structures

Let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be structures with the same signature τ and with transitive automorphism groups (that is, $\mathcal{A}_i \big/ \mathrm{Aut}(\mathcal{A}_i)$ consists of one element, for every $i \in \{1, \ldots, n\}$). In this section, we define the disjoint union of $\mathcal{A}_1, \ldots, \mathcal{A}_n$ to be the structure $\mathcal{B}$ with domain $\biguplus_{i=1}^n A_i$ and with signature $\tau \cup \{A_1, \ldots, A_n\}$, where $R^{\mathcal{B}} = \bigcup_{i=1}^n R^{\mathcal{A}_i}$ for $R \in \tau$ and with $A_i^{\mathcal{B}} = A_i$ for all $i \in \{1, \ldots, n\}$.

Definition 4.4 (Local mashup)**.** Let $\mathcal{B}$ be a structure. Let $g, h, \omega \colon B^k \to B$, let $S \subseteq B$, and let $U, V \in \mathcal{B} \big/ \mathrm{Aut}(\mathcal{B})$ be two orbits. We say that ω is an S-mashup of g and h over $\{U, V\}$ iff the following holds: there exist $\alpha, \beta \in \mathrm{Aut}(\mathcal{B})$ such that for all $a_1, \ldots, a_k \in S$, we have

$$\omega(a_1, \ldots, a_k) = \begin{cases} \alpha g(a_1, \ldots, a_k) & \text{if } (a_1, \ldots, a_k) \in U \times V^{k-1} \\ \beta h(a_1, \ldots, a_k) & \text{if } (a_1, \ldots, a_k) \in V \times U^{k-1} \end{cases}$$

Proposition 4.5. *Let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be transitive ω-categorical structures with the same signature and let $\mathcal{B}$ be the disjoint union of $\mathcal{A}_1, \ldots, \mathcal{A}_n$. Suppose that $\mathcal{B}$ has the canonisation property. Let $f \colon A^k \to A$, and let g and h be canonical and in $\overline{\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})}$. Let $i, j \in \{1, \ldots, n\}$. There exists a canonical function ζ in $\overline{\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})}$ which is for every finite set $S \subset A$ an S-mashup of g and h over $\{A_i, A_j\}$.*

Proof. We first prove that for every finite subset S of B, there exists in $\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})$ an operation ω_S which is an S-mashup of g and h over $\{A_i, A_j\}$. Let $S \subset B$ be finite. Since g and h are in $\overline{\mathrm{Aut}(\mathcal{B}) f \,\mathrm{Aut}(\mathcal{B})}$, there exist operations $\alpha, \gamma, \beta_1, \delta_1, \ldots, \beta_k, \delta_k$ in $\mathrm{Aut}(\mathcal{B})$ such that

$$\begin{aligned} \forall a_1, \ldots, a_k \in S\ \big(& g(a_1, \ldots, a_n) = \gamma f(\delta_1 a_1, \ldots, \delta_k a_k) \\ \wedge\ & h(a_1, \ldots, a_n) = \alpha f(\beta_1 a_1, \ldots, \beta_k a_k) \big) \end{aligned}$$

Define $\omega_S \colon B^k \to B$ by

$$\omega_S(a_1, \dots, a_k) := f(\epsilon_1(a_1), \dots, \epsilon_k(a_k)),$$

where

$$\epsilon_1(a) = \begin{cases} \delta_1(a) & \text{if } a \in A_i \\ \beta_1(a) & \text{if } a \notin A_i \end{cases}$$

and

$$\epsilon_\ell(a) = \begin{cases} \beta_\ell(a) & \text{if } a \in A_i \\ \delta_\ell(a) & \text{if } a \notin A_i \end{cases}$$

for $\ell > 1$. It is easy to check that ϵ_ℓ is an element of $\mathrm{Aut}(\mathcal{B})$, for every $\ell \in \{1, \dots, k\}$. This immediately gives

$$\omega_S(a_1, \dots, a_k) = \begin{cases} \gamma^{-1}(g(a_1, \dots, a_k)) & \overline{a} \in A_i \times (A_j)^{k-1} \\ \alpha^{-1}(h(a_1, \dots, a_k)) & \overline{a} \in A_j \times (A_i)^{k-1} \end{cases}$$

Thus, $\omega_S \in \mathrm{Aut}(\mathcal{B}) f \, \mathrm{Aut}(\mathcal{B})$ is an S-mashup of g and h over $\{A_i, A_j\}$.

We now prove that there exists a single operation which is an S-mashup for all finite $S \subset B$. Let $0, 1, \dots$ be an enumeration of B. For each positive integer m, consider the equivalence relation on functions $\{0, \dots, m\}^k \to B$ defined by $r \sim_m s$ iff there exists $\alpha \in \mathrm{Aut}(\mathcal{B})$ such that $r = \alpha \circ s$ (i.e., the same equivalence relation as in Proposition 2.7). For each $m \geq 0$, this relation has finite index because the action of $\mathrm{Aut}(\mathcal{B})$ on B is oligomorphic. Consider the following forest $\mathcal{F}$. For each $m \geq 0$ and each operation ω which is an $\{0, \dots, m\}$-mashup of g and h in $\mathrm{Aut}(\mathcal{B}) f \, \mathrm{Aut}(\mathcal{B})$, the forest $\mathcal{F}$ contains the vertex $(\omega|_{\{0,\dots,m\}})/{\sim_m}$. For each $m \geq 1$, if $r/{\sim_m}$ is a vertex of $\mathcal{F}$, then there is an edge $\{s/{\sim_{m-1}}, r/{\sim_m}\}$ where $s = r|_{\{0,\dots,m-1\}}$. By the first paragraph, there are infinitely many vertices in $\mathcal{F}$. Since $\sim_m$ has finite index for all $m \geq 0$, the forest is finitely branching, and has finitely many roots. By König's lemma, there exists an infinite branch in $\mathcal{F}$, which we denote by $(\omega_m/{\sim_m})_{m \geq 0}$.

We now construct a chain of functions $\zeta_m \colon \{0, \dots, m\}^k \to B$ such that $\zeta_m \subset \zeta_{m+1}$ for all $m \geq 0$, and such that ζ_m is $\sim_m$-equivalent to ω_m. For $m = 0$, take $\zeta_0 = \omega_0$. Suppose that $m > 0$ and that ζ_{m-1} is defined. There is an edge between ω_{m-1} and ω_m by hypothesis and $\zeta_{m-1} \sim_{m-1} \omega_{m-1}$, which means that there is α in $\mathrm{Aut}(\mathcal{B})$ such that $\alpha\omega_m|_{\{0,\dots,m-1\}} = \zeta_{m-1}$. Define ζ_m to be $\alpha\omega_m$. We have $\zeta_{m-1} = \zeta_m|_{\{0,\dots,m-1\}}$ and $\zeta_m \sim_m \omega_m$, as required. Let now $\zeta = \bigcup_{m \geq 0} \zeta_m$.

It remains to prove that ζ is an S-mashup of g, h for every finite $S \subset B$. Let S be such a finite set, and m be such that $m \geq \max(S)$. Since $\omega_m/{\sim_m}$ is an element of $\mathcal{F}$, there exists $\omega \in \mathrm{Aut}(\mathcal{B}) f \, \mathrm{Aut}(\mathcal{B})$ that is an $\{0, \dots, m\}$-mashup of g and h and such that $\omega|_{\{0,\dots,m\}} = \omega_m$. Let U, V be orbits of B. Let α, β be the elements in $\mathrm{Aut}(\mathcal{B})$ witnessing that ω is an $\{0, \dots, m\}$-mashup of g and h. Let $\gamma \in \mathrm{Aut}(\mathcal{B})$ be such that $\omega_m = \gamma\zeta_m$. Then

we have for all $a_1, \dots, a_k \in \{0, \dots, m\}$:

$$\begin{aligned}\zeta(a_1, \dots, a_k) &= \zeta_m(a_1, \dots, a_k) \\ &= \gamma\omega_m(a_1, \dots, a_k) \\ &= \gamma\omega(a_1, \dots, a_k) \\ &= \begin{cases} \gamma\alpha g(a_1, \dots, a_k) & \text{if } (a_1, \dots, a_k) \in U \times V^{k-1} \\ \gamma\beta h(a_1, \dots, a_k) & \text{if } (a_1, \dots, a_k) \in V \times U^{k-1} \end{cases}.\end{aligned}$$

Therefore ζ is an S-mashup of g, h, with $\gamma \circ \alpha$ and $\gamma \circ \beta$ as witnesses.

Let ζ' be canonical and in $\overline{\mathrm{Aut}(\mathcal{B})\zeta\,\mathrm{Aut}(\mathcal{B})}$, which exists by the canonisation property for $\mathrm{Aut}(\mathcal{B})$. It is immediate that ζ' is an S-mashup of g, h for every finite S. Moreover, ζ' is in $\overline{\mathrm{Aut}(\mathcal{B})\zeta\,\mathrm{Aut}(\mathcal{B})}$ and we have

$$\overline{\mathrm{Aut}(\mathcal{B})\zeta\,\mathrm{Aut}(\mathcal{B})} \subseteq \overline{\mathrm{Aut}(\mathcal{B})\{\omega_m \colon m \geq 0\}\,\mathrm{Aut}(\mathcal{B})} \subseteq \overline{\mathrm{Aut}(\mathcal{B})f\,\mathrm{Aut}(\mathcal{B})},$$

so that ζ' is in $\overline{\mathrm{Aut}(\mathcal{B})f\,\mathrm{Aut}(\mathcal{B})}$ as required. □

Proposition 4.6 (Building Mashups)**.** *Let $\mathcal{B}$ be a structure. Let $g, h \colon B^k \to B$ be canonical with respect to $\mathcal{B}$ and let $U, V \in \mathcal{B}/\mathrm{Aut}(\mathcal{B})$. Suppose that ω is canonical with respect to $\mathcal{B}$ and is an S-mashup of g, h over $\{U, V\}$ for every finite $S \subset B$. Then $\xi_1^{\mathrm{typ}}(\omega)$ is a mashup of $\xi_1^{\mathrm{typ}}(g)$ and $\xi_1^{\mathrm{typ}}(h)$ over $\{U, V\}$.*

Proof. Let $a \in U, b \in V$. Then by definition $\xi_1^{\mathrm{typ}}(\omega)(U, V, \dots, V)$ is the orbit of $\omega(a, b, \dots, b)$ under $\mathrm{Aut}(\mathcal{B})$. Since ω is by assumption an $\{a, b\}$-mashup of g and h, there exists an $\alpha \in \mathrm{Aut}(\mathcal{B})$ such that $\omega(a, b, \dots, b) = \alpha g(a, b, \dots, b)$. Hence,

$$\xi_1^{\mathrm{typ}}(\omega)(U, V, \dots, V) = \xi_1^{\mathrm{typ}}(g)(U, V, \dots, V).$$

We can prove similarly that

$$\xi_1^{\mathrm{typ}}(\omega)(V, U, \dots, U) = \xi_1^{\mathrm{typ}}(h)(V, U, \dots, U),$$

so that $\xi_1^{\mathrm{typ}}(\omega)$ is indeed a mashup of $\xi_1^{\mathrm{typ}}(g)$ and $\xi_1^{\mathrm{typ}}(h)$ over $\{U, V\}$. □

Corollary 4.7. *Let $\mathcal{A}_1, \dots, \mathcal{A}_n$ be transitive ω-categorical structures and let $\mathcal{B}$ be their disjoint union. If $\mathcal{B}$ has the canonisation property, then it has the mashup property.*

4.3 Reducts of Unary Structures

In this section we study finite-signature reducts of unary structures, i.e., we study structures $\mathcal{A}$ for which there exist subsets $U_1, \dots, U_n$ of the domain A such that the relations of $\mathcal{A}$ are first-order definable in $(A; U_1, \dots, U_n)$. We obtain a P/NP-complete dichotomy for the CSPs of reducts of unary structures, and the border between tractability and intractability agrees with the conjectured border of Conjecture 1.

Theorem 4.8. *Let $\mathcal{A}$ be a finite-signature reduct of a unary structure. Then $\mathrm{CSP}(\mathcal{A})$ is in P if the model-complete core $\mathcal{B}$ of $\mathcal{A}$ has a Siggers polymorphism modulo endomorphisms of $\mathcal{B}$, and is NP-complete otherwise.*

Without changing the class of structures that we are studying we can assume that $\{U_1, \dots, U_n\}$ forms a partition of A, and that each U_i is either infinite or a singleton $\{a\}$ for some $a \in A$. We call such a partition a *stabilised partition*. Our claim above is then that for arbitrary subsets $U_1, \dots, U_n$ of A, there exists a stabilised partition $V_1, \dots, V_m$ of A such that the structure $(A; U_1, \dots, U_n)$ is first-order definable in $(A; V_1, \dots, V_m)$.

4.3.1 The case of tame endomorphisms

We start by investigating reducts of unary structures whose endomorphisms are precisely the injective operations that preserve the sets of the partition. In particular, such structures are model-complete cores. The milestone of this section is Theorem 4.9.

Theorem 4.9. *Let $\{U_1, \dots, U_n\}$ be a stabilised partition of A. Let $\mathcal{A}$ be a reduct of $(A; U_1, \dots, U_n)$ such that $\mathrm{End}(\mathcal{A})$ is the set of injective operations that preserve $U_1, \dots, U_n$. Let $\mathscr{C}$ be the clone of polymorphisms of $\mathcal{A}$ that are canonical with respect to $(A; U_1, \dots, U_n)$. Then the following are equivalent.*

1. *there is no continuous clone homomorphism from $\mathscr{C}$ to $\mathscr{P}$;*
2. *there is no uniformly continuous clonoid homomorphism from $\mathrm{Pol}(\mathcal{A})$ to $\mathscr{P}$;*
3. *$\mathcal{A}$ has a cyclic (Siggers, weak near-unanimity) polymorphism modulo endomorphisms of $\mathcal{A}$;*
4. *$\mathcal{A}$ has a cyclic (Siggers, weak near-unanimity) polymorphism f modulo endomorphisms of $\mathcal{A}$ and f is canonical with respect to $\mathrm{Aut}(A; U_1, \dots, U_n)$.*

The proof of the theorem will be given at the end of this subsection. For now, we simply remark that the implications (1) $\Rightarrow$ (4) $\Rightarrow$ (3) $\Rightarrow$ (2) are either trivial or immediate corollaries of statements from the literature. We prove the implication from 2 to 1 by contraposition.

Let $\{U_1, \dots, U_n\}$ be a stabilised partition of A, and let $\mathcal{A}$ be a first-order reduct of $(A; U_1, \dots, U_n)$ such that $\mathrm{End}(\mathcal{A})$ is precisely the set of injective operations that preserve $U_1, \dots, U_n$. Let $\mathscr{C}$ be the subclone of $\mathrm{Pol}(\mathcal{A})$ that consists of the functions that are canonical with respect to $(A; U_1, \dots, U_n)$. Note that $\mathrm{Aut}(A; U_1, \dots, U_n)$ is dense in $\mathrm{End}(\mathcal{A})$, so that for every $i \in \{1, \dots, n\}$ the map that takes $f \in \mathrm{Pol}(\mathcal{A})$ to $f|_{U_i}$ is well-defined and is a continuous clone homomorphism: the restriction of some projection π_i^n remains the same projection, and $f|_{U_i} \circ (g_1|_{U_i}, \dots, g_k|_{U_i}) = (f \circ (g_1, \dots, g_k))|_{U_i}$ holds; the image of this clone homomorphism is a function clone $\mathrm{Pol}(\mathcal{A})_{U_i}$ over the set U_i. We show in the next two propositions that one of the following holds: there exists some $i \in \{1, \dots, n\}$ such that $\mathrm{Pol}(\mathcal{A})_{U_i} \to \mathscr{P}$, or $\mathscr{C}_1^{\mathrm{typ}} \to \mathscr{P}$, or $\mathscr{C}_2^{\mathrm{typ}}$ contains a cyclic operation.

Clearly, every permutation of U_i is an operation in $\mathrm{Pol}(\mathcal{A})_{U_i}$. Such clones have been studied in [21] in the context of constraint satisfaction problems. In particular, the authors show the following.

Theorem 4.10 (Consequence of Theorem 7 in [21])**.** *Let $\mathscr{A}$ be a closed clone over a countably infinite set A containing $\mathrm{Sym}(A)$. Then $\mathscr{A}$ has a continuous homomorphism to $\mathscr{P}$ if and only if there is no constant unary and no injective binary operation in $\mathscr{A}$.*

We say that an operation $f\colon A^k \to A$ is *injective in its ith argument* if $f(\overline{a}) \neq f(\overline{b})$ for all tuples $\overline{a}, \overline{b}$ with $a_i \neq b_i$.

Proposition 4.11. *Let A be an infinite set and let $f\colon A^k \to A$ be a function that is canonical with respect to $(A, =)$. Either f is a constant function, or there is an $i \in \{1, \ldots, k\}$ such that f is injective in its ith argument.*

Proof. For two tuples $\overline{a}, \overline{b}$, let $I_{\overline{a},\overline{b}} = \{j \in \{1, \ldots, k\} \mid a_j \neq b_j\}$. By canonicity of f, if $\overline{a}, \overline{b}, \overline{c}, \overline{d}$ are such that $I_{\overline{a},\overline{b}} = I_{\overline{c},\overline{d}}$, then $f(\overline{a}) = f(\overline{b})$ if and only if $f(\overline{c}) = f(\overline{d})$. Suppose that the second case of the statement does not apply. That is, for all $i \in \{1, \ldots, k\}$, there are tuples $\overline{a}, \overline{b}$ with $f(\overline{a}) = f(\overline{b})$ and $i \in I_{\overline{a},\overline{b}}$. We prove that for all $i \in \{1, \ldots, k\}$, there are tuples $\overline{c}, \overline{d}$ such that $f(\overline{c}) = f(\overline{d})$ and $I_{\overline{c},\overline{d}} = \{i\}$. Let $i \in \{1, \ldots, k\}$ be arbitrary. Pick $\overline{a}, \overline{b}$ such that $f(\overline{a}) = f(\overline{b})$ and such that $I_{\overline{a},\overline{b}}$ is minimal with the property that $i \in I_{\overline{a},\overline{b}}$.

Suppose for contradiction that $|I_{\overline{a},\overline{b}}| > 1$. Let $i' \in I_{\overline{a},\overline{b}} \setminus \{i\}$. Let $c_1, \ldots, c_k \in A$ be elements such that:

- $c_{i'} = a_{i'}$,
- $c_j = a_j = b_j$ for $j \in \{1, \ldots, k\} \setminus I_{\overline{a},\overline{b}}$,
- $c_j \notin \{a_j, b_j\}$ for all $j \in I_{\overline{a},\overline{b}} \setminus \{i'\}$.

Note that $I_{\overline{b},\overline{c}} = I_{\overline{a},\overline{b}}$, that $i \in I_{\overline{a},\overline{c}}$, and that $I_{\overline{a},\overline{c}} \subset I_{\overline{a},\overline{b}}$. The first equality implies that $f(\overline{b}) = f(\overline{c})$, by canonicity of f. Therefore, $f(\overline{a}) = f(\overline{c})$. This contradicts the minimality assumption on $I_{\overline{a},\overline{b}}$, which proves the claim.

We can now prove that f is constant. Let $\overline{a}, \overline{b}$ be arbitrary k-tuples. For $i \in \{0, \ldots, k\}$, define $\overline{c}^i$ to be the tuple $(b_1, \ldots, b_i, a_{i+1}, \ldots, a_k)$. For all $i \in \{0, \ldots, k-1\}$, we have $I_{\overline{c}^i,\overline{c}^{i+1}} = \{i+1\}$, so that by the claim above, we have $f(\overline{c}^i) = f(\overline{c}^{i+1})$. Note that $\overline{c}^0 = \overline{a}$, and $\overline{c}^k = \overline{b}$, so that $f(\overline{a}) = f(\overline{b})$. □

Proposition 4.12. *Let $U_1, \ldots, U_n$ be a stabilised partition of a set A. Every $f\colon A^k \to A$ interpolates modulo $\mathrm{Aut}(A; U_1, \ldots, U_n)$ an operation $g\colon A^k \to A$ that is canonical with respect to $(A; U_1, \ldots, U_n)$.*

Proof. Let $\prec$ be any linear order on A such that if $u \in U_i, v \in U_j$ and $i < j$, then $u \prec v$, and such that $\prec$ is dense and without endpoints on U_i whenever U_i is infinite. The group $\mathrm{Aut}(A; U_1, \ldots, U_n, \prec)$ is extremely amenable (this is a corollary of the fact that extreme amenability is preserved under direct products, and that the automorphism group of a countable dense linear order is extremely amenable [82]). It follows from Theorem 1 in [36] that there exists a g which satisfies the conclusion of the lemma, except that g is canonical with respect to $(A; U_1, \ldots, U_n, \prec)$.

We prove that g is also canonical with respect to $(A; U_1, \ldots, U_n)$. Since the structure $(A; U_1, \ldots, U_n)$ is homogeneous and ω-categorical, the orbits of tuples in $(A; U_1, \ldots, U_n)$ can be defined by quantifier-free formulas without disjunctions (Theorem 2.8 and Proposition 2.9). Since the signature of $(A; U_1, \ldots, U_n)$ (including the equality relation) is binary, these quantifier-free formulas can be taken to be conjunctions of binary formulas. This implies that g is canonical if and only if for all *pairs* $(a_1, b_1), \ldots, (a_k, b_k)$, $(c_1, d_1), \ldots, (c_k, d_k)$ such that (a_j, b_j) is in the same orbit as (c_j, d_j) for all $j \in \{1, \ldots, k\}$, we have that $(g(a_1, \ldots, a_k), g(b_1, \ldots, b_k))$ and $(g(c_1, \ldots, c_k), g(d_1, \ldots, d_k))$ are in the same orbit under $\mathrm{Aut}(A; U_1, \ldots, U_n)$. Note that $\mathrm{Aut}(A; U_1, \ldots, U_n)$ satisfies the following property:

(†) two pairs $(a,b),(c,d)$ are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n)$
iff a and c are in the same orbit, b and d are in the same orbit, and $a=b$ iff $c=d$.

Let $\overline{a},\overline{b},\overline{c},\overline{d}\in A^k$ be such that for all $i\in\{1,\dots,k\}$, the pairs (a_i,b_i) and (c_i,d_i) are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n)$. We first prove that $g(\overline{a})$ and $g(\overline{c})$ are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n)$. Using property (†) we know that a_i and c_i are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n)$. It follows that they are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n,\prec)$ (because if a_i and c_i belong to one of the finite sets of the partition, they must be equal from the assumption that $\{U_1,\dots,U_n\}$ is a stabilised partition). Since g is known to be canonical with respect to $(A;U_1,\dots,U_n,\prec)$, we have that $g(\overline{a})$ and $g(\overline{c})$ are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n,\prec)$, and therefore they are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n)$. Similarly we obtain that $g(\overline{b})$ and $g(\overline{d})$ are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n)$.

Therefore, it remains to check that $g(\overline{a})$ is equal to $g(\overline{b})$ iff $g(\overline{c})$ equals $g(\overline{d})$. Suppose that $g(\overline{a})=g(\overline{b})$. Let $i\in\{1,\dots,n\}$ be such that all of $g(\overline{a}),g(\overline{b}),g(\overline{c}),g(\overline{d})$ are in U_i. If U_i is finite, we have that $g(\overline{a})=g(\overline{c})$ and $g(\overline{b})=g(\overline{d})$, so the property is true. Assume now that U_i is infinite. For $j\in\{1,\dots,k\}$, let $e_j\in A$ be such that either:

- $a_j\prec b_j$ and e_i is taken to be in U_i and larger than c_j and d_j,
- $b_j\prec a_j$ and e_i is taken to be in U_i and smaller than c_j and d_j, or
- $a_j=b_j$ and $e_j=c_j$.

Note that if $a_j=b_j$ then $c_j=d_j$ so $e_j=c_j=d_j$. By definition, $(a_j,b_j),(c_j,e_j)$, and (d_j,e_j) all are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n,\prec)$ for all $j\in\{1,\dots,k\}$. We therefore have that $(g(\overline{a}),g(\overline{b})),(g(\overline{c}),g(\overline{e}))$, and $(g(\overline{d}),g(\overline{e}))$ are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n,\prec)$, whence they are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n)$. Thus, if $g(\overline{a})=g(\overline{b})$ then $g(\overline{c})=g(\overline{e})$ and $g(\overline{d})=g(\overline{e})$. In particular, we have $g(\overline{c})=g(\overline{d})$. □

Proposition 4.13. *Let $U_1,\dots,U_n$ be a stabilised partition of A and let $\mathcal{A}$ be a reduct of $(A;U_1,\dots,U_n)$ whose endomorphisms are the injective functions preserving $U_1,\dots,U_n$. Let $\mathscr{C}$ be the subclone of $\mathrm{Pol}(\mathcal{A})$ consisting of the functions that are canonical with respect to $(A;U_1,\dots,U_n)$. Suppose that neither $\mathscr{C}_1^{\mathrm{typ}}$ nor any $\mathrm{Pol}(\mathcal{A})_{U_i}$ has a continuous homomorphism to $\mathscr{P}$. Then $\mathscr{C}_2^{\mathrm{typ}}$ contains a cyclic operation.*

Proof. Note that $\mathscr{C}_1^{\mathrm{typ}}$ is idempotent since $\mathrm{End}(\mathcal{A})=\overline{\mathrm{Aut}(A;U_1,\dots,U_n)}$. Since $\mathscr{C}_1^{\mathrm{typ}}$ does not have a homomorphism to $\mathscr{P}$, Theorem 2.17 implies that there exists an operation $c\in\mathscr{C}$ of arity $k\geq 2$ such that $\xi_1^{\mathrm{typ}}(c)$ is cyclic in $\mathscr{C}_1^{\mathrm{typ}}$. For every $i\in\{1,\dots,n\}$, by assumption $\mathrm{Pol}(\mathcal{A})_{U_i}$ does not have a homomorphism to $\mathscr{P}$ and since all the functions in $\mathrm{End}(\mathcal{A})$ are injective, $\mathrm{Pol}(\mathcal{A})_{U_i}$ cannot contain a unary constant function. By Theorem 4.10, there exists a binary operation in $\mathrm{Pol}(\mathcal{A})$ that is injective when restricted to U_i (if U_i is finite it is a singleton by assumption, so such an operation also exists in this case). One sees that such a binary operation generates a k-ary operation whose restriction to U_i is again injective. Finally, by Proposition 4.12, this operation interpolates modulo $\mathrm{Aut}(A;U_1,\dots,U_n)$ a canonical function $g_i\in\mathscr{C}$ of arity k, which is still injective on U_i.

We prove by induction on m, with $1\leq m\leq n$, that there exists in $\mathscr{C}$ an operation g which is injective on $\bigcup_{i=1}^m (U_i)^k$, the case $m=1$ being dealt with by the paragraph above.

So assume that the operation g' is in $\mathscr{C}$ and is injective on $\bigcup_{i=1}^{m-1}(U_i)^k$. Define a new operation g by

$$g(x_1,\dots,x_k) := g_m(g'(\overline{x}), g'(\sigma\overline{x}),\dots,g'(\sigma^{k-1}\overline{x})),$$

where σ is the permutation $(x_1,\dots,x_k)\mapsto(x_2,\dots,x_k,x_1)$ and g_m is the k-ary canonical function whose existence is asserted in the previous paragraph. Since $\mathrm{Aut}(A;U_1,\dots,U_n)$ is dense in $\mathrm{End}(\mathcal{A})$, it is clear that if $\overline{x}\in(U_i)^k$ and $\overline{y}\in(U_j)^k$ for $i\neq j$, then $g(\overline{x})\neq g(\overline{y})$. If $\overline{x},\overline{y}\in(U_i)^k$ are two different tuples with $i\leq m-1$, we have for all j that $g'(\sigma^j\overline{x})\neq g'(\sigma^j\overline{y})$. Since g_m is canonical and there is no constant operation in $\mathrm{Pol}(\mathcal{A})$, this operation is injective in one of its arguments, by Proposition 4.11. It follows that $g(\overline{x})\neq g(\overline{y})$. If $\overline{x},\overline{y}\in(U_m)^k$, a similar argument works: since g' is canonical and non-constant, it is injective in at least one of its arguments by Proposition 4.11. Whence, for at least one $j\in\{0,\dots,k-1\}$ we have that $g'(\sigma^j\overline{x})\neq g'(\sigma^j\overline{y})$, and by injectivity of g_m on $(U_m)^k$, we obtain $g(\overline{x})\neq g(\overline{y})$. It follows that g is canonical and injective on $\bigcup_{i=1}^{m}(U_i)^k$ as required.

Define now $c'\in\mathscr{C}$ by

$$c'(\overline{x}) := g(c(\overline{x}), c(\sigma\overline{x}),\dots,c(\sigma^{k-1}\overline{x})),$$

where g is the operation built in the previous paragraph. We claim that $\xi_2^{\mathrm{typ}}(c')\in\mathscr{C}_2^{\mathrm{typ}}$ is cyclic. It is trivial to check that $\xi_1^{\mathrm{typ}}(c')$ is cyclic in $\mathscr{C}_1^{\mathrm{typ}}$. We now show that for all k-tuples $\overline{a},\overline{b}$ such that $c'(\overline{a})=c'(\overline{b})$ we have $c'(\sigma\overline{a})=c'(\sigma\overline{b})$. Suppose that $\overline{a}$ and $\overline{b}$ are given and map to the same point under c'. This means that

$$g(c(\overline{a}), c(\sigma\overline{a}),\dots,c(\sigma^{k-1}\overline{a})) = g(c(\overline{b}), c(\sigma\overline{b}),\dots,c(\sigma^{k-1}\overline{b})) \tag{4.4}$$

holds. Note that $(c(\overline{a}), c(\sigma\overline{a}),\dots,c(\sigma^{k-1}\overline{a}))$ and $g(c(\overline{b}), c(\sigma\overline{b}),\dots,c(\sigma^{k-1}\overline{b}))$ are both tuples in $\bigcup_{i=1}^{n}(U_i)^k$. By injectivity of g on this set, we therefore get that for all $j\in\{0,\dots,k-1\}$, the equality $c(\sigma^j\overline{a})=c(\sigma^j\overline{b})$ holds. By injecting this back into equation (4.4), we conclude that $c'(\sigma\overline{a})=c'(\sigma\overline{b})$.

To show that $\xi_2^{\mathrm{typ}}(c')$ is cyclic, let $(a_1,b_1),\dots,(a_k,b_k)$ be pairs of elements of A. We have to show that $(c'(\overline{a}),c'(\overline{b}))$ and $(c'(\sigma\overline{a}),c'(\sigma\overline{b}))$ are in the same orbit under G. Since $\xi_1^{\mathrm{typ}}(c')$ is cyclic, we already know that $c'(\overline{a})$ and $c'(\sigma\overline{a})$ are in the same orbit, and that $c'(\overline{b})$ and $c'(\sigma\overline{b})$ are in the same orbit. Recall that $\mathrm{Aut}(A;U_1,\dots,U_n)$ satisfies the following property: two pairs $(a,b),(c,d)$ are in the same orbit under $\mathrm{Aut}(A;U_1,\dots,U_n)$ iff a,c are in the same orbit, b,d are in the same orbit and $a=b$ iff $c=d$. So we only need to check that $c'(\overline{a})=c'(\overline{b})$ iff $c'(\sigma\overline{a})=c'(\sigma\overline{b})$. In the left-to-right direction, this is what we proved above. For the other direction, note that we can apply $k-1$ times the argument of the previous paragraph to obtain that if $c'(\sigma\overline{a})=c'(\sigma\overline{b})$, then $c'(\sigma^k(\overline{a}))=c'(\sigma^k(\overline{b}))$, i.e., $c'(\overline{a})=c'(\overline{b})$. □

Recall that we want to prove the implication (2) ⇒ (1) of Theorem 4.9 by contraposition, that is, that if there is a continuous clone homomorphism from the canonical subclone $\mathscr{C}$ of $\mathrm{Pol}(\mathcal{A})$ to $\mathscr{P}$, then there is a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{A})\xrightarrow{u.c.c.h.}\mathscr{P}$. The assumption implies that $\mathscr{C}_2^{\mathrm{typ}}$ does not contain a cyclic operation. The previous proposition implies that $\mathrm{Pol}(\mathcal{A})_{U_i}$ has a continuous clone homomorphism to $\mathscr{P}$, or that $\mathscr{C}_1^{\mathrm{typ}}$ has a continuous homomorphism to $\mathscr{P}$. In the first case, we immediately obtain a continuous clone homomorphism $\mathrm{Pol}(\mathcal{A})\to\mathscr{P}$. In the second case, we apply Corollary 4.4. In order to do so, we need to prove that the automorphism group

of the structure $(A; U_1, \dots, U_n)$ has the canonisation property and the mashup property. Note that $(A; U_1, \dots, U_n)$ is the disjoint union of n times the structure $(A; =)$, and that $(A; U_1, \dots, U_n)$ has the canonisation property by Proposition 4.12. It follows from Corollary 4.7 that $(A; U_1, \dots, U_n)$ has the mashup property.

Proof of Theorem 4.9. We prove the implications $1 \Rightarrow 4 \Rightarrow 3 \Rightarrow 2 \Rightarrow 1$. Let $\mathscr{D}$ be the clone of polymorphisms of $\mathcal{A}$; so $\mathscr{C}$ is the subclone of $\mathscr{D}$ consisting of the operations that are canonical with respect to $(A; U_1, \dots, U_n)$.

Suppose that 1 holds, that is, there is no continuous clone homomorphism from $\mathscr{C}$ to $\mathscr{P}$. The finite clone $\mathscr{C}_2^{\text{typ}}$ is idempotent, by the assumption that $\overline{\text{Aut}(A; U_1, \dots, U_n)} = \text{End}(\mathcal{A})$. By Corollary 3.7, $\mathscr{C}$ contains a pseudo-cyclic operation modulo $\text{End}(\mathcal{A})$. This proves 4.

The implication $4 \Rightarrow 3$ is trivial.

Suppose now that $\mathscr{D}$ contains a pseudo-Siggers polymorphism. Since $\mathcal{A}$ is a model-complete core, Theorem 2.21 implies that there is no uniformly continuous clonoid homomorphism $\text{Pol}(\mathcal{A}) \xrightarrow{u.c.c.h.} \mathscr{P}$, proving 2.

It remains to prove that 2 implies 1. By contraposition, let us suppose that 1 does not hold. Thus, there is a continuous clone homomorphism from $\mathscr{C}$ to $\mathscr{P}$. In particular, $\mathscr{C}$ does not contain any pseudo-cyclic operation and by Proposition prop:equations-lift, $\mathscr{C}_2^{\text{typ}}$ does not contain any cyclic operation. By Proposition 4.13, either there exists a continuous clone homomorphism $\mathscr{D}_{U_i} \to \mathscr{P}$ for some $i \in \{1, \dots, n\}$, or there is a clone homomorphism $\mathscr{C}_1^{\text{typ}} \to \mathscr{P}$. In the first case we are done: we obtain by composing with $\mathscr{D} \to \mathscr{D}_{U_i}$ a continuous clone homomorphism $\mathscr{D} \to \mathscr{P}$, so 2 does not hold. Suppose we are in the second case. Proposition 4.12 implies that $\text{Aut}(A; U_1, \dots, U_n)$ has the canonisation property. By Corollary 4.7, it also has the mashup property. Corollary 4.4 implies that there exist a uniformly continuous clone homomorphism from $\mathscr{D}$ to $\mathscr{P}$. This shows that 2 does not hold in this case either, and concludes the proof of $2 \Rightarrow 1$. □

4.3.2 The general case

In this section we conclude the proof of the dichotomy theorem for reducts $\mathcal{A}$ of unary structures $(A; U_1, \dots, U_n)$. The previous section treated the special case where $\text{End}(\mathcal{A})$ consists exactly of the injective operations preserving $U_1, \dots, U_n$. In the following, we reduce the general case to this situation.

The first step of the strategy for this is to show that we can assume without loss of generality that $\mathcal{A}$ is a model-complete core. Since reducts of unary structures are ω-categorical, and since every ω-categorical structure has a model-complete core, it suffices to prove that the model-complete core of a reduct of a unary structure is again a reduct of a unary structure (Lemma 4.14 below). The second step is to show that by adding constants in a suitable way, we obtain a reduct of a unary structure which satisfies the hypothesis of the previous section (Proposition 4.15).

Lemma 4.14. *Let $\mathcal{A}$ be a reduct of a unary structure, and let $\mathcal{B}$ be the model-complete core of $\mathcal{A}$. Then $\mathcal{B}$ is a reduct of a unary structure.*

Proof. Let $\mathcal{A}$ be a reduct of $(A; U_1, \dots, U_n)$. Suppose without loss of generality that $\mathcal{B}$ is a substructure of $\mathcal{A}$. Let h be a homomorphism from $\mathcal{A}$ to $\mathcal{B}$. We show that $\mathcal{B}$ is a reduct of $(B; U_1 \cap B, \dots, U_n \cap B)$. To this end, we prove that every permutation of B preserving the sets $U_1 \cap B, \dots, U_n \cap B$ is an automorphism of $\mathcal{B}$. Let β be such a permutation. Then β

can be extended by the identity to a permutation α of A which preserves $U_1, \dots, U_n$, and therefore α is an automorphism of $\mathcal{A}$. Thus, $h \circ \beta = h \circ \alpha|_B \colon \mathcal{B} \to \mathcal{B}$ is an endomorphism of $\mathcal{B}$, and so an embedding since $\mathcal{B}$ is a model-complete core. This implies that β is an embedding, i.e., it is an automorphism of $\mathcal{B}$. Note that $(B; U_1 \cap B, \dots, U_n \cap B)$ is ω-categorical. By Theorem 2.12, we obtain that all the relations of $\mathcal{B}$ are fo-definable in $(B; U_1 \cap B, \dots, U_n \cap B)$. $\square$

It can be the case that $\mathrm{End}(\mathcal{A})$ contains more operations than the injections preserving $U_1, \dots, U_n$ even when $\mathcal{A}$ is a reduct of $(A; U_1, \dots, U_n)$ which is a model-complete core. An example is $(A; E, \neq)$ where $A = U_1 \uplus U_2$ and $E = \{(x, y) \in A^2 \mid x \in U_1 \Leftrightarrow y \in U_2\}$. However, for every such reduct there are finitely many constants $c_1, \dots, c_n \in A$ such that the $(\mathcal{A}, c_1, \dots, c_n)$ satisfies the condition of Theorem 4.9.

Proposition 4.15. *Let $\mathcal{A}$ be a reduct of a unary structure that is a model-complete core. There exist elements $c_1, \dots, c_n \in A$ and a stabilised partition $\{V_1, \dots, V_m\}$ of A such that $(\mathcal{A}, c_1, \dots, c_n)$ is a reduct of the unary structure $(A; V_1, \dots, V_m)$ and such that the endomorphisms of $(\mathcal{A}, c_1, \dots, c_n)$ are precisely the injective functions preserving $V_1, \dots, V_m$.*

Proof. Let $\{U_1, \dots, U_n\}$ be a stabilised partition of A where n is minimal with the property that $\mathcal{A}$ is a reduct of $(A; U_1, \dots, U_n)$. Up to a permutation of the blocks, we can assume that $U_1, \dots, U_r$ are the finite blocks of the partition. For every $i \in \{1, \dots, n\}$, let $c_i \in U_i$. We claim that

$$\mathrm{Aut}(\mathcal{A}, c_1, \dots, c_n) = \mathrm{Aut}(A; U_1, \dots, U_n, c_1, \dots, c_n).$$

If $r = n$, there is nothing to prove, because of the assumption that the sets $U_1, \dots, U_n$ are either singletons are infinite. Therefore, if $r = n$, we have $\mathrm{Aut}(\mathcal{A}, c_1, \dots, c_n) = \mathrm{Aut}(A; U_1, \dots, U_n)$.

We prove that $\mathrm{Aut}(\mathcal{A}, c_1, \dots, c_r)$ preserves the binary relation

$$E := \{(x, y) \in A^2 \mid \forall i \in \{r+1, \dots, n\}, x \in U_i \Leftrightarrow y \in U_i\}.$$

Let α be an automorphism of $\mathcal{A}$. For $i, j \in \{r+1, \dots, n\}$, define $V_{ij}(\alpha)$ to be the set of elements of U_i that are mapped to U_j under α.

<u>Claim 0:</u> for every $i \in \{r+1, \dots, n\}$ and every automorphism α of $\mathcal{A}$, there exists a $j \in \{r+1, \dots, n\}$ such that $V_{ji}(\alpha)$ is infinite.

Proof. Since α is a bijection, every element of U_i has a preimage under α. Since there are only finitely many sets in the partition, one of the sets U_j contains infinitely many of those preimages, i.e., $V_{ji}(\alpha)$ is infinite. $\Diamond$

<u>Claim 1:</u> for every $i \in \{r+1, \dots, n\}$ and every automorphism α of $\mathcal{A}$, the set $V_{ii}(\alpha)$ is either finite or U_i.

Proof. Let α be an automorphism of $\mathcal{A}$, and suppose that $\emptyset \neq V_{ii}(\alpha) \neq U_i$. Since $V_{ii}(\alpha) \neq U_i$, there exists a $j \in \{r+1, \dots, n\}$ such that $V_{ij}(\alpha) \neq \emptyset$, which is equivalent to say that $V_{ji}(\alpha^{-1}) \neq \emptyset$. Suppose for contradiction that $V_{ii}(\alpha)$ is infinite. Equivalently, $V_{ii}(\alpha^{-1})$ is infinite. We claim that for every finite subset S of A, there exists an automorphism α' of $\mathcal{A}$ such that $\alpha'(S) \cap U_j = \emptyset$. This is clear: let β be an automorphism of $\mathcal{A}$ that maps $S \cap U_i$ to $V_{ii}(\alpha^{-1})$ (which is possible since $V_{ii}(\alpha^{-1})$ is infinite) and one element of $S \cap U_j$

to $V_{ji}(\alpha^{-1})$. The automorphism $\alpha'_1 := \alpha^{-1} \circ \beta$ maps one point from $S \cap U_j$ to U_i, and maps all the elements of $S \cap U_i$ to U_i. Possibly, some elements in $S \cap U_k$ for $k \notin \{i, j\}$ are mapped by α'_1 to U_j. We repeat this procedure and obtain automorphisms $\alpha'_2, \ldots, \alpha'_m$ with $m \leq |S|$, until $\alpha'_m(S) \cap U_j$ is empty. Using a standard compactness argument, we obtain an operation $e \in \overline{\mathrm{Aut}(\mathcal{A})}$ whose image does not intersect U_j. This is a contradiction to the minimality of the partition $\{U_1, \ldots, U_n\}$: the structures $e(\mathcal{A})$ and $\mathcal{A}$ are isomorphic, and the relations of $e(\mathcal{A})$ are definable in $(A \setminus U_j; U_1, \ldots, U_{j-1}, U_{j+1}, \ldots, U_n)$. Therefore $V_{ii}(\alpha)$ is finite. $\Diamond$

<u>Claim 2:</u> for every $i \in \{r+1, \ldots, n\}$ and every automorphism α of $\mathcal{A}$, the set $V_{ii}(\alpha)$ is either empty or U_i.

Proof. Suppose that for some $\alpha \in \mathrm{Aut}(\mathcal{A})$, the set $V_{ii}(\alpha)$ is not equal to U_i and is not empty. We prove that for every $k \geq 1$, there exists an automorphism α_k of $\mathcal{A}$ such that $|V_{ii}(\alpha_k)| \geq k$ and such that α_k does not preserve U_i. Let $k \geq 1$. By Claim 0, there exists a $j \in \{r+1, \ldots, n\}$ such that $V_{ji}(\alpha)$ is infinite and by Claim 1, it must be the case that $j \neq i$. Note that $V_{ij}(\alpha^{-1})$ is infinite, and that $V_{ii}(\alpha^{-1})$ is not empty. Let $x_1, \ldots, x_k$ be pairwise distinct elements in $V_{ij}(\alpha^{-1})$, and let $y \in V_{ii}(\alpha^{-1})$. Let z be an element of U_i such that $\alpha(z) \notin U_i$, which exists since $V_{ii}(\alpha) \neq U_i$. Let β be an automorphism of $\mathcal{A}$ that maps $\alpha^{-1}(y)$ to z and which leaves $\alpha^{-1}(x_1), \ldots, \alpha^{-1}(x_k)$ fixed. Then $\alpha \circ \beta \circ \alpha^{-1}$ is an automorphism of $\mathcal{A}$ such that $x_1, \ldots, x_k \in V_{ii}(\alpha \circ \beta \circ \alpha^{-1})$ and such that $(\alpha \circ \beta \circ \alpha^{-1})(y) \notin U_i$.

For each $k \geq 1$, there exists by Claim 0 a $j \in \{r+1, \ldots, n\}$ such that $V_{ji}(\alpha_k)$ is infinite. Since α_k does not preserve U_i by assumption, $V_{ii}(\alpha_k) \neq U_i$. By Claim 1, $V_{ii}(\alpha_k)$ has to be finite, so j is distinct from i. By the pigeonhole principle, there is a $j \in \{r+1, \ldots, n\}$ distinct from i such that $V_{ji}(\alpha_k)$ is infinite for infinitely many k. Therefore, using another argument one can show that there is an endomorphism of $\mathcal{A}$ in $\overline{\langle \mathrm{Aut}(A; U_1, \ldots, U_n) \cup \{\alpha_k : k \geq 1\} \rangle}$ whose image does not intersect U_j, which is a contradiction to the minimality of the partition $\{U_1, \ldots, U_n\}$. Hence, $V_{ii}(\alpha)$ is either empty or U_i. $\Diamond$

<u>Claim 3:</u> for every $i \in \{r+1, \ldots, n\}$ and every automorphism α of $\mathcal{A}$, there is exactly one $j \in \{r+1, \ldots, n\}$ such that $V_{ij}(\alpha)$ is nonempty.

Proof. Suppose that $j, j' \in \{r+1, \ldots, n\}$ are distinct and that $V_{ij}(\alpha)$ and $V_{ij'}(\alpha)$ are both nonempty, say that $\alpha(x) \in U_j$ and $\alpha(y) \in U_{j'}$. Since $V_{jj}(\alpha^{-1})$ is not U_j, it must be empty by Claim 2. Thus, there exists a k distinct from j such that $V_{jk}(\alpha^{-1})$ is infinite, which gives the existence of a $z \in U_j$ distinct from $\alpha(x)$ such that $\alpha^{-1}(z) \in U_k$. Let β be an automorphism of $\mathcal{A}$ that maps x to y, and leaves $\alpha^{-1}(z)$ fixed. Then the map $\alpha \circ \beta \circ \alpha^{-1}$ maps $\alpha(x) \in U_j$ to $\alpha(y) \in U_{j'}$, and maps $z \in U_j$ to itself. Therefore, we have that $V_{jj}(\alpha \circ \beta \circ \alpha^{-1})$ is neither empty nor equal to U_j, a contradiction to the second claim. $\Diamond$

Therefore, the relation E is preserved by $\mathrm{Aut}(\mathcal{A}, c_1, \ldots, c_r)$. This implies that each of $U_1, \ldots, U_n$ is preserved by $\mathrm{Aut}(\mathcal{A}, c_1, \ldots, c_n)$. We obtain that $(\mathcal{A}, c_1, \ldots, c_n)$ is a reduct of $(A; U_1 \setminus \{c_1\}, \ldots, U_n \setminus \{c_n\}, \{c_1\}, \ldots, \{c_n\})$ whose endomorphisms are precisely the injective functions that preserve this stabilised partition. $\square$

Corollary 4.16. *Let $\mathcal{A}$ be a reduct of a unary structure. Then there exists an expansion $\mathcal{C}$ of the model-complete core of $\mathcal{A}$ by finitely many constants such that* $\mathrm{Pol}(\mathcal{C})$ *satisfies either 1. or 2.:*

1. *there is a uniformly continuous clonoid homomorphism* $\mathrm{Pol}(\mathcal{C}) \xrightarrow{u.c.c.h.} \mathscr{P}$*;*
2. $\mathrm{Pol}(\mathcal{C})$ *contains a cyclic (equivalently: a Siggers, or a weak near-unanimity) operation f modulo* $\mathrm{End}(\mathcal{C})$*; moreover, f is canonical with respect to $\mathcal{C}$.*

Proof. Let $U_1, \ldots, U_n$ be a partition of A such that $\mathcal{A}$ is a reduct of $(A; U_1, \ldots, U_n)$. If the model-complete core $\mathcal{B}$ of $\mathcal{A}$ is finite, then we can expand by a constant for each element of $\mathcal{B}$, and the statement follows from Theorem 2.17. Otherwise, for some stabilised partition $\{V_1, \ldots, V_m\}$ of B the structure $\mathcal{B}$ is a reduct of $(B; V_1, \ldots, V_m)$, by Lemma 4.14. Then by Proposition 4.15, there are finitely many constants $c_1, \ldots, c_m$ such that $(\mathcal{B}, c_1, \ldots, c_m)$ satisfies the hypothesis of Theorem 4.9, and the statement follows directly from Theorem 4.9. □

We are now ready to give a proof of Theorem 4.8.

Proof of Theorem 4.8. Let $\mathcal{A}$ be a finite-signature reduct of $(A; U_1, \ldots, U_n)$. Let $\mathcal{B}$ be the model-complete core of $\mathcal{A}$ and let $\mathcal{C}$ be the expansion of $\mathcal{B}$ by finitely many constants given by Corollary 4.16. Since $\mathcal{B}$ is a model-complete core, the set of automorphisms of $\mathcal{B}$ is dense in the set of endomorphisms. As in the proof of $3 \Rightarrow 2$ in Theorem 4.9, we can use this fact to prove that $\mathcal{C}$ has a Siggers polymorphism modulo endomorphisms if, and only if, $\mathcal{B}$ has such a polymorphism.

- If $\mathcal{C}$ has such a polymorphism, it has a canonical one, by Theorem 4.9. Let $m \geq 3$ be greater than the arity of any relation of $\mathcal{C}$. Then $T_{\mathcal{C},m}(\mathcal{C})$ has a Siggers polymorphism, by Lemma 3.2. By Theorem 2.5, the CSP of $T_{\mathcal{C},m}(\mathcal{C})$ is in P. It follows from Theorem 3.1 that $\mathrm{CSP}(\mathcal{C})$ is in P, too.
- If $\mathcal{C}$ does not have a Siggers polymorphism modulo endomorphisms, then Corollary 4.16 gives a uniformly continuous clonoid homomorphism from $\mathrm{Pol}(\mathcal{C})$ to $\mathscr{P}$. By Corollary 2.15, $\mathrm{CSP}(\mathcal{C})$ is NP-complete.

□

We mention that the condition in Theorem 4.8 is decidable: given subsets $U_1, \ldots, U_n$ of $\mathcal{A}$ (given by the sizes of the sets in the boolean algebra they generate), it is easily seen that one can compute a finite set of bounds for the age of $(A; U_1, \ldots, U_n)$. Given first-order formulas that define the relations of $\mathcal{A}$ over $(A; U_1, \ldots, U_n)$, it is also possible to compute the model-complete core $\mathcal{B}$ of $\mathcal{A}$. Our results then imply that $\mathcal{B}$ has a Siggers polymorphism modulo endomorphisms if, and only if, it has a canonical one. Testing the existence of a canonical pseudo-Siggers polymorphism is then decidable, since it is equivalent to testing the existence of a Siggers polymorphism of $T_{\mathcal{B},m}(\mathcal{B})$ for m large enough.

Chapter 5

MMSNP: Proof of the Algebraic Dichotomy Conjecture

Monotone Monadic SNP (MMSNP) is a fragment of monadic existential second-order logic whose sentences describe problems of the form "given a structure $\mathcal{A}$, is there a colouring of the elements of $\mathcal{A}$ that avoids some fixed family of forbidden patterns?" Examples of such problems are the classical k-colourability problem for graphs (where the forbidden patterns are edges whose endpoints have the same colour), or the problem of colouring the vertices of a graph so as to avoid monochromatic triangles (Figure 5.1).

MMSNP has been introduced by Feder and Vardi [52], whose motivation was to find fragments of existential second-order logic that exhibit a complexity dichotomy between P and NP-complete. They proved that every problem described by an MMSNP sentence is equivalent under polynomial-time randomised reductions to a CSP over a finite domain. Kun [71] later improved the result by derandomising the equivalence, thus showing that MMSNP exhibits a complexity dichotomy if and only if the Feder-Vardi dichotomy conjecture holds. By Theorem 2.5, we obtain that MMSNP exhibits a complexity dichotomy.

Dalmau and Bodirsky [14] showed that every problem in MMSNP is a finite union of constraint satisfaction problems for ω-categorical structures. These structures can be expanded to finitely bounded homogeneous structures so that they fall into the scope of Conjecture 1. It is easy to see that in order to prove the MMSNP dichotomy, it suffices to prove the dichotomy for those MMSNP problems that are CSPs (see Section 5.1.1). This poses the question whether the complexity of MMSNP can be studied directly using the universal-algebraic approach, rather than the reduction of Kun which involves a complicated construction of expander structures. In particular, even though we now have a complexity dichotomy for MMSNP, it was hitherto unknown whether the CSPs in MMSNP satisfy the infinite-domain tractability conjecture.

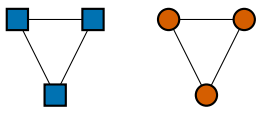

Figure 5.1: The No-Mono-Triangle problem: the input is a finite graph $\mathcal{G}$, and the question is whether there exists a colouring of the vertices of $\mathcal{G}$ with two colours that avoids monochromatic triangles. Colours are also shown with different shape of vertices for visual aid only. Non-coloured vertices will appear as white round vertices in the following.

The main result of this chapter is the confirmation of Conjecture 1 for CSPs in MMSNP. As a by-product, we obtain a new proof of the complexity dichotomy for MMSNP that does not rely on the results of Kun. To the best of our knowledge, this is the first-time that the universal-algebraic approach for infinite-domain CSPs provides a classification for a class of computational problems that has been studied in the literature before[1], and which has been introduced without having the universal-algebraic approach in mind. We also solve an open question by Lutz and Wolter [73]. Informally, we prove that the existential second-order predicates of an MMSNP sentence can be added to the original (first-order) signature of the sentence without increasing the complexity of the corresponding problem; we refer the reader to Section 5.3 for a formal statement.

The strategy for our proof is as follows. Let $\mathcal{A}$ be an ω-categorical structure such that $\mathrm{CSP}(\mathcal{A})$ is described by an MMSNP sentence Φ. First, we exhibit an MMSNP sentence Ψ in *strong normal form* such that Φ and Ψ are equivalent (and in particular, $\mathrm{CSP}(\mathcal{A})$ is described by Ψ). For this sentence Ψ, we construct another ω-categorical structure $\mathcal{C}_\Psi$ whose CSP is described by Ψ and proceed with the following steps:

1. Using the infinite-to-finite reduction from Chapter 3, we show that $\mathrm{CSP}(\mathcal{C}_\Psi)$ is in P if $\mathrm{Pol}(\mathcal{C}_\Psi)$ has a *canonical* polymorphism that behaves on the orbits of the template as a Siggers operation.

2. In order to prove that this is the only way to obtain polynomial-time tractability, we want to show that the absence of such a canonical polymorphism is equivalent to the existence of a uniformly continuous clonoid homomorphism to the clone of projections, which is known to entail NP-hardness (Corollary 2.15). We construct this map by first defining a clonoid homomorphism from the clone of canonical polymorphisms of the template to the clone of projections, followed by extending this map to the whole polymorphism clone (similarly as in Chapter 4). For this, two ingredients are necessary.

3. The first one is the fact that the template has the canonisation property as defined in Chapter 4. This requires proving that the template under consideration has an ω-categorical Ramsey expansion, which follows from recent results of Hubička and Nešetřil [58].

4. The second ingredient is the fact that every polymorphism of our template canonises in essentially one way. In Chapter 4, this was proved using the mashup property. In the case of MMSNP, we were unable to prove the mashup property for our template (although no counterexample was found). Indeed, we bypass this problem through an analysis of the binary symmetric relations that are preserved by the polymorphisms of the template.

This presentation of the strategy oversimplifies certain aspects, and we have to defer a more precise discussion to Section 5.4. This chapter contains published results from [24].

[1] `https://complexityzoo.uwaterloo.ca/Complexity_Zoo:M#mmsnp`.

5.1 MMSNP

Let τ be a relational signature (we also refer to τ as the *input signature*). *SNP* is a syntactically restricted fragment of existential second order logic. A sentence in SNP is of the form $\exists P_1, \dots, P_n . \phi$ where $P_1, \dots, P_n$ are *predicates* (i.e., relation symbols) and ϕ is a *universal* first-order-sentence over the signature $\tau \cup \{P_1, \dots, P_n\}$. *Monotone Monadic SNP without inequality*, MMSNP, is the popular restriction thereof which consists of sentences Φ of the form

$$\exists P_1, \dots, P_n \, \forall \bar{x} \bigwedge_i \neg(\alpha_i \wedge \beta_i),$$

where $P_1, \dots, P_n$ are *monadic* (i.e., unary) relation symbols not in τ, where $\bar{x}$ is a tuple of first-order variables, and for every negated conjunct:

- α_i consists of a conjunction of atomic formulas involving relation symbols from τ and variables from $\bar{x}$; and
- β_i consists of a conjunction of atomic formulas or negated atomic formulas involving relation symbols from $P_1, \dots, P_n$ and variables from $\bar{x}$.

Notice that the equality symbol is not allowed in MMSNP sentences.

Every MMSNP τ-sentence describes a computational problem: the input consists of a finite τ-structure $\mathcal{A}$, and the question is whether $\mathcal{A} \models \Phi$, i.e., whether the sentence Φ is true in $\mathcal{A}$. We sometimes identify MMSNP with the class of all computational problems described by MMSNP sentences.

5.1.1 Connected MMSNP

A pp-formula ϕ with at least one variable is called *connected* if the conjuncts of ϕ cannot be partitioned into two non-empty sets of conjuncts with disjoint sets of variables, and *disconnected* otherwise. Note that a pp-formula ϕ without equality conjuncts is connected if and only if the *Gaifman graph*[2] of the canonical database of ϕ is connected in the graph theoretic sense. A connected pp-formula is called *biconnected* if the conjuncts of ϕ cannot be partitioned into two non-empty sets of conjuncts that only share one common variable. Note that formulas with only one variable might not be biconnected, e.g., the formula $R_1(x) \wedge R_2(x)$ is not biconnected. An MMSNP τ-sentence Φ is called *connected* (or *biconnected*) if for each conjunct $\neg(\alpha \wedge \beta)$ of Φ where α is a conjunction of τ-formulas and β is a conjunction of unary formulas, the formula α is connected (or *biconnected*, respectively).

Proposition 5.1 (implicit in [52]; see also Section 6 of [76])**.** *Let Φ be an MMSNP sentence. Then Φ is logically equivalent to a finite disjunction of connected MMSNP sentences; these connected MMSNP sentences can be effectively computed from Φ.*

Proof. Let $P_1, \dots, P_k$ be the existential monadic predicates in Φ, and let τ be the input signature of Φ. Suppose that Φ has a conjunct $\neg(\alpha \wedge \beta)$ where α is a disconnected conjunction of atomic τ-formulas and β contains unary predicates only. Suppose that

[2] The *Gaifman graph* of a relational structure $\mathcal{A}$ is the undirected graph with vertex set A which contains an edge between $u, v \in A$ if and only if u and v both appear in a tuple contained in a relation of $\mathcal{A}$.

α is equivalent to $\alpha_1 \vee \alpha_2$ for non-empty formulas α_1 and α_2. Let Φ_1 be the MMSNP sentence obtained from Φ by replacing α by α_1, and let Φ_2 be the MMSNP sentence obtained from Φ by replacing α by α_2. It is then straightforward to check that every finite $(\tau \cup \{P_1, \ldots, P_k\})$-structure $\mathcal{A}$ we have that $\mathcal{A}$ satisfies the first-order part of Φ if and only if $\mathcal{A}$ satisfies the first-order part of Φ_1 or the first-order part of Φ_2. Iterating this process for each disconnected clause of ϕ, we eventually arrive at a finite disjunction of connected MMSNP sentences. □

It is well-known that the complexity classification for MMSNP can be reduced to the complexity classification for *connected* MMSNP; we add the simple proof for the convenience of the reader.

Proposition 5.2. *Let Φ be an MMSNP τ-sentence which is logically equivalent to $\Phi_1 \vee \cdots \vee \Phi_k$ for connected MMSNP τ-sentences $\Phi_1, \ldots, \Phi_k$ where k is smallest possible. Then Φ is in P if each of $\Phi_1, \ldots, \Phi_n$ is in P. If one of the Φ_i is NP-hard, then so is Φ.*

Proof. If each Φ_i can be decided in polynomial time by an algorithm A_i, then it is clear that Φ can be solved in polynomial time by running all of the algorithms $A_1, \ldots, A_k$ on the input, and accepting if one of the algorithms accepts.

Otherwise, if one of the Φ_i describes an NP-complete problem, then Φ_i can be reduced to Φ as follows. Since k was chosen to be minimal, there exists a τ-structure $\mathcal{B}$ such that $\mathcal{B}$ satisfies Φ_i, but does not satisfy Φ_j for all $j \leq n$ that are distinct from i, since otherwise we could have removed Φ_i from the disjunction $\Phi_1 \vee \cdots \vee \Phi_k$ without affecting the equivalence of the disjunction to Φ. We claim that $\mathcal{A} \uplus \mathcal{B}$ satisfies Φ if and only if $\mathcal{A}$ satisfies Φ_i. First suppose that $\mathcal{A}$ satisfies Φ_i. Since $\mathcal{B}$ also satisfies Φ_i by choice of $\mathcal{B}$, and since Φ_i is closed under disjoint unions, we have that $\mathcal{A} \uplus \mathcal{B}$ satisfies Φ_i as well. The statement follows since Φ_i is a disjunct of Φ.

For the opposite direction, suppose that $\mathcal{A} \uplus \mathcal{B}$ satisfies Φ. Since $\mathcal{B}$ does not satisfy Φ_j for all j distinct from i, $\mathcal{A} \uplus \mathcal{B}$ does not satisfy Φ_j as well, by monotonicity of Φ_j. Hence, $\mathcal{A} \uplus \mathcal{B}$ must satisfy Φ_i. By monotonicity of Φ_i, it follows that $\mathcal{A}$ satisfies Φ_i. Since $\mathcal{A} \uplus \mathcal{B}$ is for fixed $\mathcal{B}$ clearly computable from $\mathcal{A}$ in linear time this concludes our reduction from Φ_i to Φ. □

Proposition 5.3 (Corollary 1.4.15 in [11])**.** *An MMSNP sentence Φ describes a CSP if and only if Φ is logically equivalent to a connected MMSNP sentence.*

5.1.2 Templates for connected MMSNP sentences

In this section we first revisit the fact that every connected MMSNP sentence describes a CSP of an ω-categorical structure [14]. The proof uses a theorem due to Cherlin, Shelah, and Shi, stated for graphs in [48]; Theorem 5.4 below is formulated for general relational structures. Another proof of the theorem of Cherlin, Shelah, and Shi has been given by Hubička and Nešetřil [59].

For a set $\mathfrak{F}$ of finite structures, denote by $\mathrm{Forb}^{\mathrm{hom}}(\mathfrak{F})$ the set of all finite structures $\mathcal{A}$ such that for all $\mathcal{F} \in \mathfrak{F}$, there does not exist any homomorphism from $\mathcal{F}$ to $\mathcal{A}$.

Theorem 5.4 (Theorem 4 in [48])**.** *Let $\mathfrak{F}$ be a finite set of finite connected τ-structures. Then there exists a countable model-complete τ-structure $\mathcal{B}^{\mathrm{ind}}_{\mathfrak{F}}$ such that* $\mathrm{Age}(\mathcal{B}^{\mathrm{ind}}_{\mathfrak{F}}) =$

$\mathrm{Forb}^{\mathrm{hom}}(\mathfrak{F})$. *The structure* $\mathcal{B}^{\mathrm{ind}}_{\mathfrak{F}}$ *is* ω*-categorical and without algebraicity, and unique with these properties.*

Let Φ be a connected MMSNP τ-sentence. Let σ be the existentially quantified unary relation symbols in Φ, and let $\bar{\sigma}$ be the signature that contains a relation symbol $\bar{P}$ for every relation symbol $P \in \sigma$. We write $|\Phi|$ for the maximal number of variables in the clauses of Φ. For every $P \in \sigma$, add the clause $\neg(P(x) \wedge \bar{P}(x))$ to Φ. Let Φ' be the formula obtained from Φ by replacing each occurrence of $\neg P(y)$ in Φ by $\bar{P}(y)$. Then the *obstruction set for* Φ is the set $\mathfrak{F}$ of all finite connected $(\tau \cup \sigma \cup \bar{\sigma})$-structures $\mathcal{A}$ such that

- $A = \{1, \ldots, k\}$ for $k \leq |\Phi|$;
- for every $u \in A$ either $P(u)$ or $\bar{P}(u)$ holds;
- $\mathcal{A}$ falsifies a clause of Φ'.

Note that $\mathfrak{F}$ satisfies the conditions from Theorem 5.4.

Definition 5.1. Let Φ be an MMSNP sentence, and $\mathfrak{F}$ the obstruction set for Φ. Then $\mathcal{B}_\Phi$ denotes the substructure induced in $\mathcal{B}^{\mathrm{ind}}_{\mathfrak{F}}$ by all the elements b such that $\mathcal{B}^{\mathrm{ind}}_{\mathfrak{F}} \models P(b) \vee \bar{P}(b)$ for all $P \in \sigma$.

Let τ be a subset of the signature of $\mathcal{A}$; then the τ-reduct of $\mathcal{A}$ is the τ-structure obtained from $\mathcal{A}$ by dropping all relations that are not in τ, and denoted by $\mathcal{A}^\tau$. Note that reducts of ω-categorical structures are ω-categorical, and hence the structure $\mathcal{B}^\tau_\Phi$ is ω-categorical for all Φ.

Theorem 5.5 ([14])**.** *Let* Φ *be an MMSNP* τ*-sentence. Then a finite* τ*-structure* $\mathcal{A}$ *satisfies* Φ *if and only if* $\mathcal{A}$ *homomorphically maps to* $\mathcal{B}^\tau_\Phi$.

5.1.3 Statement of the main result

The main result of this chapter is the proof of the infinite-domain tractability conjecture (Conjecture 1) for CSPs in MMSNP. We actually show a stronger formulation than the conjecture since we also provide a characterisation of the polynomial-time tractable cases using pseudo-Siggers polymorphisms.

Combined with Proposition 5.1 we obtain the following theorem for MMSNP in general.

Theorem 5.6. *Let* Φ *be an MMSNP* τ*-sentence. Then* Φ *is logically equivalent to a finite disjunction* $\Phi_1 \vee \cdots \vee \Phi_k$ *of connected MMSNP sentences; for each* $i \leq k$ *there exists an* ω*-categorical structure* $\mathcal{B}_i$ *such that* Φ_i *describes* $\mathrm{CSP}(\mathcal{B}_i)$*, and either*

- $\mathrm{Pol}(\mathcal{B}_i)$ *has a uniformly continuous clonoid homomorphism to* $\mathscr{P}$*, for some* $i \in \{1, \ldots, k\}$*, and* Φ *is NP-complete, or*
- $\mathrm{Pol}(\mathcal{B}_i)$ *contains a pseudo-Siggers polymorphism, for each* $i \in \{1, \ldots, k\}$*, and* Φ *is in P.*

In particular, every problem in MMSNP is in P or NP-complete.

5.2 Normal Forms

We recall and adapt a normal form for MMSNP sentences that was initially proposed by Feder and Vardi in [51, 52] and later extended in [76]. The normal form has been invented by Feder and Vardi to show that for every connected MMSNP sentence Φ there is a polynomial-time equivalent finite-domain CSP. In their proof, the reduction from an MMSNP sentence to the corresponding finite-domain CSP is straightforward, but the reduction from the finite-domain CSP to Φ is tricky: it uses the fact that hard finite-domain CSPs are already hard when restricted to high-girth instances. The fact that MMSNP sentences in normal form are *biconnected* is then the key to reduce high-girth instances to the problem described by Φ.

In our work, the purpose of the normal form is the reduction of the classification problem to MMSNP sentences that are *precoloured* in a sense that will be made precise in Section 5.3, which is later important to apply the universal-algebraic approach. Moreover, we describe a new *strong normal form* that is based on recolourings introduced by Madelaine [75]. Recolourings have been applied by Madelaine to study the computational problem whether one MMSNP sentence implies another. In our context, the importance of strong normal forms is that the templates that we construct for MMSNP sentences in strong normal form, expanded with the inequality relation $\neq$, are model-complete cores (Theorem 5.25). Let us mention that in order to get this result, the biconnectivity of the MMSNP sentences in normal form is essential (e.g., the proof of Theorem 5.25 uses Corollary 5.15, which uses Lemma 5.13, which uses Lemma 5.8, which crucially uses biconnectivity of Φ).

5.2.1 The normal form for MMSNP

Every connected MMSNP sentence can be rewritten to a connected MMSNP sentence of a very particular shape, and this shape will be crucial for the results that we prove in the following sections.

Definition 5.2 (originates from [52]; also see [76])**.** Let Φ be an MMSNP sentence where $M_1, \dots, M_n$, for $n \geq 1$, are the existentially quantified predicates (also called the *colours* in the following). Then Φ is said to be in *normal form* if it is connected and

1. (Every vertex has a colour) the first conjunct of Φ is
$$\neg\big(\neg M_1(x) \wedge \cdots \wedge \neg M_n(x)\big)\,;$$
2. (Every vertex has at most one colour) Φ contains the conjunct
$$\neg\big(M_i(x) \wedge M_j(x)\big)$$
for all distinct $i, j \in \{1, \dots, n\}$;
3. (Clauses are fully coloured) for each conjunct $\neg\phi$ of Φ except the first, and for each variable x that appears in ϕ, there is an $i \leq n$ such that ϕ has a literal of the form $M_i(x)$;
4. (Clauses are biconnected) if a conjunct $\neg\phi$ of Φ is not of the form as described in item 1 and 2, the formula ϕ is biconnected;

5. (Small clauses are explicit) any $(\tau \cup \{M_1, \dots, M_n\})$-structure $\mathcal{A}$ with at most k elements satisfies the first-order part of Φ if $\mathcal{A}$ satisfies all conjuncts of Φ with at most k variables.

Note that when Φ is in normal form then in all conjuncts $\neg\phi$ of Φ except for the first we can drop conjuncts where predicates appear negatively in ϕ; hence, we assume henceforth that ϕ is a conjunction of atomic formulas. We illustrate item 4 and item 5 in this definition with the following examples.

Example 1. Let Φ be the connected MMSNP sentence

$$\forall a, b, c, d, e. \neg\big(E(a,b) \wedge E(b,c) \wedge E(c,d) \wedge E(d,e) \wedge E(e,a)\big)$$

which is in fact a first-order formula. The canonical database of

$$E(x_1, x_2) \wedge E(x_2, x_3) \wedge E(x_3, x_4) \wedge E(x_4, x_3) \wedge E(x_3, x_1)$$

has only four elements, does not satisfy Φ, but the only conjunct of Φ has five elements. So this is an example that satisfies all items except item 5 in the definition of normal forms.

However, Φ is logically equivalent to the following MMSNP formula, and it can be checked that this formula is in normal form.

$$\begin{aligned}\exists M_1 \forall x_0, \dots, x_4 \big(\neg(\neg M_1(x_0)) \wedge \neg(&\bigwedge_{0 \le i \le 4} M_1(x_i) \wedge E(x_i, x_{i+1 \bmod 5})) \\ \wedge \neg(&\bigwedge_{0 \le i \le 2} M_1(x_i) \wedge E(x_i, x_{i+1 \bmod 3})) \\ \wedge \neg(&M_1(x_0) \wedge E(x_0, x_0))\big).\end{aligned}$$

Adding clauses to an MMSNP sentence to obtain an equivalent sentence that satisfies item 5 can make a biconnected sentence not biconnected, as we see in the following example.

Example 2. Let Φ be the following biconnected MMSNP sentence.

$$\forall a, b, c, d. \neg\big(E(a,b) \wedge E(b,d) \wedge E(a,c) \wedge E(c,d)\big)$$

Note that Φ does not satisfy item 5 (it has implicit small clauses) and in fact is equivalent to

$$\forall a, b, d. \neg\big(E(a,b) \wedge E(b,d)\big)$$

which is not biconnected.

Lemma 5.7. *Every connected MMSNP sentence Φ is equivalent to an MMSNP sentence Ψ in normal form, and Ψ can be computed from Φ.*

Proof. We transform Φ in several steps (their order is important).

1: Biconnected clauses. Suppose that Φ contains a conjunct $\neg\phi$ such that ϕ is not biconnected, i.e., ϕ can be written as $\phi_1(x, \bar{y}) \wedge \phi_2(x, \bar{z})$ for tuples of variables $\bar{y}$ and $\bar{z}$ with disjoint sets of variables, and where ϕ_1 and ϕ_2 are conjunctions of atomic formulas. Then we introduce a new existentially quantified predicate P, and replace $\neg\phi$ by $\neg(\phi_1(x, \bar{y}) \wedge P(x)) \wedge \neg(\phi_2(x, \bar{z}) \wedge \neg P(x))$. Repeating this step, we can establish item 4 in the definition of normal forms.

2: Making implicit small clauses explicit. Let $\neg\phi(x_1, \ldots, x_n)$ be a conjunct of Φ that is not the first conjunct. Let x be a variable that does not appear among $x_1, \ldots, x_n$, and consider the formula $\phi(y_1, \ldots, y_n)$ where y_i is either x_i or x, and suppose that $y_i = y_j = x$ for at least two different $i, j \leq n$. If $\phi(y_1, \ldots, y_n)$ is biconnected, then add $\neg\phi(y_1, \ldots, y_n)$ to Φ. Otherwise, $\phi(y_1, \ldots, y_n)$ can be written as $\phi_1(x, \bar{z}_1) \wedge \phi_2(x, \bar{z}_2)$. We then apply the procedure from step 1 with the formula $\neg\phi(y_1, \ldots, y_n)$. In this way we can produce an equivalent MMSNP sentence that still satisfies item 4 (biconnected clauses). When we repeat this in all possible ways the procedure eventually terminates, and we claim that the resulting sentence Ψ satisfies additionally item 5. To see this, let $\mathcal{A}$ be a $(\tau \cup \{M_1, \ldots, M_n\})$-structure with at most k elements which does not satisfy some conjunct $\neg\phi$ of Φ. Pick the conjunct $\neg\phi$ from Φ with the least number of variables and this property. Then there are $a_1, \ldots, a_l \in A$ such that $\mathcal{A}$ satisfies $\phi(a_1, \ldots, a_l)$. If $l \leq k$, we are done. Otherwise, there must be $i, j \leq l$ such that $a_i = a_j$. If the conjunct $\neg\phi(y_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_{j-1}, x, x_{j+1}, \ldots, y_l)$ is biconnected, it has been added to Φ, and it has less variables than ϕ, a contradiction. Otherwise, our procedure did split the conjunct, and inductively we see that a clause that it not satisfied by $\mathcal{A}$ and has less variables than ϕ has been added to Φ.

3: Predicates as colours. Next, we want to ensure the property that Φ contains for each pair of distinct existentially quantified monadic predicates M_i, M_j the negated conjunct

$$\neg\big(M_i(x) \wedge M_j(x)\big),$$

and when $M_1, \ldots, M_c$ are all the existentially quantified predicates, then Φ contains the negated conjunct

$$\neg\big(\neg M_1(x) \wedge \cdots \wedge \neg M_c(x)\big).$$

We may transform every MMSNP sentence into an equivalent MMSNP sentence of this form, via the addition of further monadic predicates (2^n predicates starting from n monadic predicates). If $n = 0$ then Φ was a first-order formula; in this case, to have a unified treatment of all cases, we introduce a single existentially quantified predicate M_1, too.

4: Fully coloured clauses. Finally, if $\neg\phi$ is a conjunct of Φ and x a variable from ϕ such that x does not appear in any literal of the form $M_i(x)$ in ϕ, then we replace $\neg\phi$ by the conjuncts

$$\neg(\phi \wedge M_1(x)) \wedge \cdots \wedge \neg(\phi \wedge M_n(x)).$$

We do this for all conjuncts of Φ and all such variables, and obtain an MMSNP sentence that finally satisfies all the items from the definition of normal forms. □

Example 3. We revisit an MMSNP sentence from Example 2,

$$\forall a, b, c.\, \neg\big(E(a, b) \wedge E(b, c)\big)\,.$$

An equivalent MMSNP sentence Ψ in normal form is

$$\begin{aligned}\exists M_1, M_2\, \forall x, y\, \big(&\neg(\neg M_1(x) \wedge \neg M_2(x))\\ &\wedge \neg(M_1(x) \wedge M_2(x))\\ &\wedge \neg(M_1(x) \wedge R(x,x))\\ &\wedge \neg(M_2(x) \wedge E(x,x))\\ &\wedge \neg(M_1(x) \wedge M_1(y) \wedge E(x,y))\\ &\wedge \neg(M_2(x) \wedge M_2(y) \wedge E(x,y))\\ &\wedge \neg(M_2(x) \wedge M_1(y) \wedge E(x,y))\big).\end{aligned}$$

The following lemma states a key property that we have achieved with our normal form (in particular, we use the biconnectivity assumption).

Lemma 5.8. *Let ϕ be the first-order part of an MMSNP τ-sentence in normal form with colour set σ and let $\psi_1(x, \bar{y})$ and $\psi_2(x, \bar{z})$ be two conjunctions of atomic $(\tau \cup \sigma)$-formulas such that*

- *$\bar{y}$ and $\bar{z}$ are vectors of disjoint sets of variables;*
- *the canonical databases of ψ_1 and of ψ_2 satisfy ϕ;*
- *the canonical database $\mathcal{A}$ of $\psi_1(x, \bar{y}) \wedge \psi_2(x, \bar{z})$ does not satisfy ϕ.*

Then ψ_1 must contain a literal $M_i(x)$ and ψ_2 must contain a literal $M_j(x)$ for distinct colours M_i and M_j of ϕ.

Proof. First observe that all vertices of $\mathcal{A}$ must be coloured since all vertices of the canonical databases of ψ_1 and of ψ_2 are coloured (because they satisfy ϕ). Therefore, since $\mathcal{A}$ does not satisfy ϕ, there is a conjunct $\neg\phi'$ of ϕ and $a_1, \ldots, a_l \in A$ such that $\mathcal{A} \models \phi'(a_1, \ldots, a_l)$. Pick the conjunct such that l is minimal. Since both the canonical database of ψ_1 and of ψ_2 satisfy ϕ, not all of $a_1, \ldots, a_l$ can lie in the canonical database of ψ_1, or in the canonical database of ψ_2. If ϕ' is of the form $M_i(x) \wedge M_j(x)$ for $i \neq j$ then we are done. Otherwise, since ϕ' is biconnected, there are $i, j \leq n$ such that $a_i = a_j = x$. In this case, the structure $\mathcal{A}'$ induced by $a_1, \ldots, a_l$ in $\mathcal{A}$ has strictly less then l elements. Since Φ is in normal form, and since $\mathcal{A}'$ does not satisfy ϕ, by item 5 in the definition of normal forms there must be a conjunct $\neg\phi''$ of ϕ with at most $|A'|$ variables such that ϕ'' holds in $\mathcal{A}'$. This contradicts the choice of ϕ'. □

5.2.2 Templates for sentences in normal form

Let Φ be an MMSNP τ-sentence in normal form. Let σ be the set of colours of Φ. We will now construct an ω-categorical $(\tau \cup \sigma)$-structure $\mathcal{C}_\Phi$ for an MMSNP sentence Φ in normal form; this structure will have several important properties:

1. a structure $\mathcal{A}$ satisfies Φ if and only if $\mathcal{A}$ homomorphically maps to $\mathcal{C}_\Phi^\tau$;
2. for every colour M of Φ, $M^{\mathcal{C}_\Phi}$ is an orbit of elements under $\mathrm{Aut}(\mathcal{C}_\Phi)$; moreover, every orbit is of this form.

3. $(\mathcal{C}_\Phi, \neq)$ is a model-complete core;

4. if Φ is furthermore in *strong normal form* (to be introduced in Section 5.2.3) then even $(\mathcal{C}_\Phi^\tau, \neq)$ is a model-complete core.

If Φ is an MMSNP sentence in normal form, it is more natural to consider a variant of the notion of an obstruction set introduced in Section 5.1.2, which we call *coloured obstruction set*, because when Φ is in normal form we do not have to introduce a new symbol for the negation of each existentially quantified predicate to construct a template.

Definition 5.3. Let Φ be an MMSNP τ-sentence in normal form. The *coloured obstruction set* for Φ is the set $\mathfrak{F}$ of all canonical databases for formulas ϕ such that $\neg\phi$ is a conjunct of Φ, except for the first conjunct.

Theorem 5.4 has the following variant in the category of injective homomorphisms.

Theorem 5.9. *Let $\mathfrak{F}$ be a finite set of finite connected τ-structures. Then there exists a τ-structure $\mathcal{B}_\mathfrak{F}^{\text{hom}}$ such that*

- *a finite τ-structure $\mathcal{A}$ homomorphically and injectively maps to $\mathcal{B}_\mathfrak{F}^{\text{hom}}$ if and only if $\mathcal{A} \in \text{Forb}^{\text{hom}}(\mathfrak{F})$;*

- *$(\mathcal{B}_\mathfrak{F}^{\text{hom}}; \neq)$ is a model-complete core.*

The structure $\mathcal{B}_\mathfrak{F}^{\text{hom}}$ is ω-categorical, and is unique up to isomorphism with these properties.

Proof. Let $(\mathcal{B}_\mathfrak{F}^{\text{hom}}, \neq)$ be the model-complete core of $(\mathcal{B}_\mathfrak{F}^{\text{ind}}, \neq)$; by Theorem 2.18 the structure $(\mathcal{B}_\mathfrak{F}^{\text{hom}}; \neq)$ is unique up to isomorphism, and ω-categorical. Let $\mathcal{A}$ be a finite τ-structure. If $\mathcal{A} \in \text{Forb}^{\text{hom}}(\mathfrak{F})$, then $\mathcal{A}$ embeds into $\mathcal{B}_\mathfrak{F}^{\text{ind}}$ by Theorem 5.4, and since $(\mathcal{B}_\mathfrak{F}^{\text{ind}}, \neq)$ is homomorphically equivalent to $(\mathcal{B}_\mathfrak{F}^{\text{hom}}, \neq)$, there is an injective homomorphism from $\mathcal{A}$ to $\mathcal{B}_\mathfrak{F}^{\text{hom}}$. These reverse implication can be shown similarly, and this shows the first item. □

The structure $\mathcal{B}_\mathfrak{F}^{\text{ind}}$ from Theorem 5.4 and the structure $\mathcal{B}_\mathfrak{F}^{\text{hom}}$ from Theorem 5.9 might or might not be isomorphic, as we see in the following example.

Example 4. The structure $\mathcal{B}_\mathfrak{F}^{\text{hom}}$ might be isomorphic to the structure $\mathcal{B}_\mathfrak{F}^{\text{ind}}$: it is for example easy to verify that for $\mathfrak{F} := \{K_3\}$ the structure $\mathcal{B}_\mathfrak{F}^{\text{ind}}$ is a model-complete core, and therefore isomorphic to $\mathcal{B}_\mathfrak{F}^{\text{hom}}$.

In general, however, the two structures are not isomorphic. Consider for example the signature $\tau = \{E\}$ for E binary and $\mathfrak{F} := \{L\}$ where $L := (\{0\}; \{(0,0)\})$, i.e., L is the canonical database of $E(x,x)$. Then all finite τ-structures embed into $\mathcal{B}_\mathfrak{F}^{\text{ind}}$, but $\mathcal{B}_\mathfrak{F}^{\text{hom}}$ satisfies $\forall x, y\, (E(x,y) \vee x = y)$, i.e., $\mathcal{B}_\mathfrak{F}^{\text{hom}}$ is the countably infinite complete graph.

Definition 5.4. Let Φ be an MMSNP τ-sentence in normal form and let $\mathfrak{F}$ be the coloured obstruction set of Φ. Then $\mathcal{C}_\Phi$ denotes the substructure of $\mathcal{B}_\mathfrak{F}^{\text{hom}}$ induced by the coloured elements of $\mathcal{B}_\mathfrak{F}^{\text{hom}}$.

The τ-reduct $\mathcal{C}_\Phi^\tau$ of the structure $\mathcal{C}_\Phi$ that we constructed for an MMSNP sentence Φ in normal form is indeed a template for the CSP described by Φ.

Lemma 5.10. *Let Φ be an MMSNP τ-sentence in normal form and let $\mathcal{A}$ be a τ-structure. Then the following are equivalent.*

(1) $\mathcal{A} \models \Phi$;

(2) $\mathcal{A}$ homomorphically and injectively maps to $\mathcal{C}_\Phi^\tau$;

(3) $\mathcal{A}$ homomorphically maps to $\mathcal{C}_\Phi^\tau$.

Proof. Let ρ be the colour set and let $\mathfrak{F}$ be the coloured obstruction set of Φ. (1) $\Rightarrow$ (2). If $\mathcal{A}$ satisfies Φ it has a $(\tau \cup \sigma)$-expansion $\mathcal{A}'$ such that no structure in $\mathfrak{F}$ homomorphically maps to $\mathcal{A}'$. So $\mathcal{A}'$ homomorphically and injectively maps to $\mathcal{B}_{\mathfrak{F}}^{\text{hom}}$ by Theorem 5.9. Moreover, every element of $\mathcal{A}'$ is contained in one predicate from σ (because of the first conjunct of Φ) and hence the image of the embedding must lie in $\mathcal{C}_\Phi$.

(2) $\Rightarrow$ (3) is trivial. For (3) $\Rightarrow$ (1), let h be the homomorphism from $\mathcal{A}$ to $\mathcal{C}_\Phi^\tau$. Expand $\mathcal{A}$ to a $(\tau \cup \sigma)$-structure $\mathcal{A}'$ by colouring each element $a \in A$ by the colour of $h(a)$ in $\mathcal{C}_\Phi$; then there is no homomorphism from a structure $\mathcal{F} \in \mathfrak{F}$ to $\mathcal{A}'$, since the composition of such a homomorphism with h would give a homomorphism from $\mathcal{F}$ to $\mathcal{B}_{\mathfrak{F}}^{\text{ind}}$, a contradiction. The expansion $\mathcal{A}'$ also satisfies the first conjunct of Φ, and hence $\mathcal{A} \models \Phi'$. $\square$

In the following we prove that $\mathcal{C}_\Phi$ indeed has the properties that we announced at the beginning of this section. We start with some remarkable properties of the structure $\mathcal{B}_{\mathfrak{F}}^{\text{ind}}$ (Section 5.2.2) and continue with properties of $\mathcal{C}_\Phi$ (Section 5.2.2).

Properties of Cherlin-Shelah-Shi structures

An existential formula is called *primitive* if it does not contain disjunctions.

Lemma 5.11. *For every $k \in \mathbb{N}$, the orbits of k-tuples in $\mathcal{B}_{\mathfrak{F}}^{\text{ind}}$ can be defined by $\phi_1 \wedge \phi_2$ where ϕ_1 is a pp-formula and ϕ_2 is a conjunction of negated atomic formulas.*

Proof. It suffices to prove the statement for k-tuples $\bar{a}$ with pairwise distinct entries. Since $\mathcal{B}_{\mathfrak{F}}^{\text{ind}}$ is ω-categorical and model-complete, there is an existential definition $\phi(\overline{x})$ of the orbit of $\overline{a}$ in $\mathcal{B}_{\mathfrak{F}}^{\text{ind}}$. Since ϕ defines an orbit of k-tuples it can be chosen to be primitive. Moreover, since $\bar{a}$ is a tuple with pairwise distinct entries, ϕ can be chosen to be without conjuncts of the form $x = y$ (it is impossible that both x and y are among the free variables $x_1, \ldots, x_n$; if one of the variables is existentially quantified, we can replace all occurrences of it by the other variable and obtain an equivalent formula). Let ϕ_1 be the pp-formula obtained from ϕ by deleting all the negated conjuncts. Let ϕ_2 be conjunction of all negated atomic formulas that hold on $\bar{a}$. Clearly, ϕ implies $\phi_1 \wedge \phi_2$.

Let $\bar{b}$ be a tuple that satisfies $\phi_1 \wedge \phi_2$; we have to show that $\bar{b}$ satisfies ϕ. Let $\psi(x_1, \ldots, x_n)$ be the existential definition of the orbit of $\bar{b}$. Again, we may assume that ψ is disjunction-free and free of literals of the form $x = y$. Let ψ_1 be the formula obtained from ψ by dropping negated conjuncts. Let $\mathcal{A}$ be the canonical database of $\phi_1 \wedge \psi_1$ (which is well-defined since both ϕ_1 and ψ_1 are primitive positive and do not involve literals of the form $x = y$). We have $\mathcal{B}_{\mathfrak{F}}^{\text{ind}} \models \phi_1(\bar{b}) \wedge \psi_1(\bar{b})$, so $\mathcal{A}$ does not homomorphically embed any structure from $\mathfrak{F}$. By definition of $\mathcal{B}_{\mathfrak{F}}^{\text{ind}}$ (Theorem 5.4), there exists an embedding e of $\mathcal{A}$ into $\mathcal{B}_{\mathfrak{F}}^{\text{ind}}$. Then e provides witnesses for the existentially quantified variables in $\phi \wedge \psi$ showing that $\mathcal{B}_{\mathfrak{F}}^{\text{ind}} \models (\phi \wedge \psi)(e(x_1), \ldots, e(x_n))$ because for those witnesses the negated

conjuncts will also be satisfied. Hence, ϕ and ψ define the same orbit of n-tuples. In particular, t satisfies ϕ which is what we wanted to show. □

When $\mathcal{B}$ is a structure, we write $\mathcal{B}^*$ for the expansion of $\mathcal{B}$ by all pp-formulas.

Corollary 5.12. *The structure $(\mathcal{B}^{\text{ind}}_{\mathfrak{F}})^*$ is homogeneous.*

Proof. Let $\bar{a}$, $\bar{b}$ be two k-tuples of elements of $(\mathcal{B}^{\text{ind}}_{\mathfrak{F}})^*$ such that the map that sends a_i to b_i, for $i \in \{1, \dots, k\}$, is an isomorphism between the substructures induced by $\{a_1, \dots, a_n\}$ and by $\{b_1, \dots, b_n\}$ in $(\mathcal{B}^{\text{ind}}_{\mathfrak{F}})^*$. Then $\bar{a}$ and $\bar{b}$ satisfy in particular the same negated atomic formulas, and they also satisfy the same primitive positive formulas in $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$ since α must preserve the relations that we have introduced for these formulas in $(\mathcal{B}^{\text{ind}}_{\mathfrak{F}})^*$. The statement now follows from Lemma 5.11. □

Definition 5.5. A relational structure $\mathcal{B}$ is said to be *1-homogeneous* if it has the property that when $a, b \in B$ satisfy the same unary relations in $\mathcal{B}$, then there exists an automorphism of $\mathcal{B}$ that maps a to b.

Lemma 5.13. *Let Φ be an MMSNP sentence in normal form with coloured obstruction set $\mathfrak{F}$. Then $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$ is 1-homogeneous.*

Proof. Let a_1 and a_2 be two elements that induce isomorphic 1-element substructures of $\mathcal{B}^{\text{ind}}_{\Phi}$. Since $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$ is model-complete, the orbit of a_i, for $i = 1$ and $i = 2$, has a primitive definition $\psi_i(x)$ in $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$. Pick elements for the existentially quantified variables in ψ_i that witness the truth of $\psi_i(a_i)$, and let ψ'_i be the canonical query of the structure induced by a_i and those elements in $\mathcal{B}^{\text{ind}}_{\Phi}$.

Suppose for contradiction that a_1 and a_2 are in different orbits of $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$. This means that $\psi_1(x) \wedge \psi_2(x)$, and therefore also $\psi'_1(x) \wedge \psi'_2(x)$, is unsatisfiable in the structure $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$. Since a_1 and a_2 induce isomorphic 1-element substructures, the contrapositive of Lemma 5.8 shows that already the canonical database of ψ'_1 or of ψ'_2 does not satisfy the first-order part of Φ, a contradiction. □

Properties of our templates for MMSNP

Some properties that we have derived for $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$ transfer to $\mathcal{B}^{\text{hom}}_{\mathfrak{F}}$ and $\mathcal{C}_{\Phi}$.

Lemma 5.14. *Let Φ be an MMSNP sentence in normal form with coloured obstruction set $\mathfrak{F}$. Then $\mathcal{B}^{\text{hom}}_{\mathfrak{F}}$ is 1-homogeneous.*

Proof. We already know that $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$ is 1-homogeneous. Let f be an injective homomorphism from $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$ to $\mathcal{B}^{\text{hom}}_{\mathfrak{F}}$ and g an injective homomorphism from $\mathcal{B}^{\text{hom}}_{\mathfrak{F}}$ to $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$. Let u and v be two elements of $\mathcal{B}^{\text{hom}}_{\mathfrak{F}}$ that induce isomorphic 1-element substructures. Then $g(u)$ and $g(v)$ must induce isomorphic 1-element substructures, too, since otherwise the injection $e := f \circ g$ would not preserve all first-order formulas, in contradiction to the assumption that $(\mathcal{B}^{\text{hom}}_{\mathfrak{F}}, \neq)$ is a model-complete core. By the 1-homogeneity of $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$ (Lemma 5.13) there exists $\alpha \in \text{Aut}(\mathcal{B}^{\text{ind}}_{\mathfrak{F}})$ such that $\alpha(g(u)) = g(v)$. The mapping $e' := f \circ \alpha \circ g$ is an endomorphism of $(\mathcal{B}^{\text{hom}}_{\mathfrak{F}}, \neq)$, and since $(\mathcal{B}^{\text{hom}}_{\mathfrak{F}}, \neq)$ is a model-complete core there exists

$\beta \in \mathrm{Aut}(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}, \neq)$ such that $\beta(u) = e'(u)$. There also exists a $\gamma \in \mathrm{Aut}(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}, \neq)$ such that $\gamma(u) = e(v)$. Then

$$\begin{aligned}\gamma^{-1}(\beta(u)) &= \gamma^{-1}(f(\alpha(g(u)))) \\ &= \gamma^{-1}(f(g(v))) \\ &= \gamma^{-1}(e(v)) = v\end{aligned}$$

and so u and v are in the same orbit of $\mathrm{Aut}(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}})$. □

Corollary 5.15. *Let Φ be an MMSNP sentence in normal form. Then $\mathcal{C}_\Phi$ is 1-homogeneous.*

Proof. Let $\mathfrak{F}$ be the coloured obstruction set for Φ. Recall that $\mathcal{C}_\Phi$ is a substructure of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$. Let x and y be two elements of $\mathcal{C}_\Phi$ that induce isomorphic 1-element substructures. By Lemma 5.13, x and y lie in the same orbit of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$. When x and y are in the same orbit of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$, they are clearly also in the same orbit of $\mathcal{C}_\Phi$ since automorphisms of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ respect the domain of $\mathcal{C}_\Phi$. □

Lemma 5.16. *Let Φ be in normal form with colours $M_1, \ldots, M_n$. Let a and b be two elements of $\mathcal{C}_\Phi$ that induce non-isomorphic one-element structures in $\mathcal{C}_\Phi$. Then there are distinct $i, j \in \{1, \ldots, n\}$ such that $\mathcal{C}_\Phi \models M_i(a) \wedge M_j(b)$.*

Proof. By definition of $\mathcal{C}_\Phi$ there are $i, j \in \{1, \ldots, n\}$ such that $M_i(a)$ and $M_j(b)$. Let $\mathfrak{F}$ be the coloured obstruction set for Φ. Since $(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}, \neq)$ is a model-complete core, there is a pp-definition $\psi_1(x)$ of the orbit of a in $(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}, \neq)$, and similarly a primitive positive definition $\psi_2(x)$ of the orbit of b in $(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}, \neq)$. Pick witnesses for the existentially quantified variables that show that $\psi_1(a)$ and $\psi_2(b)$ hold, and let $\psi_1'(x)$ and $\psi_2'(x)$ be the pp-formulas in the language of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ that we obtain from ψ_1 and ψ_2 by

1. dropping the conjuncts that involve the symbol $\neq$, and
2. adding conjuncts of the form $M(x)$ for every existentially quantified variable, where M is the colour of the witness that we picked above.

Clearly, the canonical databases of ψ_1' and of ψ_2' satisfy the first-order part ϕ of Φ. We claim that the canonical database of $\psi_1'(x) \wedge \psi_2'(x)$ does not satisfy ϕ. Then Lemma 5.8 implies that $i \neq j$ and we are done.

To show the claim, suppose for contradiction that $\psi_1'(x) \wedge \psi_2'(x)$ is satisfiable. Then the canonical database of this formula homomorphically maps to $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$, and by the first item of Theorem 5.9 also injectively homomorphically map to $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$. Hence, the formula $\psi_1(x) \wedge \psi_2(x)$ is satisfiable as well (any injective homomorphism gives a satisfying assignment). But $\psi_1(x) \wedge \psi_2(x)$ cannot be satisfiable in $(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}, \neq)$ because a and b must lie in different orbits of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$. □

Note that Lemma 5.16 would be false if instead of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ we would have used $\mathcal{B}^{\mathrm{ind}}_{\mathfrak{F}}$ in the definition of $\mathcal{C}_\Phi$, as shown by the following example.

Example 5. Let τ be the signature that only contains the two unary predicates P and Q. Let Φ be the MMSNP τ-sentence in normal form with an empty coloured obstruction set $\mathfrak{F}$. Then $\mathcal{B}^{\mathrm{ind}}_{\mathfrak{F}}$ would have four orbits, but just one colour, so there are vertices of the same colour that lie in different orbits. But $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ has only one orbit, since all elements of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ must lie both in P and in Q.

The previous two lemmas jointly imply the following, which will become important in later sections.

Corollary 5.17. *Let Φ be in normal form. Then the colours of Φ denote the orbits of* $\mathrm{Aut}(\mathcal{C}_\Phi)$.

The final goal of this section is to prove that for MMSNP sentences Φ in normal form the structure $(\mathcal{C}_\Phi, \neq)$ is a model-complete core. To this end, we need the following.

Lemma 5.18. *Let Φ be an MMSNP τ-sentence in normal form and $\mathfrak{F}$ be the coloured obstruction set for Φ. Let $\bar{a}$ be a k-tuple of elements of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ which has an entry a_i that does not satisfy the first conjunct of Φ. Then $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}} \models R(\bar{a})$ for every $R \in \tau$ of arity k.*

Proof. Let $\mathcal{B}$ be the structure obtained from $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ by adding $\bar{a}$ to $R \in \tau$. We claim that $\mathcal{B}$ homomorphically maps to $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$. By ω-categoricity of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$, it suffices to prove that every finite substructure $\mathcal{B}'$ of the countable structure $\mathcal{B}$ homomorphically maps to $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$. No structure from $\mathfrak{F}$ homomorphically maps to $\mathcal{B}'$, since

- coloured obstructions from conjuncts as in item 2 of the definition of normal forms are satisfied by $\mathcal{B}$ since $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ satisfies the conjunct, and $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ and $\mathcal{B}$ coincide with respect to the unary relations;
- all other coloured obstructions cannot map to $\mathcal{B}$ since they are fully coloured (item 3 of the definition of normal forms) and the element a_i is by assumption not coloured.

Therefore $\mathcal{B}'$ homomorphically maps to $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ by the first item in the definition of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ from Theorem 5.9. Since the identity is a homomorphism from $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ to $\mathcal{B}$, and $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ is a model-complete core, we therefore must have that $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}} \models R(\bar{a})$. □

Lemma 5.19. *Let Φ be an MMSNP τ-sentence in normal form. Then $(\mathcal{C}_\Phi, \neq)$ is a model-complete core.*

Proof. Let $M_1, \ldots, M_n$ be the colours of Φ, and let $\mathfrak{F}$ be the coloured obstruction set for Φ. Let e be an endomorphism of $\mathcal{C}_\Phi$ and let $\bar{b}$ be a tuple of elements of $\mathcal{C}_\Phi$. We have to show that there exists an automorphism β of $\mathcal{C}_\Phi$ such that $\beta(\bar{b}) = e(\bar{b})$. We extend e to all elements of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ by setting $e(a) := a$ for all uncoloured elements a of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$, and verify that the resulting map e' is an endomorphism of $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$. Clearly, e' preserves M_i for all $i \leq n$. Let $R \in \tau$, and let $\bar{a}$ be such that $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}} \models R(\bar{a})$. If all entries of $\bar{a}$ are elements of $\mathcal{C}_\Phi$ then $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}} \models R(e'(\bar{a}))$ since $e'(\bar{a}) = e(\bar{a})$ and e' is an endomorphism. On the other hand, if $\bar{a}$ has an entry a_i which is not in $\mathcal{C}_\Phi$, then $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}} \models R(e'(\bar{a}))$ by Lemma 5.18. Since $(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}, \neq)$ is a model-complete core there exists an $\alpha \in \mathrm{Aut}(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}})$ such that $\alpha(\bar{b}) = e(\bar{b})$. The restriction β of α to $\mathcal{C}_\Phi$ is an automorphism of $\mathcal{C}_\Phi$ with the desired property. □

The canonisation property

Finally, we claim that all the templates considered so far can be extended by a linear order in a way that the resulting expansions have the canonisation property, as defined in Chapter 4. For this, we use the following result by Bodirsky and Pinsker [36] proving that every ω-categorical *Ramsey* structure has the canonisation property, and recent results by

Hubička and Nešetřil [59] proving that the expansion of our templates by an appropriate linear order are Ramsey structures. We choose not to delve here into the intricacies of Ramsey theory, and even the definition of Ramsey structure is omitted here as it is only a tool for us to connect the results from [36] and [59]. We invite the reader to consult the excellent papers [12] and [81] for an exposition to the topic.

Theorem 5.20 (Bodirsky, Pinsker [36])**.** *Let $\mathcal{B}$ be a countable ω-categorical Ramsey structure, and let $\mathcal{C}$ be ω-categorical. Then for any map $h\colon B^k \to C$ there exists a function in $\overline{\mathrm{Aut}(\mathcal{C})h\,\mathrm{Aut}(\mathcal{B})}$ that is canonical from $\mathcal{B}$ to $\mathcal{C}$.*

In order to use the results from [59], we first introduce another structure $\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}$. Let $\mathfrak{F}$ be a finite set of finite connected τ-structures of size $\leq m$. Consider, for every structure $\mathcal{A} \in \mathrm{Forb}^{\mathrm{hom}}(\mathfrak{F})$, the expansion $\mathcal{A}^*$ of $\mathcal{A}$ by all relations of arity $\leq m$ that are pp-definable in $\mathcal{A}$. Let $\mathfrak{A}$ be the set of all substructures of $\mathcal{A}^*$ for some $\mathcal{A} \in \mathrm{Forb}^{\mathrm{hom}}(\mathfrak{F})$, and let ρ be the signature of these structures. It was proved by Hubička and Nešetřil [57] that $\mathfrak{A}$ is an amalgamation class, and therefore is has a Fraïssé limit $\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}$ by Theorem 2.10. Consider moreover the class $\vec{\mathfrak{A}}$ of $(\rho \cup \{<\})$-structures obtained by endowing structures in $\mathfrak{A}$ by every possible linear order. This class is again an amalgamation class and its Fraïssé limit can be viewed as an expansion $(\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}, <)$ of $\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}$ by a linear order that we call the *free* linear order.

Theorem 5.21 (Hubička, Nešetřil [59])**.** *For every finite set $\mathfrak{F}$ of finite connected τ-structures, the structure $(\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}, <)$ is Ramsey.*

Similarly, one can extend $\mathcal{B}^{\mathrm{ind}}_{\mathfrak{F}}$, $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$, and $\mathcal{C}_\Phi$ by free linear orders. It follows from general Ramsey-theoretic results [12] that the Ramsey property of $(\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}, <)$ transfers to the structures $(\mathcal{B}^{\mathrm{ind}}_{\mathfrak{F}}, <)$, $(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}, <)$, and $(\mathcal{C}_\Phi, <)$.

Corollary 5.22. *Let Φ be an MMSNP sentence in normal form with coloured obstruction set $\mathfrak{F}$. The structures $(\mathcal{B}^{\mathrm{ind}}_{\mathfrak{F}}, <)$, $(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}, <)$, and $(\mathcal{C}_\Phi, <)$ are Ramsey and have the canonisation property.*

5.2.3 The strong normal form

Let Φ_1 and Φ_2 be two MMSNP τ-sentences in normal form with colour sets σ_1 and σ_2, respectively. For $r\colon \sigma_1 \to \sigma_2$ and a $(\tau \cup \sigma_1)$-structure $\mathcal{A}$ we write $r(\mathcal{A})$ for the structure obtained from $\mathcal{A}$ by renaming each predicate $P \in C_1$ to $r(P) \in C_2$.

Definition 5.6. A *recolouring (from Φ_1 to Φ_2)* is given by a function $r\colon \sigma_1 \to \sigma_2$ such that for every $(\tau \cup \sigma_1)$-structure $\mathcal{A}$, if a coloured obstruction of Φ_2 homomorphically maps to $r(\mathcal{A})$, then a coloured obstruction of Φ_1 homomorphically maps to $\mathcal{A}$. A recolouring $r\colon \sigma_1 \to \sigma_2$ is said to be *proper* if r is non-injective.

Example 6. Consider the MMSNP sentence Φ given by

$$\exists M_1, M_2\ \forall x\ \big((M_1(x) \vee M_2(x)) \wedge (\neg M_1(x) \vee \neg M_2(x))\big)$$

and note that this sentence is in normal form. There is a proper recolouring r from Φ to Φ, e.g., the map given by $r(M_1) = r(M_2) = M_1$.

Example 7. Consider the MMSNP $\{E\}$-sentence

$$\exists P\, \forall x, y\, \neg\big(\neg P(x) \wedge E(x,y) \wedge \neg P(y)\big)$$

It is not yet in normal form; an equivalent MMSNP sentence Φ in normal form is

$$\begin{aligned}\exists M_1, M_2\, \forall x, y\, \big(&\neg(\neg M_1(x) \wedge \neg M_2(x)) \wedge \\ &\neg(M_1(x) \wedge M_2(x)) \wedge \neg(M_1(x) \wedge E(x,y) \wedge M_1(y))\big)\end{aligned}$$

A proper recolouring from Φ to Φ is given by $r(M_1) = r(M_2) = M_2$. To verify that r is indeed a recolouring, consider the conjunct $\neg\phi_1 = \neg(M_1(x) \wedge E(x,y) \wedge M_1(y))$: when $\mathcal{B}_1$ is the canonical database of ϕ_1 then there does not exist any $(\tau \cup \sigma_1)$-structure $\mathcal{A}$ such that $r(\mathcal{A}) = \mathcal{B}_1$. For the conjunct $\neg\phi_2 = \neg(M_1(x) \wedge M_2(x))$, when B_2 is the canonical database of ϕ_2, there is again no $(\tau \cup \sigma_1)$-structure $\mathcal{A}$ such that $r(\mathcal{A}) = \mathcal{B}_2$.

In contrast, the map given by $r(M_1) = r(M_2) = M_1$ is *not* a recolouring: consider the canonical database $\mathcal{A}$ of the formula $M_1(x) \wedge E(x,y) \wedge M_2(y)$. It satisfies the quantifier-free part of Φ, but $r(\mathcal{A})$ is isomorphic to the canonical database of $\phi = (M_1(x) \wedge E(x,y) \wedge M_1(y))$, and $\neg\phi$ is a conjunct of Φ.

Lemma 5.23. *Let Φ_1 and Φ_2 be MMSNP τ-sentences in normal form. If r is a recolouring from Φ_1 to Φ_2, then every τ-structure that satisfies Φ_1 also satisfies Φ_2.*

Proof. Let τ be the signature of Φ_1 and Φ_2, and let σ_1 be the existentially quantified predicates of Φ_1. Let $\mathcal{A}$ be a finite model of Φ_1. We have to show that $\mathcal{A} \models \Phi_2$. Let σ_1 be the existentially quantified predicates of Φ_1. Let $\mathcal{A}'$ be the $(\tau \cup \sigma_1)$-expansion of $\mathcal{A}$ witnessing the truth of Φ_1 in $\mathcal{A}$. Since r is a recolouring, the structure $r(\mathcal{A}')$ does not embed any coloured obstruction of Φ_2, hence $\mathcal{A} \models \Phi_2$. □

An MMSNP sentence Φ is defined to be in *strong normal form* if it is in normal form and there is no proper recolouring from Φ to Φ.

Example 8. The MMSNP sentence Ψ from Example 3 is not only in normal form, but even in strong normal form.

Example 9. Example 6 was in normal form, but not in strong normal form. An equivalent formula in strong normal form is

$$\exists M_1 \forall x. \neg(\neg M_1(x)).$$

Theorem 5.24. *For every connected MMSNP sentence Φ there exists an equivalent connected MMSNP Ψ in strong normal form, and Ψ can be effectively computed from Φ.*

Proof. By Lemma 5.7, we can assume that Φ is already given in normal form; let σ be the colours of Φ. To compute a strong normal form for Φ we exhaustively check for proper recolourings from Φ to Φ.

If there is no such recolouring we are done. Otherwise, let r be such a proper recolouring. Let Ψ be the MMSNP sentence obtained from Φ by performing the following for each colour M not in the image of r:

1. drop all conjuncts $\neg\phi$ of Φ such that M appears positively in ϕ,

2. remove the literal in which M appears negatively from the first conjunct of Φ, and

3. remove M from the existential quantifier prefix of Φ.

(Step 1 and 2 amount to replacing M by *false*.) Since the identity map is clearly a recolouring from Ψ to Φ, Lemma 5.23 implies that Ψ is equivalent to Φ. We now repeat the procedure with Ψ instead of Φ. Since Ψ has less existential predicates than Φ this procedure must eventually terminate with an MMSNP sentence in strong normal form that is equivalent to the sentence we started with. □

Theorem 5.25. *Let Φ be an MMSNP sentence in strong normal form and with input signature τ. Then $(\mathcal{C}_\Phi^\tau, \neq)$ is a model-complete core.*

Proof. Let $\mathcal{C}$ be the model-complete core of $(\mathcal{C}_\Phi^\tau, \neq)$, and let h be a homomorphism from $(\mathcal{C}_\Phi^\tau, \neq)$ to $\mathcal{C}$. Since $\mathcal{C}$ is isomorphic to a substructure of $(\mathcal{C}_\Phi^\tau, \neq)$ we can assume in the following that $\mathcal{C}$ equals such a substructure. It suffices to show that $\mathcal{C}$ and $(\mathcal{C}_\Phi^\tau, \neq)$ satisfy the same first-order formulas, as this implies that $\mathcal{C}$ and $(\mathcal{C}_\Phi^\tau, \neq)$ are isomorphic by ω-categoricity. By Corollary 5.22, there exists a function

$$g \in \overline{\{\beta \circ h \circ \alpha \mid \alpha \in \mathrm{Aut}(\mathcal{C}_\Phi, <), \beta \in \mathrm{Aut}(\mathcal{C})\}}$$

that is canonical as a function from $(\mathcal{C}_\Phi, <)$ to $\mathcal{C}$, and an endomorphism of $(\mathcal{C}_\Phi^\tau, \neq)$ (recall that $\mathcal{C}$ is a substructure of $(\mathcal{C}_\Phi^\tau, \neq)$).

Since g is canonical, it induces a function on the orbits of elements under $\mathrm{Aut}(\mathcal{C}_\Phi)$, which are exactly the colours of Φ by Corollary 5.17. This induced function must be a permutation, otherwise we would obtain a proper recolouring of Φ, contradicting the assumption that Φ is in strong normal form. Thus, if n is the number of colours in Φ, we obtain that g^n acts as the identity on the orbits of elements and is therefore an endomorphism of $(\mathcal{C}_\Phi, \neq)$, which is a model-complete core (Lemma 5.19). Thus, $g^n \in \overline{\mathrm{Aut}(\mathcal{C}_\Phi, \neq)}$ so that on every finite subset of $\mathcal{C}_\Phi$, g is invertible by an element of $\mathrm{End}(\mathcal{C}_\Phi, \neq)$ ($\alpha \circ g^{n-1}$ is an inverse for an appropriate $\alpha \in \mathrm{Aut}(\mathcal{C}_\Phi, \neq)$) and it follows that $g \in \overline{\mathrm{Aut}(\mathcal{C}_\Phi, \neq)}$. In particular, g preserves the truth of every first-order formula, so that $\mathcal{C}$ and $(\mathcal{C}_\Phi, \neq)$ are isomorphic. □

We give an example that shows that the assumption that Φ is in *strong* normal form in Theorem 5.25 is necessary.

Example 10. Consider again the MMSNP sentence

$$\exists P\, \forall x, y.\, \neg\big(\neg P(x) \wedge E(x, y) \wedge \neg P(y)\big)$$

from Example 7; as we have observed, it is not in strong normal form. And indeed, the domain of $(\mathcal{C}_\Phi^\tau, \neq)$ consists of two countably infinite sets such there are no edges within the first set, and otherwise all edges are present. Clearly, this structure is not a model-complete core since there are endomorphisms whose range does not contain any element from the first set.

5.3 Precoloured MMSNP

An MMSNP τ-sentence Φ in normal form is called *precoloured* if, informally, for each colour of Φ there is a corresponding unary relation symbol in τ that forces elements to have this colour. In this section we show that every MMSNPsentence is polynomial-time equivalent to a precoloured MMSNP sentence; this answers a question posed in [73]. We first formally introduce precoloured MMSNP and state some basic properties in Section 5.3.1. We then prove a stronger result than the complexity statement above: we show that the Bodirsky-Pinsker tractability conjecture is true for CSPs in MMSNP if and only if it is true for CSPs in precoloured MMSNP (Theorem 5.31). In order to prove this stronger result we relate in Section 5.3.2 the algebraic properties of $\mathcal{C}_\Phi^\tau$ with the algebraic properties of the expansion of $\mathcal{C}_\Phi^\tau$ by the inequality relation $\neq$. The main results are stated in Section 5.3.3. In Section 5.3.4 we complete the proofs of the results in this section.

5.3.1 Basic properties of precoloured MMSNP

Formally, an MMSNP τ-sentence Φ is *precoloured* if it is in normal form and for every colour M of Φ there exists a unary symbol $P_M \in \tau$ such that for every colour M' of Φ which is distinct from M the formula Φ contains the conjunct $\neg(P_M(x) \wedge M'(x))$.

Lemma 5.26. *Every precoloured MMSNP sentence is in strong normal form.*

Proof. Let Φ be a precoloured MMSNP sentence with colour set σ. We will show that every recolouring $r\colon \sigma \to \sigma$ of Φ must be the identity. Let $M \in \sigma$, and let A be the canonical database of $P_M(x) \wedge M(x)$. Note that A does not homomorphically embed any coloured obstruction of Φ. But if $M' := r(M) \neq M$, then $r(A)$ homomorphically embeds the canonical database of $P_M(x) \wedge M'(x)$, in contradiction to the assumption that r is a recolouring. Hence, $r(M) = M$ for all $M \in \sigma$. □

Finally, we prove an important property that will be used in Section 5.4: the colours in a precoloured MMSNP sentence Φ denote (all) the orbits of $\mathrm{Aut}(\mathcal{C}_\Phi^\tau)$.

Lemma 5.27. *Let Φ be a precoloured MMSNP sentence. Then for each colour M, the symbol P_M and M both interpret the same orbit of $\mathrm{Aut}(\mathcal{C}_\Phi) = \mathrm{Aut}(\mathcal{C}_\Phi^\tau)$, and each orbit is denoted by some colour M of Φ.*

Proof. By Lemma 5.19 the structure $(\mathcal{C}_\Phi; \neq)$ is a model-complete core. Note that the ω-categorical structures $(\mathcal{C}_\Phi; \neq, M)$ and $(\mathcal{C}_\Phi; \neq, P_M)$ have the same CSP, and hence they are homomorphically equivalent. The fact that ω-categorical model-complete cores are up to isomorphism unique then implies that M and P_M have the same interpretation in $\mathcal{C}_\Phi$. Since Φ is in particular in normal form, Corollary 5.17 states that M and P_M denote an orbit of $\mathrm{Aut}(\mathcal{C}_\Phi) = \mathrm{Aut}(\mathcal{C}_\Phi^\tau)$, and that each orbit of $\mathrm{Aut}(\mathcal{C}_\Phi)$ is denoted by some color of $\mathcal{C}_\Phi$. □

5.3.2 Adding inequality

Let Φ be an MMSNP sentence in normal form. We first show that adding the inequality relation to $\mathcal{C}_\Phi^\tau$ does not increase the complexity of its CSP.

Proposition 5.28. $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$ *and* $\mathrm{CSP}(\mathcal{C}_\Phi^\tau, \neq)$ *are polynomial-time equivalent.*

Proof. If a given instance of $\mathrm{CSP}(\mathcal{C}_\Phi^\tau, \neq)$, viewed as a pp-sentence, contains conjuncts of the form $x \neq x$, then the instance is unsatisfiable. Otherwise, we only consider the constraints using relations from τ, and let $\mathcal{A}$ be the canonical database of those constraints. If $\mathcal{A}$ has no homomorphism to $\mathcal{C}_\Phi^\tau$ then the instance is unsatisfiable. Otherwise, by Lemma 5.10 there is an injective homomorphism from $\mathcal{A}$ to $\mathcal{C}_\Phi^\tau$. The injectivity implies that the homomorphism also satisfies all the inequality constraints, so we have a polynomial-time reduction from $\mathrm{CSP}(\mathcal{C}_\Phi^\tau, \neq)$ to $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$. □

We would now like to prove that $\mathcal{C}_\Phi^\tau$ satisfies the Bodirsky-Pinsker conjecture if and only if $(\mathcal{C}_\Phi^\tau, \neq)$ does. However, we do not know whether $(\mathcal{C}_\Phi^\tau, \neq)$ in general has a pp-construction in $\mathcal{C}_\Phi^\tau$. But we can prove the following, which turns out to be sufficient.

Proposition 5.29. *There exists a uniformly continuous clonoid homomorphism from* $\mathrm{Pol}(\mathcal{C}_\Phi^\tau)$ *to* $\mathscr{P}$ *if, and only if, there exists a uniformly continuous clonoid homomorphism from* $\mathrm{Pol}(\mathcal{C}_\Phi^\tau, \neq)$ *to* $\mathscr{P}$.

In the proof of this lemma, we need the following proposition.

Lemma 5.30. *Let $\mathcal{A}$ be any structure that has a homomorphism g to $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$. Then there exists an injective homomorphism $h\colon \mathcal{A} \to \mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$ such that for all tuples $\overline{a}$ from $\mathcal{A}$ and all existential formulas ϕ without equality literals, if $\phi(g(\overline{a}))$ holds in $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$, then $\phi(h(\overline{a}))$ also holds in $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$. Moreover, for all injective tuples $\overline{a}, \overline{b}$ from $\mathcal{A}$, if $g(\overline{a})$ and $g(\overline{b})$ lie in the same orbit in $\mathrm{Aut}(\mathcal{B}_\mathfrak{F}^{\mathrm{ind}})$ then $h(\overline{a})$ and $h(\overline{b})$ lie in the same orbit in $\mathrm{Aut}(\mathcal{B}_\mathfrak{F}^{\mathrm{ind}})$.*

Proof. Assume first that $\mathcal{A}$ is finite with domain A. Build a new structure $\mathcal{A}'$ as follows. For every $\overline{a}$ in $\mathcal{A}$ and existential formula $\phi(\overline{x}) := \exists y_1, \ldots, y_s.\psi(\overline{x}, \overline{y})$ such that $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}} \models \phi(g(\overline{a}))$ holds, pick elements $b_1, \ldots, b_s$ of $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$ such that $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}} \models \psi(g(\overline{a}), b_1, \ldots, b_s)$. Let A' be the set consisting of A as well as new elements $a'_1, \ldots, a'_s$, and define $g(a'_i) := b_i$. Let $\mathcal{A}'$ be the $(\tau \cup \sigma)$-structure on A' obtained by pulling back the relations from the structure induced by $g(A')$ in $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$. We therefore have that g is a homomorphism $\mathcal{A}' \to \mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$. It follows that there exists an embedding $h\colon \mathcal{A}' \to \mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$.

We prove the first part of the statement. Let $\phi(\overline{x}) := \exists y_1, \ldots, y_s.\psi(\overline{x}, \overline{y})$ be an existential formula not containing equality literals (positive or negative). Assume that $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}} \models \phi(g(\overline{a}))$. By construction and the fact that ϕ does not contain equality literals, this is equivalent to $\mathcal{A}' \models \psi(\overline{a}, a'_1, \ldots, a'_s)$ for some elements $a'_1, \ldots, a'_s \in A'$. Since h is an embedding, this implies $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}} \models \exists y_1, \ldots, y_s.\psi(h(\overline{a}), \overline{y})$, i.e., $\phi(h(\overline{a}))$ holds in $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$.

We now prove the second part of the statement. Let $\overline{a}, \overline{b}$ be injective tuples from $\mathcal{A}$. Since $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$ is ω-categorical and by Theorem 2.6, the orbit of the tuple $g(\overline{a})$ has a first-order definition $\phi(\overline{x})$. Since $\mathcal{B}_\mathfrak{F}^{\mathrm{ind}}$ is model-complete and ϕ defines an orbit, we can assume that ϕ is existential without disjunctions, of the form $\exists y_1, \ldots, y_s \left(\psi_1(\overline{x}, \overline{y})) \wedge \psi_2(\overline{x})\right)$ with ψ_1 quantifier-free and without equality literals, and ψ_2 a conjunction of literals of the form $x_i \neq x_j$. Since h is injective and the tuples $\overline{a}$ and $\overline{b}$ are injective, $\psi_2(h(\overline{a}))$ and $\psi_2(h(\overline{b}))$ hold. Moreover, since ψ_1 is without equality literals, the previous paragraph gives us that both $\exists y_1, \ldots, y_s.\psi_1(h(\overline{a}), \overline{y})$ and $\exists y_1, \ldots, y_s.\psi(h(\overline{b}), \overline{y})$ hold. Therefore, $h(\overline{a})$ and $h(\overline{b})$ lie in the same orbit of $\mathrm{Aut}(\mathcal{B}_\mathfrak{F}^{\mathrm{ind}})$.

In case $\mathcal{A}$ is infinite, it suffices to apply a compactness argument using the statement for finite substructures of $\mathcal{A}$. □

Proof of Proposition 5.29. Let $\mathcal{K}_3$ be the complete graph on $\{R, G, B\}$. It is known that $\mathrm{Pol}(\mathcal{K}_3, R, G, B) = \mathscr{P}$. Thus, to prove Proposition 5.29, it suffices by Theorem 2.14 to prove that $\mathcal{K}_3$ is pp-constructible in $\mathcal{C}_\Phi^\tau$ if, and only if, it is pp-constructible in $(\mathcal{C}_\Phi^\tau, \neq)$. Suppose then that $\mathcal{K}_3$ is homomorphically equivalent to a pp-power $\mathcal{A}$ of $(\mathcal{C}_\Phi^\tau, \neq)$. Let $\phi_E(x_1, \ldots, x_d, y_1, \ldots, y_d)$ be the defining pp-formula of the edge relation of $\mathcal{A}$. Without loss of generality, we can suppose that ϕ_E does not contain literals of the form $x_i = x_j$, $y_i = y_j$, or $x_i = y_j$ with $i \neq j$, as otherwise we can take a smaller d.

Let ψ_E be the formula ϕ_E where all the inequality literals have been removed (note that a literal $x \neq x$ cannot appear, for otherwise the edge relation of $\mathcal{A}$ is empty, and $\mathcal{K}_3$ would not have a homomorphism to $\mathcal{A}$). Let $\mathcal{B}$ be the pp-power of $\mathcal{C}_\Phi^\tau$ defined by ψ_E. Observe that $\mathcal{B}$ contains all the edges of $\mathcal{A}$, so $\mathcal{B}$ contains a triangle.

<u>Claim:</u> $\mathcal{B}$ does not contain any loop.

Proof. Suppose the contrary, and let $\overline{a} \in B$ be such that $\mathcal{C}_\Phi^\tau \models \psi_E(\overline{a}, \overline{a})$. Let $D = \{b_1, \ldots, b_d, c_1, \ldots, c_d\}$ be a set with at most $2d$ elements, where $b_i = c_i$ iff the literal $x_i = y_i$ is in ϕ_E. Let $g\colon b_i, c_i \mapsto a_i$ for all $i \in \{1, \ldots, d\}$. Let $\mathcal{D}$ be the $(\tau \cup \sigma)$-structure on D obtained by pulling back the relations from the structure induced by $g(D)$ in $\mathcal{C}_\Phi$. Note that all the elements of $\mathcal{D}$ are coloured. By Lemma 5.30, there is an injective homomorphism $g'\colon \mathcal{D} \to \mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}$ with the additional property that $g'(\overline{b})$ and $g'(\overline{c})$ are in the same orbit in $\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}$, because $g(\overline{b})$ and $g(\overline{c})$ are in the same orbit (they are actually equal). By composing with an appropriate $\alpha \in \mathrm{Aut}(\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}})$, we can assume that $g'(\overline{b})$ and $g'(\overline{c})$ are in the same orbit in $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}, <)$. Compose with an injective homomorphism $h\colon \mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}} \to \mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}$ that is canonical from $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}, <)$ to $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, <)$ to get an injective homomorphism $g''\colon \mathcal{D} \to \mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}$ such that $g''(\overline{b})$ and $g''(\overline{c})$ are in the same orbit in $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, <)$. Note that all the elements of the image of g'' are coloured, because all the elements of $\mathcal{D}$ are coloured. So the image of g'' lies in $\mathcal{C}_\Phi$.

We prove that $\phi_E(g''(\overline{b}), g''(\overline{c}))$ holds in $\mathcal{C}_\Phi^\tau$. Indeed, $\mathcal{C}_\Phi^\tau \models \psi_E(g(\overline{b}), g(\overline{c}))$. We want to use Lemma 5.30, except that ψ_E can contain literals of the form $x_i = y_i$. Therefore an application of Lemma 5.30 only gives us that the tuple $(g'(\overline{b}), g'(\overline{c}))$ satisfies the equality-free part of ψ_E. But if $x_i = y_i$ is in ψ_E (and in ϕ_E), by construction we chose $b_i = c_i$, so that $g'(b_i) = g'(c_i)$. It follows that $\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}} \models \psi_E(g'(\overline{b}), g'(\overline{c}))$. This implies that $\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}} \models \psi_E(g''(\overline{b}), g''(\overline{c}))$ and by injectivity of g'', the pair $(g''(\overline{b}), g''(\overline{c}))$ also satisfies $x_i \neq y_j$ whenever $x_i = y_j$ is not in ϕ_E. In particular, if $x_i \neq y_j$ *is* in ϕ_E, we have $g''(b_i) \neq g''(c_j)$. Hence, $\mathcal{C}_\Phi^\tau \models \phi_E(g''(\overline{b}), g''(\overline{c}))$ holds.

Let now $\chi\colon \mathcal{A} \to \mathcal{K}_3$ be a homomorphism, that we can moreover suppose to be canonical from $(\mathcal{C}_\Phi, <)$ to $(\mathcal{K}_3, R, G, B)$ by Corollary 5.22. Since χ is canonical, we have that $\chi(g''(\overline{b})) = \chi(g''(\overline{c}))$. This contradicts the fact that χ is a homomorphism $\mathcal{A} \to \mathcal{K}_3$. Therefore, $\mathcal{B}$ has no loops. $\Diamond$

We now prove that every finite substructure $\mathcal{S}$ of $\mathcal{B}$ has a homomorphism to $\mathcal{K}_3$ (which proves, by compactness, that $\mathcal{B}$ has a homomorphism to $\mathcal{K}_3$). Let $\overline{s}^1 = (s_1^1, \ldots, s_d^1), \ldots, \overline{s}^K = (s_1^K, \ldots, s_d^K)$ be a list of the elements of $\mathcal{S}$. Let $\theta(\overline{x}^1, \ldots, \overline{x}^K)$ be the formula with Kd free variables that is a conjunction of the formulas $\psi_E(\overline{x}^i, \overline{x}^j)$ for all $i, j \in \{1, \ldots, K\}$ such that $\mathcal{C}_\Phi^\tau \models \psi_E(\overline{s}^i, \overline{s}^j)$. This pp-formula is satisfiable in $\mathcal{B}$ (by mapping x_j^i to s_j^i), so it is also satisfiable in $\mathcal{B}$ by an assignment g that satisfies $g(x_j^i) \neq g(x_l^k)$ whenever $x_j^i = x_l^k$ is not a literal of ψ_E (and of ϕ_E). Let $\overline{t}^i := (g(x_1^i), \ldots, g(x_d^i))$. Let $\mathcal{T}$ be the structure

induced by $\{\bar{t}^1, \dots, \bar{t}^K\}$ in $\mathcal{B}$. We have a homomorphism $\mathcal{S} \to \mathcal{T}$, since $\mathcal{T}$ satisfies the canonical query of $\mathcal{S}$. If $\mathcal{C}_\Phi^\tau \models \psi_E(\bar{t}^i, \bar{t}^j)$, then $i \neq j$ because $\mathcal{B}$ has no loops. As we have seen above, $\mathcal{C}_\Phi^\tau \models \psi_E(\bar{t}^i, \bar{t}^j) \wedge \bigwedge_{i,j,k,l} t_j^i \neq t_l^k$ where the conjunction ranges over all indices i, j, k, l such that the literal $x_j^i = x_l^k$ is not in ϕ_E. Hence, $\mathcal{C}_\Phi^\tau \models \phi_E(\bar{t}^i, \bar{t}^j)$. Therefore, $\mathcal{T}$ is a weak subgraph of $\mathcal{A}$, which homomorphically maps to $\mathcal{K}_3$. We obtain a homomorphism $\mathcal{S} \to \mathcal{K}_3$.

Thus, $\mathcal{K}_3$ is homomorphically equivalent to a pp-power of $\mathcal{C}_\Phi^\tau$. □

5.3.3 The standard precolouration

Let Φ be an MMSNP sentence in strong normal form with colour set σ, and let Ψ be the following precoloured MMSNP sentence: we obtain Ψ from Φ by adding for each $M \in \sigma$ a new input predicate P_M and adding the conjunct $\neg(P_M(x) \wedge M'(x))$ for each colour $M' \in \sigma \setminus \{M\}$. We call this sentence the *standard precolouration of* Φ.

Theorem 5.31. *Let Φ be an MMSNP sentence in strong normal form with input signature τ. Let Ψ be the standard precolouration of Φ, and let ρ be the input signature of Ψ. Then $\mathcal{C}_\Psi^\rho$ is pp-constructible in $(\mathcal{C}_\Phi^\tau, \neq)$, and $\mathcal{C}_\Phi^\tau$ is pp-constructible in $\mathcal{C}_\Psi^\rho$ (in fact, $\mathcal{C}_\Phi^\tau$ is isomorphic to a reduct of $\mathcal{C}_\Psi^\rho$). Moreover, there exists a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{C}_\Phi^\tau) \xrightarrow{u.c.c.h.} \mathscr{P}$ if, and only if, there exists a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{C}_\Psi^\rho) \xrightarrow{u.c.c.h.} \mathscr{P}$.*

The proof of this theorem will be given in Section 5.3.4. We first point out an immediate consequence.

Corollary 5.32. *Let Φ be an MMSNP sentence in strong normal form, and let Ψ be its standard precolouration. Then Φ and Ψ describe polynomial-time equivalent problems.*

Proof. It is clear that the problem described by Φ reduces to the problem described by Ψ.

We now prove that there is a polynomial-time reduction in the other direction. Let τ and ρ be the input signatures of Φ and Ψ. Since $\mathcal{C}_\Psi^\rho$ is pp-constructible in $(\mathcal{C}_\Phi^\tau, \neq)$ by Theorem 5.31, we have that $\mathrm{CSP}(\mathcal{C}_\Psi^\rho)$ reduces in polynomial-time to $\mathrm{CSP}(\mathcal{C}_\Phi^\tau, \neq)$, by Lemma 2.3. Moreover, by Proposition 5.28, there is a polynomial-time reduction from $\mathrm{CSP}(\mathcal{C}_\Phi^\tau, \neq)$ to $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$. Therefore, $\mathrm{CSP}(\mathcal{C}_\Psi^\rho)$ reduces to $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$. □

5.3.4 Proof of the precolouring theorem

Let $\mathcal{A}$ be a properly coloured $(\tau \cup \sigma)$-structure, i.e., every element appears in the interpretation of precisely one symbol from σ. For an element $a \in A$, denote by $\mathcal{A}[a \mapsto *]$ the structure obtained by uncolouring a. For $M \in \sigma$ and a tuple $\bar{a}$ of elements $\mathcal{A}$, denote by $\mathcal{A}[\bar{a} \mapsto M]$ the structure obtained by uncolouring the elements of $\bar{a}$, and giving them the colour M. Let $C(\mathcal{A}, a)$ be the subset of $\mathcal{C}_\Phi$ containing all elements c such that there exists a homomorphism

$$h \colon \mathcal{A}[a \mapsto *] \to \mathcal{C}_\Phi$$

that satisfies $h(a) = c$. Note that $C(\mathcal{A}, a)$ is, by 1-homogeneity of $\mathcal{C}_\Phi$, a union of colours. So we can also see $C(\mathcal{A}, a)$ as the union of $M^{\mathcal{C}_\Phi}$ for $M \in \sigma$ such that $\mathcal{A}[a \mapsto M]$ is $\mathfrak{F}$-free.

Lemma 5.33. *Suppose that Φ is in strong normal form, and let M be a colour of Φ. Then $M^{\mathcal{C}_\Phi} = \bigcap C(\mathfrak{F}, a)$ where the intersection ranges over all $\mathcal{F} \in \mathfrak{F}$ and $a \in F$ such that $M^{\mathcal{C}_\Phi} \subseteq C(\mathcal{F}, a)$.*

Proof. The left-to-right inclusion is clear. We prove the other inclusion. To do this, it suffices to show that for every $M' \in \sigma \setminus \{M\}$, there exists $\mathcal{G} \in \mathfrak{F}$ and $b \in G$ such that $M^{\mathcal{C}_\Phi} \subseteq C(\mathcal{G}, b)$ but $(M')^{\mathcal{C}_\Phi} \not\subseteq C(\mathcal{G}, b)$. Let $r\colon \sigma \to \sigma$ be defined by $r(M) = M'$ and $r(N) = N$ for all $N \in \sigma \setminus \{M\}$. Since Φ is in strong normal form and r is not surjective, it cannot be a recolouring of Φ. This means that there exists a $\mathfrak{F}$-free structure $\mathcal{A}$ and $\mathcal{F} \in \mathfrak{F}$ such that there exists a homomorphism $h\colon \mathcal{F} \to r(\mathcal{A})$. Let $a_1, \dots, a_k$ be the elements of F that are mapped to $M^{\mathcal{A}}$ by h. In $r(\mathcal{A})$, these elements are in M', so since h is a homomorphism and $\mathcal{F}$ is completely coloured, we have that $a_1, \dots, a_k \in (M')^{\mathcal{F}}$. Moreover, since $\mathcal{A}$ is $\mathfrak{F}$-free, the structure $\mathcal{F}[a_1, \dots, a_k \mapsto M]$ is $\mathfrak{F}$-free. Let $0 \leq j \leq k$ be minimal such that $\mathcal{F}[a_1, \dots, a_j \mapsto M]$ is $\mathfrak{F}$-free. Since $\mathcal{F} \in \mathfrak{F}$, we have $j \geq 1$. Let now $\mathcal{G} \in \mathfrak{F}$ be such that there exists $g\colon \mathcal{G} \to \mathcal{F}[a_1, \dots, a_{j-1} \mapsto M]$, which exists by minimality of j. Note that a_j is in the image of g, otherwise g would be a homomorphism $g\colon \mathcal{G} \to \mathcal{F}[a_1, \dots, a_j \mapsto M]$, in contradiction to the choice of j. Thus, let $b \in G$ be such that $g(b) = a_j$, and note that $b \in (M')^{\mathcal{G}}$, so that $(M')^{\mathcal{C}_\Phi} \not\subseteq C(\mathcal{G}, b)$. Since g is a homomorphism $\mathcal{G}[b \mapsto M] \to F[a_1, \dots, a_j \mapsto M]$, the structure $\mathcal{G}[b \mapsto M]$ is $\mathfrak{F}$-free. This implies that $M^{\mathcal{C}_\Phi} \subseteq C(\mathcal{G}, b)$. We therefore found a $\mathcal{G} \in \mathfrak{F}$ and $b \in G$ such that $M^{\mathcal{C}_\Phi} \subseteq C(\mathcal{G}, b)$ but $(M')^{\mathcal{C}_\Phi} \not\subseteq C(\mathcal{G}, b)$. □

If the sets of the form $C(\mathcal{F}, a)$ were pp-definable in an expansion of $(\mathcal{C}_\Phi, \neq)$ by finitely many constants, we would be done for the proof of Theorem 5.31 since the intersection in Lemma 5.33 is finite. We show how to approximate these sets by pp-definable subsets.

For $M \in \sigma$, let $P(M)$ be the set of pairs $(\mathcal{F}, a)$ such that $M^{\mathcal{C}_\Phi} \subseteq C(\mathcal{F}, a)$. Let $(\mathcal{F}, a) \in P(M)$. Let $a_1, \dots, a_k$ be the elements of $\mathcal{F}$ that are distinct from a. Let $\phi_{\mathcal{F}}(a, a_1, \dots, a_k)$ be the canonical query of $\mathcal{F}^\tau$. Let $M_1, \dots, M_k$ be the colours of these elements in $\mathcal{F}$. Fix the formula

$$\psi_{\mathcal{F},a}(x, U_1, \dots, U_k) := \exists y_1, \dots, y_k \left(\phi_{\mathcal{F}}(x, y_1, \dots, y_k) \wedge \bigwedge_{i \in \{1,\dots,k\}} U_i(y_i) \right),$$

in the language $\tau \cup \{U_1, \dots, U_k\}$. Let $\chi_M^{(0)}$ be $M(x)$. We define $\chi_M^{(n)}$ inductively. For $n \geq 0$, let

$$\chi_M^{(n+1)}(x) := \bigwedge_{(\mathcal{F},a) \in P(M)} \psi_{\mathcal{F},a}(x, \chi_{M_1}^{(n)}, \dots, \chi_{M_k}^{(n)}).$$

Lemma 5.34. *For any $n \in \mathbb{N}$ and $M \in \sigma$ the formula $\chi_M^{(n)}(x)$ defines $M^{\mathcal{C}_\Phi}$ over $\mathcal{C}_\Phi$.*

Proof. We prove the result by induction, the case $n = 0$ being trivial. Suppose that the result is proved for some $n \geq 0$. From Lemma 5.33 and the induction hypothesis follows that $\chi_M^{(n+1)}(x)$ defines a subset of $M^{\mathcal{C}_\Phi}$, so we just have to prove that the formula is satisfiable (then by 1-homogeneity of $\mathcal{C}_\Phi$, we get that $\chi_M^{(n+1)}$ defines $M^{\mathcal{C}_\Phi}$). By Lemma 5.8, if $\chi_M^{(n+1)}$ is not satisfiable then there must exist $(\mathcal{F}, a) \in P(M)$ such that $\psi_{\mathcal{F},a}(x, \chi_{M_1}^{(n)}, \dots, \chi_{M_k}^{(n)})$ is not satisfiable, i.e.,

$$\phi_{\mathcal{F}}(x, y_1, \dots, y_k) \wedge \bigwedge_{i \in \{1,\dots,k\}} \chi_{M_i}^{(n)}(y_i)$$

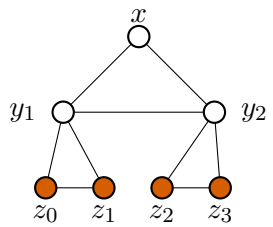

Figure 5.2: Illustration of the formula $\chi_M^{(2)}(x)$, for the MMSNP sentence of Example 11. All the variables except for x are existentially quantified.

is not satisfiable, where $M_1, \ldots, M_k$ are the colours in $\mathcal{F}$ of the elements other than a. By Lemma 5.8 again, and since $\phi_{\mathcal{F}}(x, y_1, \ldots, y_k)$ is clearly satisfiable, there must exist $i \in \{1, \ldots, k\}$ such that $\chi_{M_i}^{(n)}(y_i)$ is not satisfiable, in contradiction to our induction hypothesis. Therefore, $\chi_M^{(n+1)}$ is satisfiable. □

Example 11. We show in Figure 5.2 the construction of the formula $\chi_M^{(2)}$ in the case of the MMSNP sentence given by the obstructions in Figure 5.1, where M is represented with round vertices. Note that if $\mathcal{F}$ is the triangle with coloured square vertices in Figure 5.1 and a is a vertex of this triangle then $C(\mathcal{F}, a) = M^{\mathcal{C}_\Phi}$. Note that each y_i must be coloured with a square vertex (otherwise the triangle with coloured round vertices would appear), so that x necessarily belongs to $M^{\mathcal{C}_\Phi}$. This shows that $\chi_M^{(2)}(x)$ defines a subset of $M^{\mathcal{C}_\Phi}$.

Let $n > |\Phi|$. It is a consequence of Lemma 5.34 that for each $M \in \sigma$, the formula $\chi_M^{(n)}(x)$ is satisfiable in $\mathcal{C}_\Phi$. Let $\mathcal{A}$ be the canonical query of $\chi_M^{(n)}(x)$ where we additionally colour the elements of $\mathcal{A}$ according to an arbitrary satisfying assignment for $\chi_M^{(n)}$. Then $\mathcal{A}$ homomorphically maps to $\mathcal{C}_\Phi^\tau$, so by Lemma 5.10 it also injectively maps to $\mathcal{C}_\Phi^\tau$. We deduce from this that $\chi_M^{(n)}$ is satisfiable by an injective assignment h. For every $M' \in \sigma$, replace in $\chi_M^{(n)}$ each literal $M'(y)$ (the vertices at the bottom level, in Figure 5.2) by the literal $y = h(y)$. The resulting formula, $\tilde{\chi}_M(x)$, is then a pp-formula in an expansion of $\mathcal{C}_\Phi^\tau$ by finitely many constants $\bar{c}$.

Lemma 5.35. *The formula $\tilde{\chi}_M(x)$ defines a subset of $M^{\mathcal{C}_\Phi}$ in $(\mathcal{C}_\Phi^\tau, \bar{c})$.*

Proof. Immediate from Lemma 5.34 and the definition of $\tilde{\chi}_M$. □

We claim that the formulas $\tilde{\chi}$ define a universal substructure of $\mathcal{C}_\Phi$, in the sense that any structure $\mathcal{A}$ that has a homomorphism to $\mathcal{C}_\Phi$ has a homomorphism h to $\mathcal{C}_\Phi$ such that $\mathcal{C}_\Phi \models \tilde{\chi}_M(h(a))$ for every $a \in M^{\mathcal{A}}$.

Proposition 5.36. *Let $\mathcal{A}$ be a finite structure that has a homomorphism to $\mathcal{C}_\Phi$, and let $\phi_{\mathcal{A}}(a_1, \ldots, a_k)$ be the canonical query of $\mathcal{A}$. Let M_i be the colour of a_i in $\mathcal{A}$. Let $n > |\Phi|$. Then the formula*

$$\phi_{\mathcal{A}}(x_1, \ldots, x_k) \wedge \bigwedge_{1 \leq i \leq k} \tilde{\chi}_{M_i}(x_i)$$

is satisfiable in $(\mathcal{C}_\Phi, \bar{c})$.

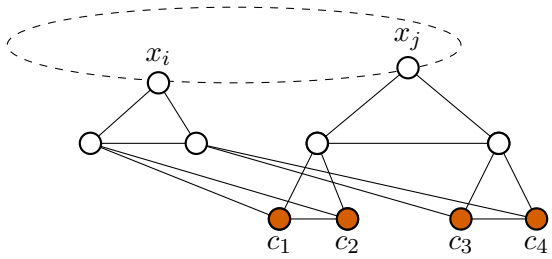

Figure 5.3: Depiction of the canonical database $\mathcal{A}'$ of the formula ρ in the proof of Proposition 5.36. The vertices on the ellipse are the elements $\overline{x}$ of $\mathcal{A}$. The vertices at the bottom are the variables $\overline{y}$. The only variables shared between different $\tilde{\chi}$ formulas are the variables $\overline{y}$.

See Figure 5.3 for an illustration.

Proof. Let $\psi(\overline{x}) \wedge \theta(\overline{x})$ be a formula describing the orbit of the tuple $\overline{c}$ in $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$ where $\psi(\overline{x})$ is a pp-formula in the language of $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$ and $\theta(\overline{x})$ is a conjunction of negated atomic formulas (that such a formula exists is a consequence of Lemma 5.11). Let $\overline{y}$ be a tuple of fresh variables with the same length as $\overline{c}$. We prove that the formula

$$\rho(\overline{x}, \overline{y}) := \phi_{\mathcal{A}}(x_1, \ldots, x_k) \wedge \bigwedge_{1 \leq i \leq k} \tilde{\chi}_{M_i}(x_i, \overline{y}) \wedge \psi(\overline{y})$$

is satisfiable in $\mathcal{C}_\Phi$, where we modified the formulas $\tilde{\chi}$ by replacing every constant symbol in them by the corresponding y variable.

Suppose that ρ is not satisfiable, and let $\mathcal{A}'$ be its canonical database. Therefore, there exists $\mathcal{F} \in \mathfrak{F}$ and a homomorphism $h \colon \mathcal{F} \to \mathcal{A}'$. Since $\mathcal{F}$ is connected, the image of h cannot contain both vertices from $\overline{x}$ and vertices from $\overline{y}$, because the shortest path between an x variable and a y variable is at least n, which has been chosen to be greater than the number of elements of $\mathcal{F}$. Suppose that the image of h does not contain any y variable (in Figure 5.3, this means that the image of h does not touch any node at the bottom of the picture). Note that if one removes the variables $\overline{y}$, each x_i becomes an articulation point (i.e., removing x_i disconnects the structure, for any i). By applying Lemma 5.8 at each x_i, we obtain that at least one of $\phi_{\mathcal{A}}$ or the canonical database of some formula $\tilde{\chi}$ cannot be $\mathfrak{F}$-free, which is a contradiction because the formulas $\tilde{\chi}$ are satisfiable by Lemma 5.34 and $\phi_{\mathcal{A}}$ is satisfiable as well.

If the image of h does not contain any of $x_1, \ldots, x_k$, we immediately obtain a contradiction because $\overline{c}$ satisfies

$$\psi(\overline{c}) \wedge \bigwedge_{1 \leq i \leq k} \exists x_i (\tilde{\chi}_{M_i}(x_i, \overline{c})).$$

Whence, let h be an embedding of $\mathcal{A}'$ into $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$. Since $h(\overline{y})$ satisfies ψ and h is an embedding, $h(\overline{y})$ satisfies $\psi \wedge \theta$, which implies that $\overline{c}$ and $h(\overline{y})$ are in the same orbit in $\mathcal{B}^{\text{ind}}_{\mathfrak{F}}$. Without loss of generality, we can assume that $h(\overline{y}) = \overline{c}$. Let g be any injective homomorphism $\mathcal{B}^{\text{ind}}_{\mathfrak{F}} \to \mathcal{C}_\Phi$. The restriction of g to $\mathcal{C}_\Phi \subseteq \mathcal{B}^{\text{ind}}_{\mathfrak{F}}$ is an embedding, since $(\mathcal{C}_\Phi, \neq)$ is a model-complete core. Therefore, $(g \circ h)(\overline{y})$ and $\overline{c}$ are in the same orbit, and

without loss of generality we can assume that $(g \circ h)(\overline{y}) = \overline{c}$. In conclusion, $(g \circ h)|_{\{x_1,\dots,x_k\}}$ is a satisfying assignment to the formula in the statement of the proposition. □

Proof of Theorem 5.31. We first show that $\mathcal{C}^{\rho}_{\Psi}$ is pp-constructible in $(\mathcal{C}^{\tau}_{\Phi}, \neq)$. Let $\mathcal{D}$ be the expansion with signature ρ of the structure $\mathcal{C}^{\tau}_{\Phi}$ such that for every color $M \in \sigma$ of Φ the symbol $P_M \in \rho$ denotes the relation defined by the formula $\tilde{\chi}_M$ from Lemma 5.35. Since $(\mathcal{C}^{\tau}_{\Phi}, \neq)$ is a model-complete core and $\mathcal{D}$ is pp-definable in $\mathcal{C}^{\tau}_{\Phi}$ after having added finitely many constants, we obtain that $\mathcal{D}$ is pp-constructible from $(\mathcal{C}^{\tau}_{\Phi}, \neq)$. Hence, it suffices to show that $\mathcal{D}$ and $\mathcal{C}^{\rho}_{\Psi}$ are homomorphically equivalent. We first show that $\mathcal{D}$ satisfies Ψ. Consider the expansion of $\mathcal{D}$ where $M \in \sigma$ denotes $M^{\mathcal{C}_{\Phi}}$. This expansion satisfies for distinct $M, M' \in \sigma$ the clause $\forall x. \neg(P_M(x) \wedge M'(x))$ of Ψ as a consequence of Lemma 5.35. The expansion clearly satisfies all other conjuncts of Ψ. Therefore, $\mathcal{D}$ satisfies Ψ and we obtain a homomorphism $\mathcal{D} \to \mathcal{C}^{\rho}_{\Psi}$. Conversely, Proposition 5.36 gives that every finite substructure of $\mathcal{C}^{\rho}_{\Psi}$ has a homomorphism to $\mathcal{D}$. By the ω-categoricity of $\mathcal{D}$, we get a homomorphism from $\mathcal{C}^{\rho}_{\Psi}$ to $\mathcal{D}$.

To prove that $\mathcal{C}^{\tau}_{\Phi}$ is pp-constructible in $\mathcal{C}^{\rho}_{\Psi}$, it suffices to note that the structures $\mathcal{C}^{\tau}_{\Phi}$ and $\mathcal{C}^{\tau}_{\Psi}$ are isomorphic (since $(\mathcal{C}^{\tau}_{\Phi}, \neq)$ and $(\mathcal{C}^{\tau}_{\Psi}, \neq)$ are model-complete cores and have the same CSP), and that $\mathcal{C}^{\tau}_{\Psi}$ is obtained from $\mathcal{C}^{\rho}_{\Psi}$ by dropping the relations from $\rho \setminus \tau$, and is in particular a pp-power of $\mathcal{C}^{\rho}_{\Psi}$.

These pp-constructions give uniformly continuous clonoid homomorphisms $\mathrm{Pol}(\mathcal{C}^{\rho}_{\Psi}) \to \mathrm{Pol}(\mathcal{C}^{\tau}_{\Phi})$ and $\mathrm{Pol}(\mathcal{C}^{\tau}_{\Phi}, \neq) \to \mathrm{Pol}(\mathcal{C}^{\rho}_{\Psi})$ (Theorem 2.14). From the former homomomorphism we get that if there is a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{C}^{\tau}_{\Phi}) \to \mathscr{P}$, there is also one $\mathrm{Pol}(\mathcal{C}^{\rho}_{\Psi}) \to \mathscr{P}$. The latter homomorphism gives us that if there exists a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{C}^{\rho}_{\Psi}) \to \mathscr{P}$, there is one $\mathrm{Pol}(\mathcal{C}^{\tau}_{\Phi}, \neq) \to \mathscr{P}$. We conclude by Proposition 5.29. □

5.4 An Algebraic Dichotomy for MMSNP

We prove in this section that MMSNP exhibits a complexity dichotomy, that is, that every problem in MMSNP is in P or NP-complete. Moreover, we show that the tractability border can be described in terms of clonoid homomorphisms to $\mathscr{P}$, thus confirming Conjecture 1 for the class of constraint satisfaction problems in MMSNP.

Theorem 5.37. *Let $\mathcal{B}$ be an ω-categorical structure such that* $\mathrm{CSP}(\mathcal{B})$ *is in MMSNP. Then exactly one of the following holds:*

(i) there is no uniformly continuous clonoid homomorphism from $\mathrm{Pol}(\mathcal{B})$ *to $\mathscr{P}$, and* $\mathrm{CSP}(\mathcal{B})$ *is solvable in polynomial time,*

(ii) there is a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{B}) \xrightarrow{u.c.c.h.} \mathscr{P}$*, and* $\mathrm{CSP}(\mathcal{B})$ *is NP-complete.*

We briefly describe the road to proving Theorem 5.37. In virtue of Theorem 5.31 and Corollary 5.32, it suffices to focus on the case that $\mathrm{CSP}(\mathcal{B})$ is described by a precoloured MMSNP sentence. For each precoloured sentence Φ, we consider the structure $\mathcal{C}^{\tau}_{\Phi}$ whose CSP is described by Φ. We prove that the complexity of $\mathrm{CSP}(\mathcal{C}^{\tau}_{\Phi})$ and the existence of a clonoid homomorphism $\mathrm{Pol}(\mathcal{C}^{\tau}_{\Phi}) \xrightarrow{u.c.c.h.} \mathscr{P}$ are determined by the existence of a clone

homomorphism $\mathscr{C} \to \mathscr{P}$, where $\mathscr{C}$ is the subset of $\mathrm{Pol}(\mathcal{C}_\Phi^\tau)$ that contains the functions that are canonical with respect to $(\mathcal{C}_\Phi, <)$.

From now on, we fix a precoloured MMSNP sentence Φ with coloured obstruction set $\mathfrak{F}$, input signature τ, and colour signature σ.

We finish this section by stating a consequence of assuming that Φ is precoloured and in normal form on the set $\mathscr{C}_1^{\mathrm{typ}}$.

Proposition 5.38. *Let Φ be a precoloured MMSNP sentence in normal form. Let $\mathscr{C}$ be the set of polymorphisms of $\mathcal{C}_\Phi^\tau$ that are canonical with respect to $(\mathcal{C}_\Phi, <)$. Then all functions in $\mathscr{C}_1^{\mathrm{typ}}$ are idempotent.*

Proof. The orbits of $\mathrm{Aut}(\mathcal{C}_\Phi)$ are in one-to-one correspondence with the colours from Φ (by Corollary 5.17 since Φ is in normal form). Since Φ is precoloured and by Lemma 5.27, the symbols $P_M \in \tau$ and $M \in \sigma$ have the same interpretation in $\mathcal{C}_\Phi$. This implies that all polymorphisms of $\mathcal{C}_\Phi^\tau$ (and in particular, the ones that are canonical with respect to $(\mathcal{C}_\Phi, <)$) preserve the orbits of elements of $\mathcal{C}_\Phi^\tau$. Therefore, every function in $\mathscr{C}_1^{\mathrm{typ}}$ is idempotent. □

5.4.1 The tractable case

In this section, we prove that $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$ is polynomial-time tractable, under the assumption that $\mathcal{C}_\Phi^\tau$ has a polymorphism that is canonical with respect to $(\mathcal{C}_\Phi, <)$ and whose behaviour on orbits of elements is Siggers. For that we use Corollary 3.11.

Proposition 5.39. *Let $\mathscr{C}$ be the clone of functions in $\mathrm{Pol}(\mathcal{C}_\Phi^\tau)$ that are canonical with respect to $(\mathcal{C}_\Phi, <)$. Suppose that $\mathscr{C}_1^{\mathrm{typ}}$ does not have a clonoid homomorphism to $\mathscr{P}$. Then $\mathrm{Pol}(\mathcal{C}_\Phi)$ contains an operation that is pseudo-Siggers modulo $\overline{\mathrm{Aut}(\mathcal{C}_\Phi, <)}$ and canonical with respect to $(\mathcal{C}_\Phi, <)$.*

Proof. Since $\mathscr{C}_1^{\mathrm{typ}}$ does not have a clonoid homomorphism to $\mathscr{P}$, Theorem 2.17 (1. $\Rightarrow$ 5.) implies that there exists an $f \in \mathscr{C}$ such that $\xi_1^{\mathrm{typ}}(f)$ is Siggers in $\mathscr{C}_1^{\mathrm{typ}}$. It will be convenient to use the notation $\pi(a, b, a, c, b, c) := (b, a, c, a, c, b)$. Let $\mathcal{A}$ be the $(\tau \cup \sigma \cup \{<\})$-structure obtained from $(\mathcal{C}_\Phi^\tau)^6$ as follows.

The colours and precolours. For $M_0, M_1, \ldots, M_6 \in \sigma$ and $(a_1, \ldots, a_6) \in A$ such that $a_i \in M_i^{\mathcal{C}_\Phi}$ for all $i \in \{1, \ldots, 6\}$ and $\xi_1^{\mathrm{typ}}(f)(M_1, \ldots, M_6) = M_0$, declare that $(a_1, \ldots, a_6) \in A$ is in $M_0^{\mathcal{A}}$ and in $P_{M_0}^{\mathcal{A}}$.

The order. Let B be the domain of $\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}$. Let $s\colon (B, <)^6 \to (B, <)$ be an injective map that is pseudo-Siggers modulo $\overline{\mathrm{Aut}(B, <)}$. Such a map can be constructed by considering the digraph on B^6 with arcs

$$\{((x, y, x, z, y, z), \pi(x, y, x, z, y, z)) \mid x, y, z \in \mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}\}.$$

Note that this graph is a disjoint union of arcs and loops. Let $<$ be any linear order on B^6 such that if (u_1, v_1) and (u_2, v_2) are arcs then $u_1 < u_2$ if and only if $v_1 < v_2$ (it is easy to see that such a linear order exists for any directed graph without cycles and with outdegree and indegree at most one). This linear order embeds into $(B, <)$ and gives the desired injective map. Declare that $(a_1, \ldots, a_6) < (b_1, \ldots, b_6)$ holds in $\mathcal{A}$ iff $s(a_1, \ldots, a_6) < s(b_1, \ldots, b_6)$. Since s is injective, this defines a linear order on $\mathcal{A}$.

The structure $\mathcal{A}$ is linearly ordered, satisfies Φ, and all its elements are precoloured, so $\mathcal{A}$ embeds into the $(\tau \cup \sigma \cup \{<\})$-reduct $\mathcal{D}$ of $\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}$, via a map $e\colon \mathcal{A} \hookrightarrow \mathcal{D}$. By Corollary 5.22, we can assume that e is canonical from $(\mathcal{C}_\Phi, <)$ to $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}, <)$. There is an injective homomorphism h from $\mathcal{D}$ to $\mathcal{C}_\Phi$, and again we can pick h to be canonical from $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}; <)$ to $(\mathcal{C}_\Phi, <)$. It is clear that $f' := h \circ e$ is canonical with respect to $(\mathcal{C}_\Phi, <)$. We claim that it is pseudo-Siggers modulo $\overline{\mathrm{Aut}(\mathcal{C}_\Phi, <)}$.

We have to show that for all $m \in \mathbb{N}$ and all $\overline{a}^1, \dots, \overline{a}^m \in A^6$ the m-tuples $(f'(\overline{a}^1), \dots, f'(\overline{a}^m))$ and $(f'(\pi\overline{a}^1), \dots, f'(\pi\overline{a}^m))$ lie in the same orbit of $\mathrm{Aut}(\mathcal{C}_\Phi, <)$. Since h is canonical, it suffices to prove that $(e(\overline{a}^1), \dots, e(\overline{a}^m))$ and $(e(\pi\overline{a}^1), \dots, e(\pi\overline{a}^m))$ lie in the same orbit in $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}, <)$. By the homogeneity of $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}, <)$ we have to prove that the two tuples satisfy the same atomic formulas in $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}, <)$. Suppose that $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}, <) \models R(e(\overline{a}^1), \dots, e(\overline{a}^s))$ for an s-ary relation symbol $R \in \tau \cup \sigma \cup \{<\}$. Since e is an embedding this means that $R(\overline{a}^1, \dots, \overline{a}^s)$ also holds in $\mathcal{A}$. If $R \in \tau$ then the definition of $\mathcal{A}$ implies that for all $i \in \{1, \dots, 6\}$, we have $\mathcal{C}_\Phi \models R(a_i^1, \dots, a_i^s)$. This immediately implies that $\mathcal{A} \models R(\pi\overline{a}^1, \dots, \pi\overline{a}^s)$, so that by applying e we obtain $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}, <) \models R(f'(\pi\overline{a}^1), \dots, f'(\pi\overline{a}^s))$. Consider now the case that R is a symbol M' from σ (so that $s = 1$). By the definition of $\mathcal{A}$ this implies that the entries of $\overline{a}^1 = (a, b, a, c, b, c)$ are such that $a \in {M_1}^{\mathcal{C}_\Phi}, b \in {M_2}^{\mathcal{C}_\Phi}, c \in {M_3}^{\mathcal{C}_\Phi}$ for $M_1, M_2, M_3 \in \sigma$ and

$$\xi_1^{\mathrm{typ}}(f)(M_1, M_2, M_1, M_3, M_2, M_3) = M'.$$

Since $\xi_1^{\mathrm{typ}}(f)$ is Siggers, we also have

$$\xi_1^{\mathrm{typ}}(f)(M_2, M_1, M_3, M_1, M_3, M_2) = M'.$$

Therefore, we also get that $\pi\overline{a}^1 = (b, a, c, a, c, b)$ belongs to $(M')^{\mathcal{A}}$, so that $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}, <) \models M'(f'(\pi\overline{a}^1))$. Finally, if R is the order symbol, it means that $\overline{a}^1 < \overline{a}^2$ holds in $\mathcal{A}$. By definition, this is true if and only if $s(\overline{a}^1) < s(\overline{a}^2)$. Since s is pseudo-Siggers modulo $(B, <)$, we have $s(\pi\overline{a}^1) < s(\pi\overline{a}^2)$, so that $\mathcal{A} \models \pi\overline{a}^1 < \pi\overline{a}^2$. Finally, composing with e gives that $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{HN}}, <) \models e(\pi\overline{a}^1) < e(\pi\overline{a}^2)$. □

In order to use Corollary 3.11, it remains to prove that $\mathcal{C}_\Phi^\tau$ is a reduct of a finitely bounded homogeneous structure, which we now show in a series of lemmas.

Proposition 5.40. *The structure $\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}$ has a homogeneous expansion by* finitely many *pp-definable relations. Moreover, the expansion is finitely bounded.*

Proof. Let m be the size of the largest structure in $\mathfrak{F}$. We show that the expansion of $\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}$ by all relations with a pp-definition having at most m variables (free or existentially quantified) is homogeneous. Since we assume that pp-formulas are in prenex normal form, there is only a bounded number of such formulas. Let t_1 and t_2 be two n-tuples of $\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}$ with pairwise distinct entries such that t_1 and t_2 lie in different orbits. Since $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, \neq)$ is a model-complete core, the orbits of t_1 and of t_2 are pp-definable, and hence there are pp-formulas ϕ_1 and ϕ_2 such that $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, \neq) \models \phi_1(t_1)$ and $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, \neq) \models \phi_2(t_2)$ but $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, \neq) \not\models \exists x_1, \dots, x_n(\phi_1(\bar{x}) \wedge \phi_2(\bar{x}))$. So there exists a structure $\mathcal{A} \in \mathfrak{F}$ that homomorphically embeds into the canonical database of $\phi_1(\bar{x}) \wedge \phi_2(\bar{x})$. But then $\phi_i(\bar{x})$ must imply for $i = 1$ and $i = 2$ a pp-formula ψ_i with at most m variables such that $\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}} \not\models \exists x_1, \dots, x_n(\psi_1(\bar{x}) \wedge \psi_2(\bar{x}))$. Hence, the orbits of injective tuples are determined

by pp-definition having at most m variables, showing homogeneity of the expansion by the relations defined by those formulas.

We finally claim that this expansion is finitely bounded. Clearly, we still have the bounds $\mathfrak{F}$. Additionally, for every pp-formula $\phi(x_1, \ldots, x_k)$ with at most m variables and every k-ary relation symbol R introduced for an inequivalent pp-formula, we have the canonical database of $\phi(x_1, \ldots, x_n) \wedge R(x_1, \ldots, x_n)$ as a new bound. These are finitely many bounds, and they fully describe the expansion, showing the claim. □

Corollary 5.41. *The structure $\mathcal{C}_\Phi$ has a homogeneous expansion by finitely many pp-definable relations. Moreover, the expansion is finitely bounded.*

Proof. By Proposition 5.40, $\mathcal{B}^{\text{hom}}_{\mathfrak{F}}$ has a homogeneous finitely bounded expansion $\mathcal{B}$ by pp-definable relations. The restriction of $\mathcal{D}$ to the coloured elements is still homogeneous, and has the additional bounds excluding all finite one-element structures that are not coloured, so it is finitely bounded, too. □

Theorem 5.42. *If there is no clone homomorphism $\mathscr{C}_1^{\text{typ}} \to \mathscr{P}$, then* $\text{CSP}(\mathcal{C}_\Phi^\tau)$ *is in* P.

Proof. Proposition 5.40 gives a finitely bounded homogeneous expansion $\mathcal{D}$ of $\mathcal{C}_\Phi$ by pp-definable relations, so $\text{Pol}(\mathcal{D}) = \text{Pol}(\mathcal{C}_\Phi)$. Proposition 4.13 states that $\text{Pol}(\mathcal{C}_\Phi^\tau)$ contains an operation that is pseudo-Siggers modulo $\overline{\text{Aut}(\mathcal{C}_\Phi)} = \overline{\text{Aut}(\mathcal{D})}$ and that is canonical with respect to $\mathcal{C}_\Phi$ (and therefore also with respect to $\mathcal{D}$). By Corollary 3.11, $\text{CSP}(\mathcal{C}_\Phi^\tau)$ is in P. □

5.4.2 The hard case

Let Φ be a precoloured MMSNP sentence and let $\mathscr{C}$ be the clone of polymorphisms of $\mathcal{C}_\Phi^\tau$ that are canonical with respect to $(\mathcal{C}_\Phi, <)$. In this section, we deal with the case that there exists a clone homomorphism $\xi\colon \mathscr{C}_1^{\text{typ}} \to \mathscr{P}$, and prove that there exists a uniformly continuous clonoid homomorphism $\phi\colon \text{Pol}(\mathcal{C}_\Phi^\tau) \xrightarrow{u.c.c.h.} \mathscr{P}$.

As in Chapter 4, we use the canonisation property for $(\mathcal{C}_\Phi, <)$ in order to give a candidate for ϕ. However, we were not able to prove that $(\mathcal{C}_\Phi, <)$ has the mashup property, so that we need other tools to prove that ϕ is well-defined. We develop these tools now.

Let ρ be a subset of σ such that ρ is preserved by $\mathscr{C}_1^{\text{typ}}$ (we identify the relation symbols with the domain of $\mathscr{C}_1^{\text{typ}}$). Let Θ be an equivalence relation on ρ that is preserved by $\mathscr{C}_1^{\text{typ}}$ and with two equivalence classes $S, T \subseteq \rho$. We call $\{S, T\}$ a *subfactor* of $\mathscr{C}_1^{\text{typ}}$. The clone $\mathscr{C}_1^{\text{typ}}$ naturally induces a clone on the two-element set $\{S, T\}$. If this clone is (isomorphic to) the projection clone $\mathscr{P}$, then we call $\{S, T\}$ a *trivial subfactor*. Proposition 4.1 implies that $\mathscr{C}_1^{\text{typ}}$ has a clone homomorphism to $\mathscr{P}$ if, and only if, $\mathscr{C}_1^{\text{typ}}$ has a trivial subfactor $\{S, T\}$. Note that if $\{S, T\}$ is a subfactor of $\mathscr{C}_1^{\text{typ}}$, the subset $S^{\mathcal{C}_\Phi} \cup T^{\mathcal{C}_\Phi}$ of $\mathcal{C}_\Phi$ is preserved by every operation in $\mathscr{C}$ (where we write $S^{\mathcal{C}_\Phi}$ for $\bigcup_{R \in S} R^{\mathcal{C}_\Phi}$ and similarly for $T^{\mathcal{C}_\Phi}$).

Let X be a pp-definable subset of $\mathcal{C}_\Phi$. A binary symmetric relation $N \subseteq X^2$ defines an undirected graph on σ: there is an edge between M and M' iff there exist $x \in M^{\mathcal{C}_\Phi}$ and $y \in M'^{\mathcal{C}_\Phi}$ such that $(x, y) \in N$. If N is pp-definable in $\mathcal{C}_\Phi$, we call the resulting graph on σ a *definable colour graph over* X. In the following technical propositions, we prove that the existence of a trivial subfactor $\{S, T\}$ of $\mathscr{C}_1^{\text{typ}}$ implies the existence of definable colour graphs with an edge from S to T and without loops (Proposition 5.45). Refining this even further, we show the existence of such a graph whose connected components are of three

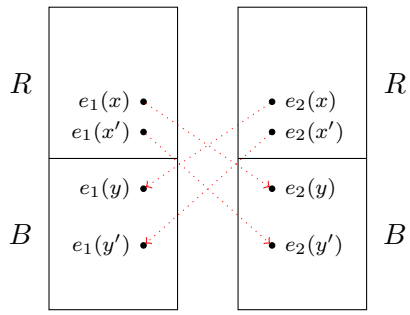

Figure 5.4: Illustration of Lemma 5.43. The four pairs shown by directed edges are in the same orbit under $\mathrm{Aut}(\mathcal{C}_\Phi, <)$.

types: contained in S, contained in T, and bipartite with the bipartition being induced by S and T (Proposition 5.46).

Lemma 5.43. *For every pair of colours $R, B \in \sigma$, there are endomorphisms e_1 and e_2 of $\mathcal{C}_\Phi$ such that for all $(x_1, x_2), (y_1, y_2) \in R^{\mathcal{C}_\Phi} \times B^{\mathcal{C}_\Phi}$, the pairs $(e_1(x_1), e_2(x_2))$, $(e_1(y_1), e_2(y_2))$, $(e_2(x_1), e_1(x_2))$, and $(e_2(y_1), e_1(y_2))$ are in the same orbit under $\mathrm{Aut}(\mathcal{C}_\Phi, <)$.*

For an illustration, see Figure 5.4.

Proof. We build the endomorphisms by compactness, showing that partial homomorphisms with the given properties exist for every finite substructure $\mathcal{F}$ of $\mathcal{C}_\Phi$. Let $\mathcal{G}$ be the disjoint union of 2 copies of $\mathcal{F}$, with domain $F \times \{1, 2\}$. We prepare a new structure $\mathcal{H}$ which contains $\mathcal{G}$ as a substructure. For all elements x and x' of $\mathcal{G}$ of the same colour, take a fresh copy $\mathcal{G}'$ of $\mathcal{G}$ and add to $\mathcal{H}$ this fresh copy, where the vertex corresponding to x in $\mathcal{G}'$ is glued on top of the vertex corresponding to x' in the original copy of $\mathcal{G}$. This way, every two elements of the original $\mathcal{G}$ that are in the same colour satisfy the same pp-formulas in $\mathcal{H}$. It is also clear that $\mathcal{H}$ is $\mathfrak{F}$-free, since Φ is in normal form. Since $\mathcal{H}$ is $\mathfrak{F}$-free, the expansion $\mathcal{H}^*$ of $\mathcal{H}$ by all relations with a pp-definition with at most m variables embeds into $\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}$ (where m denotes the size of the largest structure in $\mathfrak{F}$).

Let $<$ be any linear order on $\mathcal{G}$ such that $(x, 1) < (y, 2)$ and $(x, 2) < (y, 1)$ for all $x \in R^{\mathcal{F}}$ and $y \in B^{\mathcal{F}}$. Complete $<$ arbitrarily into a linear order on $\mathcal{H}$, so that there exists an embedding e of $(\mathcal{H}^*, <)$ into $(\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}, <)$. By the homogeneity of $(\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}, <)$, the pairs

$$(e(x, 1), e(y, 2))$$
$$(e(x', 1), e(y', 2))$$
$$(e(x, 2), e(y, 1))$$

are all in the same orbit in $(\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}, <)$, for all $x, x' \in R^{\mathcal{F}}$ and $y, y' \in B^{\mathcal{F}}$. Let $e' \colon \mathcal{G} \to \mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ be obtained by composing e with an injective homomorphism of the $(\tau \cup \sigma)$-reduct of $\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}$ to $\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}$ that is canonical from $(\mathcal{B}^{\mathrm{HN}}_{\mathfrak{F}}, <)$ to $(\mathcal{B}^{\mathrm{hom}}_{\mathfrak{F}}, <)$ (we use Theorem 5.21 and Theorem 5.20). Since all the vertices of $\mathcal{G}$ are coloured, the image of e' is included in $\mathcal{C}_\Phi$. We obtain a homomorphism h from $\mathcal{G}$ to $\mathcal{C}_\Phi$ such that the given pairs are in the same orbit under $\mathrm{Aut}(\mathcal{C}_\Phi, <)$. For $i \in \{1, 2\}$, define the partial endomorphisms e_i of $\mathcal{C}_\Phi$ by $x \mapsto h(x, i)$. It is easy to check that these partial endomorphisms satisfy the required properties. $\square$

In the following proof, we need a slightly different notion of canonicity. A function $f\colon A^k \to A$ is said to be *diagonally canonical* with respect to $\mathcal{A}$ if for all m-tuples $\bar{t}^1, \dots, \bar{t}^m$ and every automorphism $\alpha \in \mathrm{Aut}(\mathcal{A})$, there exists $\beta \in \mathrm{Aut}(\mathcal{A})$ such that $\beta f(\bar{t}^1, \dots, \bar{t}^m) = f(\alpha\bar{t}^1, \dots, \alpha\bar{t}^m)$. The results from [36] and the fact that $(\mathcal{C}_\Phi, <)$ is a Ramsey structure (Corollary 5.22) imply the following.

Theorem 5.44. *Let $f \in \mathrm{Pol}(\mathcal{C}_\Phi)$. There exists a polymorphism $g \in \mathrm{Pol}(\mathcal{C}_\Phi)$ that is diagonally canonical with respect to $(\mathcal{C}_\Phi, <)$ and $g \in \overline{\{\beta f(\alpha, \dots, \alpha) \mid \alpha, \beta \in \mathrm{Aut}(\mathcal{C}_\Phi, <)\}}$.*

Proposition 5.45. *Let Φ be a precoloured MMSNP sentence in strong normal form and let $\mathscr{C}$ be the clone of polymorphisms of $\mathcal{C}_\Phi$ that are canonical with respect to $(\mathcal{C}_\Phi, <)$. Let $\{S, T\}$ be a trivial subfactor of $\mathscr{C}_1^{\mathrm{typ}}$. Then for every pp-definable subset X of $\mathcal{C}_\Phi$ such that $X \cap S^{\mathcal{C}_\Phi} \neq \emptyset$ and $X \cap T^{\mathcal{C}_\Phi} \neq \emptyset$, there exists a loopless definable colour graph over X containing an edge from S to T.*

Proof. Let $X \subseteq \mathcal{C}_\Phi$ be pp-definable and such that $X \cap S^{\mathcal{C}_\Phi}$ and $X \cap T^{\mathcal{C}_\Phi}$ are non-empty. We prove the result by contradiction, assuming that every definable colour graph over X that contains an edge from S to T also contains a loop. The crux of the proof is to show that this assumption implies the existence of a canonical polymorphism $\hat{h}$ of $\mathcal{C}_\Phi$ such that for all $x, y \in X$ the equivalence $\hat{h}(x, y) \in S^{\mathcal{C}_\Phi} \Leftrightarrow \hat{h}(y, x) \in S^{\mathcal{C}_\Phi}$ holds. This argument is similar to the proof of Lemma 4.4 in [4].

First, we show that for every finite subset A of $\mathcal{C}_\Phi$, there exists a binary polymorphism f of $\mathcal{C}_\Phi$ such that the following property (†) holds for all $a, b \in A \cap X$:

$$f(a, b), f(b, a) \in S^{\mathcal{C}_\Phi} \cup T^{\mathcal{C}_\Phi} \text{ implies } (f(a, b) \in S^{\mathcal{C}_\Phi} \Leftrightarrow f(b, a) \in S^{\mathcal{C}_\Phi}). \quad (\dagger)$$

For a binary polymorphism f of $\mathcal{C}_\Phi$, denote by $C(f) = \{(a, b) \in A^2 \mid \exists \alpha \in \mathrm{Aut}(\mathcal{C}_\Phi) : f(a, b) = \alpha f(b, a)\}$. Let f be such that $C(f)$ is maximal. Suppose that f does not satisfy (†). This means that there exist $a, b \in A \cap X$ such that $f(a, b), f(b, a) \in S^{\mathcal{C}_\Phi} \cup T^{\mathcal{C}_\Phi}$ and such that $f(a, b) \in S^{\mathcal{C}_\Phi}$ and $f(b, a) \in T^{\mathcal{C}_\Phi}$. Let N be the smallest binary relation containing $(f(a, b), f(b, a)), (f(b, a), f(a, b))$ and being preserved by the polymorphisms of $\mathcal{C}_\Phi$. Note that $N \subseteq X^2$, since a and b are in X and X is preserved by all the polymorphisms of $\mathcal{C}_\Phi$. Since $\mathcal{C}_\Phi$ is ω-categorical, this relation has a pp-definition in $\mathcal{C}_\Phi$ (Theorem 2.12). Moreover, it is symmetric and $(f(a, b), f(b, a)) \in N \cap (S^{\mathcal{C}_\Phi} \times T^{\mathcal{C}_\Phi})$. By hypothesis, the colour graph defined by N contains a loop. This implies that there exist $g \in \mathrm{Pol}(\mathcal{C}_\Phi)$ and $\alpha \in \mathrm{Aut}(\mathcal{C}_\Phi)$ such that $g(f(a, b), f(b, a)) = \alpha g(f(b, a), f(a, b))$. Define $f'(x, y) := g(f(x, y), f(y, x))$ for all $x, y \in \mathcal{C}_\Phi$. It is clear from the above that $(a, b) \in C(f')$. Moreover, we have $C(f) \subseteq C(f')$. Indeed, let $(a', b') \in C(f)$. Then $f(a', b'), f(b', a')$ are in the same orbit, and since Φ is precoloured, this implies that $f'(a', b')$ and $f'(b', a')$ are in the same orbit. This contradicts the maximality of $C(f)$, so that it must be the case that f satisfies (†).

Using a standard compactness argument (see the proof of Proposition 2.7), we obtain a binary polymorphism f of $\mathcal{C}_\Phi$ that satisfies (†) for all $a, b \in X$.

Let g be any polymorphism obtained by diagonally canonising f, using Theorem 5.44. We claim that g still satisfies (†) on X. Indeed, let $a, b \in X$ and suppose that $g(a, b), g(b, a) \in S^{\mathcal{C}_\Phi} \cup T^{\mathcal{C}_\Phi}$. There exist $\alpha, \beta \in \mathrm{Aut}(\mathcal{C}_\Phi)$ such that $g(a, b) = \alpha f(\beta a, \beta b)$ and $g(b, a) = \alpha f(\beta b, \beta a)$. Since $S^{\mathcal{C}_\Phi}$ and $T^{\mathcal{C}_\Phi}$ are union of colours, they are preserved by automorphisms of $\mathcal{C}_\Phi$. We conclude that $f(\beta a, \beta b), f(\beta b, \beta a) \in S^{\mathcal{C}_\Phi} \cup T^{\mathcal{C}_\Phi}$. Since f satisfies

(†) on X, the equivalence $f(\beta a, \beta b) \in S^{\mathcal{C}_\Phi} \Leftrightarrow f(\beta b, \beta a) \in S^{\mathcal{C}_\Phi}$ holds. It follows that $g(a,b) \in S^{\mathcal{C}_\Phi} \Leftrightarrow g(b,a) \in S^{\mathcal{C}_\Phi}$, so that g also satisfies (†) on X.

Let $R \in S, B \in T$ be such that $R^{\mathcal{C}_\Phi} \subseteq X$ and $B^{\mathcal{C}_\Phi} \subseteq X$. Let e_1, e_2 be the endomorphisms of $\mathcal{C}_\Phi$ given by Lemma 5.43. Define $h(x,y) := g(e_1(x), e_2(y))$ for all $x, y \in \mathcal{C}_\Phi$. Note that h is 1-canonical on $R^{\mathcal{C}}_\Phi \cup B^{\mathcal{C}}_\Phi$: for $(a,b),(a',b') \in R^{\mathcal{C}_\Phi} \times B^{\mathcal{C}_\Phi}$, the pairs $(e_1(a), e_2(b))$ and $(e_1(a'), e_2(b'))$ are in the same orbit of $(\mathcal{C}_\Phi, <)$, according to Lemma 5.43. Since g is diagonally canonical, this implies that $h(a,b)$ and $h(a',b')$ are in the same orbit. Similarly, for $(a,b),(a',b') \in B^{\mathcal{C}_\Phi} \times R^{\mathcal{C}_\Phi}$, the pairs $(e_1(a), e_2(b))$ and $(e_1(a'), e_2(b'))$ are in the same orbit of $(\mathcal{C}_\Phi, <)$. Moreover, h satisfies (†) on $R^{\mathcal{C}_\Phi} \cup B^{\mathcal{C}_\Phi}$. Indeed, let $(a,b) \in R^{\mathcal{C}_\Phi} \times B^{\mathcal{C}_\Phi}$ be such that $h(a,b)$ and $h(b,a)$ are in $S^{\mathcal{C}_\Phi} \cup T^{\mathcal{C}_\Phi}$. Then $g(e_1(a), e_2(b))$ and $g(e_1(b), e_2(a))$ are in $S^{\mathcal{C}_\Phi} \cup T^{\mathcal{C}_\Phi}$. Since g is diagonally canonical and $(e_1(b), e_2(a))$ and $(e_2(b), e_1(a))$ are in the same orbit, we have that also $g(e_2(b), e_1(a))$ is in $S^{\mathcal{C}_\Phi} \cup T^{\mathcal{C}_\Phi}$. By (†), we have $g(e_1(a), e_2(b)) \in S^{\mathcal{C}_\Phi}$ if, and only if, $g(e_2(b), e_1(a)) \in S^{\mathcal{C}_\Phi}$. By definition, this implies that $h(a,b) \in S^{\mathcal{C}_\Phi} \Leftrightarrow h(b,a) \in S^{\mathcal{C}_\Phi}$ holds. So h satisfies (†) on $R^{\mathcal{C}_\Phi} \cup B^{\mathcal{C}_\Phi}$.

Let now $\hat{h}$ be obtained by canonising h with respect to $(\mathcal{C}_\Phi, <)$. Since h was already 1-canonical on $R^{\mathcal{C}_\Phi} \cup B^{\mathcal{C}_\Phi}$, the restrictions of $\xi_1^{\text{typ}}(h)$ and $\xi_1^{\text{typ}}(\hat{h})$ to $\{R, B\}$ are equal. This implies that $\hat{h}$ still satisfies (†) on $R^{\mathcal{C}_\Phi} \cup B^{\mathcal{C}_\Phi}$. By assumption, $S^{\mathcal{C}_\phi} \cup T^{\mathcal{C}_\Phi}$ is preserved by $\hat{h}$. This implies that for all $a \in R^{\mathcal{C}_\Phi}, b \in B^{\mathcal{C}_\Phi}$, we have that $\hat{h}(a,b) \in S^{\mathcal{C}_\Phi} \Leftrightarrow \hat{h}(b,a) \in S^{\mathcal{C}_\Phi}$. Finally, since the partition $\{S, T\}$ is preserved by $\hat{h}$ by assumption, for all $a, a' \in S^{\mathcal{C}_\Phi}$ and $b, b' \in T^{\mathcal{C}_\Phi}$ we must have that $\hat{h}(a,b) \in S^{\mathcal{C}_\Phi}$ iff $\hat{h}(a',b') \in S^{\mathcal{C}_\Phi}$, and similarly for $\hat{h}(b,a)$ and $\hat{h}(b',a')$. This finishes the construction of $\hat{h}$.

Note that the function induced by $\hat{h}$ on the subfactor $\{S, T\}$ is binary and symmetric. But since $\{S, T\}$ is a trivial subfactor of $\mathscr{C}_1^{\text{typ}}$, the clone induced by $\mathscr{C}_1^{\text{typ}}$ on $\{S, T\}$ only contains projections. We have reached the desired contradiction. □

Proposition 5.46. *Let Φ be a precoloured MMSNP sentence in normal form and let $\mathscr{C}$ be the clone of polymorphisms of $\mathcal{C}_\Phi$ that are canonical with respect to $(\mathcal{C}_\Phi, <)$. Let $\{S, T\}$ be a trivial subfactor of $\mathscr{C}_1^{\text{typ}}$. Then there exists a binary symmetric relation N that is pp-definable in $\mathcal{C}_\Phi$, defines a colour graph with an edge from S to T, and does not contain a path of even length between S and T.*

Proof. Consider the set $\mathcal{S}$ of all triples (S, T, N) such that:

- $\{S, T\}$ is a trivial subfactor of $\mathscr{C}_1^{\text{typ}}$;
- N is a binary symmetric relation pp-definable in $\mathcal{C}_\Phi$ whose colour graph is loopless;
- if the colour graph defined by N is bipartite, then N contains an edge between S and T.

By Proposition 5.45 applied with $X := \mathcal{C}_\Phi$, the set $\mathcal{S}$ is nonempty. Pick a triple $(S, T, N) \in \mathcal{S}$ such that the *support* of N (i.e., the set of $x \in \mathcal{C}_\Phi$ such that there exists $(x, y) \in N$ for some y) intersects the fewest number of orbits of $\mathcal{C}_\Phi$. We claim that N then satisfies the conclusion of the proposition.

We first claim that the colour graph defined by N is bipartite. For the sake of contradiction, suppose that it is not the case. If ℓ is the smallest length of an odd cycle in the colour graph defined by N, then the relation $N^{\circ(\ell-2)}$ is again symmetric, pp-definable in $\mathcal{C}_\Phi$, and such that the colour graph it defines is loopless and contains a triangle. We

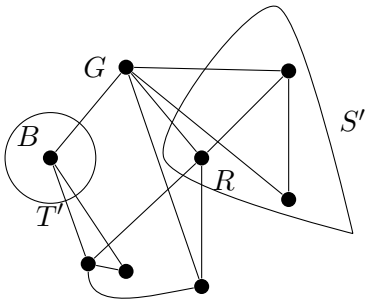

Figure 5.5: An illustration of the colour graph defined by N in the proof of Proposition 5.46. The formula $\exists y(G(y) \wedge N(x,y))$ defines in $\mathcal{C}_\Phi$ a subset that does not intersect $G^{\mathcal{C}_\Phi}$ and intersects $R^{\mathcal{C}_\Phi}$ and $B^{\mathcal{C}_\Phi}$, a contradiction to the minimality of $(S,T,N) \in \mathcal{S}$.

therefore assume without loss of generality that $\ell = 3$. The set defined by the unary formula

$$\phi(x) := \exists y, z, x'(N(x,y) \wedge N(y,z) \wedge N(z,x') \wedge x \sim x')$$

is preserved by $\mathscr{C}$, where the meaning of $x \sim x'$ is that x and x' are in the same orbit of $\mathrm{Aut}(\mathcal{C}_\Phi)$. The corresponding subset ρ of σ is therefore preserved by $\mathscr{C}_1^{\mathrm{typ}}$, and consists of the set of colours that belong to a triangle in the colour graph defined by N. In particular, ρ is not empty. When restricted to ρ, the colour graph is therefore not bipartite and does not contain a loop. It follows from [42, Theorem 1] that there is a trivial subfactor $\{S', T'\}$ of $\mathscr{C}_1^{\mathrm{typ}}$ such that $S' \cup T' \subseteq \rho$. Let now $R \in S'$, and let $N'(x,y)$ be defined by

$$N(x,y) \wedge \exists z, z' \left(R(z) \wedge R(z') \wedge N^{\circ 2}(z,x) \wedge N^{\circ 2}(z',y)\right),$$

which is a pp-definition in $\mathcal{C}_\Phi$. Again, $N' \subseteq N$ is symmetric and the colour graph that it defines is loopless and contains a triangle, so that $(S', T', N') \in \mathcal{S}$. By minimality of (S,T,N), the support of N' must be equal to the support of N, so that the colour graphs defined by N' and by N are the same. In particular, since every colour in T' is in the support of N, every such $B \in T'$ has a path of length 2 to R. Pick an arbitrary $B \in T'$ and let G be the midpoint of a path of length 2 between R and B. The situation is described in Figure 5.5. Then the formula $\theta(x) := \exists y(G(y) \wedge N(x,y))$ defines a proper subset Y of the support of N (proper because the colour graph is loopless, which implies that no element in G satisfies θ), and this subset intersects S' and T'. By Proposition 5.45 applied with $X := Y$, we obtain a new binary symmetric relation $M \subseteq Y^2$ whose colour graph is loopless and contains an edge between S' and T'. In particular, $(S', T', M) \in \mathcal{S}$, a contradiction to the minimality of (S,T,N).

Therefore, the colour graph defined by N is bipartite. Since $(S,T,N) \in \mathcal{S}$, it must be that N contains an edge between S and T. If N does not satisfy the conclusion of the proposition, there must be a path of even length $2k$ between S and T. Let G be the midpoint of this path. Let $Y \subseteq \mathcal{C}_\Phi$ be the subset consisting of the elements of $\mathcal{C}_\Phi$ that are reachable by a path of length k from an element in G. This set is pp-definable in $\mathcal{C}_\Phi$, and intersects S and T. Moreover, it is a proper subset of the support of N. Indeed, if k is odd then no element of G belongs to Y, otherwise the colour graph defined by N would contain a cycle of length k. If k is even, then no direct neighbour of an element in G

belongs to Y, for otherwise N would contain a cycle of length $k+1$. By Proposition 5.45 applied with $X := Y$, we obtain a new binary symmetric relation $M \subseteq Y^2$ whose colour graph is loopless and contains an edge between S' and T'. In particular, $(S', T', M) \in \mathcal{S}$, a contradiction to the minimality of (S, T, N). Thus, it must be that (S, T, N) satisfies the conclusion of the statement. □

Theorem 5.47. *Let Φ be a precoloured MMSNP sentence in strong normal form. Let $\mathscr{C}$ be the clone of polymorphisms of $\mathcal{C}_\Phi$ that are canonical with respect to $(\mathcal{C}_\Phi, <)$. If there is a clone homomorphism $\mathscr{C}_1^{\text{typ}} \to \mathscr{P}$, then there exists a uniformly continuous clonoid homomorphism from ϕ from $\operatorname{Pol}(\mathcal{C}_\Phi)$ to $\mathscr{P}$ and $\operatorname{CSP}(\mathcal{C}_\Phi^\tau)$ is NP-hard. Moreover, ϕ is constant on sets of the form $\overline{\operatorname{Aut}(\mathcal{C}_\Phi) f \operatorname{Aut}(\mathcal{C}_\Phi)}$ for $f \in \operatorname{Pol}(\mathcal{C}_\Phi)$.*

Proof. As we have mentioned before, if the finite idempotent clone $\mathscr{C}_1^{\text{typ}}$ has a homomorphism to $\mathscr{P}$, then $\mathscr{C}_1^{\text{typ}}$ has a trivial subfactor $\{S, T\}$ (Proposition 4.1).

Let $\xi \colon \mathscr{C}_1^{\text{typ}} \to \mathscr{P}$ be the clone homomorphism defined as follows. Let $R \in S$ and $B \in T$ be arbitrary. For a k-ary $f \in \mathscr{C}_1^{\text{typ}}$, let $i \in \{1, \dots, k\}$ be the unique index such that $f(B, ..., B, R, B, ..., B) \in S$, where the argument R is in the ith position. Such an i exists because of the assumption that $\{S, T\}$ is a trivial subfactor of $\mathscr{C}_1^{\text{typ}}$. Define $\xi(f)$ to be the ith projection. Note that the definition of ξ does not depend on the choice of R and B, by the fact that the equivalence relation on $S \cup T$ whose equivalence classes are S and T is assumed to be preserved by the operations in $\mathscr{C}_1^{\text{typ}}$. It is straightforward to check that the map ξ thus defined is a clone homomorphism.

Let $X \subseteq \mathcal{C}_\Phi$ and $N \subseteq X^2$ be the pp-definable relations given by Proposition 5.46. Fix $f \in \operatorname{Pol}(\mathcal{C}_\Phi)$ a k-ary operation and g, h two operations in $\overline{\operatorname{Aut}(\mathcal{C}_\Phi, <) f \operatorname{Aut}(\mathcal{C}_\Phi, <)}$. As explained in the beginning of this section, it suffices to prove that $\xi(g_1^{\text{typ}}) = \xi(h_1^{\text{typ}})$. For ease of notation, assume that $\xi(g_1^{\text{typ}})$ is the first projection, the general case being treated in the same way. Since ξ is the clone homomorphism induced by $\{S, T\}$, this means that for all $R \in S$ and $B \in T$, we have $g_1^{\text{typ}}(R, B, \dots, B) \in S$. In order to prove that $\xi(h_1^{\text{typ}})$ is also the first projection, it suffices to prove that there exists $R \in S$ and $B \in T$ such that $h_1^{\text{typ}}(R, B, \dots, B) \in S$. Let $R \in S$ and $B \in T$ be adjacent colours in the colour graph defined by N. Let $(a_1, \dots, a_k)$ be any tuple in $R^{\mathcal{C}_\Phi} \times B^{\mathcal{C}_\Phi} \times \cdots \times B^{\mathcal{C}_\Phi}$. Since f interpolates g and h modulo $\operatorname{Aut}(\mathcal{C}_\Phi, <)$, there are automorphisms $\alpha, \beta_1, \dots, \beta_k$ such that

$$g(a_1, \dots, a_k) = \alpha f(\beta_1 a_1, \dots, \beta_k a_k)$$

and automorphisms $\gamma, \delta_1, \dots, \delta_k$ such that

$$h(a_1, \dots, a_k) = \gamma f(\delta_1 a_1, \dots, \delta_k a_k).$$

Let $\mathcal{S}$ be the substructure of $\mathcal{C}_\Phi$ induced by $\{\beta_1 a_1, \dots, \beta_k a_k, \delta_1 a_1, \dots, \delta_k a_k\}$. Since $(\mathcal{C}_\Phi, \neq)$ is a model-complete core (Lemma 5.19), by Proposition 2.19 by Proposition 2.19 the orbit of the tuple $(\beta_1 a_1, \dots, \beta_k a_k, \delta_1 a_1, \dots, \delta_k a_k)$ has a pp-definition $\theta(x_1, \dots, x_k, y_1, \dots, y_k)$ in $(\mathcal{C}_\Phi, \neq)$. Let θ^* be θ where the atomic conjuncts involving $\neq$ have been removed. Let $\phi_N(x, y)$ be a pp-formula defining the relation $N \subseteq (\mathcal{C}_\Phi)^2$ in $\mathcal{C}_\Phi$. Fix an integer ℓ such that $2\ell > |\Phi|$. For every $i \in \{1, \dots, k\}$, let $z_1^i, \dots, z_{2\ell-1}^i$ be fresh variables. In the following, we also write z_0^i for x_i and $z_{2\ell}^i$ for y_i. Let $\psi(x_1, \dots, x_k, y_1, \dots, y_k)$ be the pp-formula whose conjuncts are:

- $\theta^*(x_1, \dots, x_k, y_1, \dots, y_k)$,

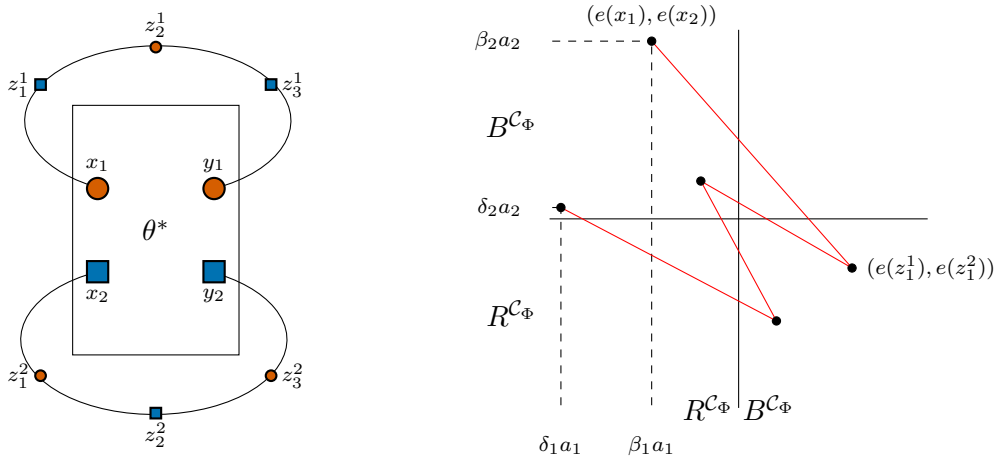

Figure 5.6: Proof of Theorem 5.47: A depiction of ψ (left) in the case that $k = 2$ and $2\ell = 4$, and a view of $(R^{\mathcal{C}_\Phi} \cup B^{\mathcal{C}_\Phi})^2$ (right). The red edges on the right represent the relation N; these edges connect the images of the drawn points under f.

- $\phi_N(z_j^i, z_{j+1}^i)$, for every $i \in \{1, \dots, k\}$ and $j \in \{0, \dots, 2\ell - 1\}$,
- $R(z_j^1)$ for even $j \in \{1, \dots, 2\ell - 1\}$ and $B(z_j^1)$ for odd $j \in \{1, \dots, 2\ell - 1\}$,
- for $i \in \{2, \dots, k\}$, the conjunct $B(z_j^i)$ for even $j \in \{1, \dots, 2\ell - 1\}$ and $R(z_j^i)$ for odd $j \in \{1, \dots, 2\ell - 1\}$.

We claim that ψ is satisfiable in $\mathcal{C}_\Phi$. We first prove that it is satisfiable in $\mathcal{B}_{\mathfrak{F}}^{\text{ind}}$, where $\mathfrak{F}$ is the coloured obstruction set of Φ. Let $\mathcal{S}'$ be the canonical database of ψ (see Figure 5.6). By Lemma 5.8, ψ is satisfiable if and only if all the biconnected components of $\mathcal{S}'$ are $\mathfrak{F}$-free. Suppose that there exists an obstruction $\mathcal{F} \in \mathfrak{F}$ and a homomorphism $e\colon \mathcal{F} \to \mathcal{S}'$. By the choice of ℓ we have that $|\mathcal{F}| < 2\ell$. Since Φ is in normal form, its obstructions are biconnected and we can suppose that the image of the homomorphism e is a biconnected component of $\mathcal{S}'$. It follows that either the image of e is included in $\mathcal{S}$, or it is included in the subset induced by the canonical database of some $N(z_j^i, z_{j+1}^i)$ for some $i \in \{1, \dots, k\}$ and $j \in \{0, \dots, 2\ell - 1\}$. But the assumption on N is that there is $(a, b) \in N$ such that $a \in R^{\mathcal{C}_\Phi}$ and $b \in B^{\mathcal{C}_\Phi}$. Therefore, the conjunct $\phi_N(z_j^i, z_{j+1}^i)$ is satisfiable by an assignment that maps z_j^i and z_{j+1}^i to the appropriate colours. We conclude that there exists an embedding e of $\mathcal{S}'$ into $\mathcal{B}_{\mathfrak{F}}^{\text{ind}}$.

Let $d\colon \mathcal{B}_{\mathfrak{F}}^{\text{ind}} \to \mathcal{B}_{\mathfrak{F}}^{\text{hom}}$ be an injective homomorphism (whose existence follows from Theorem 5.9). Note that the image of the restriction of d to the substructure $\mathcal{C}_\Phi$ of $\mathcal{B}_{\mathfrak{F}}^{\text{ind}}$ is in $\mathcal{C}_\Phi$ since d must preserve the colours. Since $d \circ e$ is injective, the tuple

$$(e(x_1), \dots, e(x_k), e(y_1), \dots, e(y_k))$$

satisfies θ. This means that $d \circ e\colon \mathcal{S}' \to \mathcal{C}_\Phi$ is a satisfying assignment that maps

$$(x_1, \dots, x_k, y_1, \dots, y_k)$$

to a tuple that is in the same orbit as $(\beta_1 a_1, \dots, \beta_k a_k, \delta_1 a_1, \dots, \delta_k a_k)$. By composing with an automorphism of $\mathcal{C}_\Phi$, we can suppose that $(x_1, \dots, x_k, y_1, \dots, y_k)$ is exactly this tuple.

It must therefore be the case that $f(\beta_1 a_1, \dots, \beta_k a_k)$ and $f(\delta_1 a_1, \dots, \delta_k a_k)$ are connected by an N-path of even length, that is, there are $b_1, \dots, b_{2\ell-1} \in \mathcal{C}_\Phi$ such that

$$\begin{aligned} (b_j, b_{j+1}) &\in N \quad \text{for all } j \in \{1, \dots, 2\ell\} \\ (f(\beta_1 a_1, \dots, \beta_k a_k), b_1) &\in N \quad \text{and} \\ (b_{2\ell-1}, f(\delta_1 a_1, \dots, \delta_k a_k)) &\in N \quad \text{(see Figure 5.6, right side).} \end{aligned}$$

By assumption, such a path of even length cannot connect S and T, so that $f(\delta_1 a_1, \dots, \delta_k a_k)$ is in a colour $Y \in S$. We obtain that $h(R, B, \dots, B) \in S$, as desired.

Therefore, the map $\phi\colon \mathrm{Pol}(\mathcal{C}_\Phi) \to \mathscr{P}$ defined by $\phi(f) := \xi(g_1^{\mathrm{typ}})$ for an arbitrary $g \in \overline{\mathrm{Aut}(\mathcal{C}_\Phi, <) f \,\mathrm{Aut}(\mathcal{C}_\Phi, <)}$ is well-defined. As in Theorem 4.3, ϕ is readily seen to be constant on sets of the form $\overline{\mathrm{Aut}(\mathcal{C}_\Phi) f \,\mathrm{Aut}(\mathcal{C}_\Phi)}$ for $f \in \mathrm{Pol}(\mathcal{C}_\Phi)$. In order to prove continuity of ϕ, let $(f_n)_{n\in\mathbb{N}}$ be a sequence of polymorphisms of $\mathcal{C}_\Phi$ converging to f. Let h_n be canonical and in $\overline{\mathrm{Aut}(\mathcal{C}_\Phi, <) f_n \,\mathrm{Aut}(\mathcal{C}_\Phi, <)}$ for all $n \in \mathbb{N}$, and let g be canonical and in $\overline{\mathrm{Aut}(\mathcal{C}_\Phi, <) f \,\mathrm{Aut}(\mathcal{C}_\Phi, <)}$. Since there are only finitely many behaviours of canonical functions, there is an infinite set $I \subseteq \mathbb{N}$ such that $\xi_1^{\mathrm{typ}}(h_i) = \xi_1^{\mathrm{typ}}(h_j)$ for all $i, j \in I$. Now, it remains to apply the same as in the third paragraph. Indeed, when $a_1, \dots, a_k$ are fixed, for arbitrarily large $n \in I$ we obtain automorphisms $\alpha, \beta_1, \dots, \beta_k, \gamma, \delta_1, \dots, \delta_k \in \mathrm{Aut}(\mathcal{C}_\Phi)$ such that

$$g(a_1, \dots, a_k) = \alpha f_n(\beta_1 a_1, \dots, \beta_k a_k)$$

and

$$h_n(a_1, \dots, a_k) = \gamma f_n(\delta_1 a_1, \dots, \delta_k a_k).$$

By the argument in the third paragraph, we then obtain $\xi(\xi_1^{\mathrm{typ}}(h_n)) = \xi(\xi_1^{\mathrm{typ}}(g)) = \phi(f)$. Thus, the sequence $(\phi(f_n))_{n\in\mathbb{N}} = (\xi(\xi_1^{\mathrm{typ}}(h_n)))_{n\in\mathbb{N}}$ converges to $\xi(\xi_1^{\mathrm{typ}}(g)) = \phi(f)$ and ϕ is continuous. Finally, the fact that ϕ is uniformly continuous follows from Proposition 6.4 in [6]. □

5.4.3 The dichotomy: conclusion

Summing up the results of the previous two sections, we obtain the following dichotomy for precoloured MMSNP sentences.

Theorem 5.48. *Let Φ be a precoloured MMSNP sentence. Let $\mathscr{C}$ be the clone of polymorphisms of $\mathcal{C}_\Phi^\tau$ that are canonical with respect to $(\mathcal{C}_\Phi, <)$. Then one of the following equivalent statements holds:*

(1) there is a clone homomorphism $\mathscr{C}_1^{\mathrm{typ}} \to \mathscr{P}$;

(2) there is a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{C}_\Phi^\tau) \to \mathscr{P}$ that is invariant under left-composition by $\overline{\mathrm{Aut}(\mathcal{C}_\Phi)}$;

and $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$ is NP-complete, or one of the following equivalent statements holds:

(a) $\mathscr{C}_1^{\mathrm{typ}}$ contains a Siggers operation;

(b) $\mathscr{C}$ contains a pseudo-Siggers operation modulo $\overline{\mathrm{Aut}(\mathcal{C}_\Phi, <)}$;

(c) $\mathrm{Pol}(\mathcal{C}_\Phi^\tau)$ *contains a pseudo-Siggers operation modulo* $\overline{\mathrm{Aut}(\mathcal{C}_\Phi, <)}$.

and $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$ *is in* P.

Proof. The implication from (a) to (b) follows from Proposition 5.39. The implication from (b) to (c) is trivial. Clearly, (c) implies the negation of (2). The implication $\neg(2) \Rightarrow \neg(1)$ is Theorem 5.47, and $\neg(1)$ implies (a) by Theorem 2.17. □

Note that item (a) is for given Φ clearly algorithmically decidable. Via the facts about precolourings from Section 5.3, Theorem 5.48 implies a more general result about MMSNP sentences in normal form, Theorem 5.50 below. In order to show that the two cases in Theorem 5.50 are disjoint, we need the following transfer for the existence of pseudo-Siggers polymorphisms of $\mathrm{Pol}(\mathcal{C}_\Phi^\tau)$.

Proposition 5.49. *The structure* $\mathcal{C}_\Phi^\tau$ *has a pseudo-Siggers polymorphism modulo* $\overline{\mathrm{Aut}(\mathcal{C}_\Phi)}$ *if, and only if, it has an injective polymorphism that is pseudo-Siggers modulo* $\overline{\mathrm{Aut}(\mathcal{C}_\Phi, <)}$.

Proof. Let $s\colon (\mathcal{C}_\Phi^\tau)^6 \to \mathcal{C}_\Phi^\tau$ be the given pseudo-Siggers. Let $\mathcal{B}$ be the $(\tau \cup \sigma)$-expansion of $(\mathcal{C}_\Phi^\tau)^6$ where $(a_1, \ldots, a_6)$ has the same colour as $s(a_1, \ldots, a_6)$ in $\mathcal{C}_\Phi$. We view $\mathcal{C}_\Phi$ as a substructure of $\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}$, and consequently s as a homomorphism $\mathcal{B} \to \mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}$. By Lemma 5.30, we obtain an injective homomorphism $t\colon \mathcal{B} \to \mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}$ such that for all injective tuples $\overline{a}, \overline{b}$ in $\mathcal{B}$, if $s(\overline{a})$ and $s(\overline{b})$ are in the same orbit in $\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}$ then so are $t(\overline{a})$ and $t(\overline{b})$ (call this property (†)).

We claim that for every finite substructure $\mathcal{A}$ of $\mathcal{C}_\Phi^\tau$, there exists an injective homomorphism $t_A\colon \mathcal{A}^6 \to \mathcal{C}_\Phi^\tau$ that is pseudo-Siggers modulo $\mathrm{Aut}(\mathcal{C}_\Phi, <)$. Let $\overline{a}$ be the tuple whose entries are of the form (x, y, x, z, y, z) for $x, y, z \in A$ (that is, $\overline{a}$ is a tuple of 6-tuples). Let $\overline{b}$ be the tuple whose entries are of the form (y, x, z, x, z, y) (using the same enumeration of the elements (x, y, z) of A^3 as in $\overline{a}$). Since s is pseudo-Siggers modulo $\mathrm{Aut}(\mathcal{C}_\Phi)$, the tuples $s(\overline{a})$ and $s(\overline{b})$ lie in the same orbit of $\mathrm{Aut}(\mathcal{C}_\Phi)$, so they lie in the same orbit of $\mathrm{Aut}(\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}})$ by Lemma 5.19. By (†), we obtain that $t(\overline{a})$ and $t(\overline{b})$ lie in the same orbit of $\mathrm{Aut}(\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}})$. Moreover, since t is injective, there exists $\alpha \in \mathrm{Aut}(\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}})$ such that the tuples $(\alpha t)(\overline{a})$ and $(\alpha t)(\overline{b})$ lie in the same orbit of $\mathrm{Aut}(\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}, <)$. Let $h\colon (\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}, \neq) \to (\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, \neq)$ be an injective homomorphism that is canonical from $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}, <)$ to $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, <)$. We claim that $t_A := h \circ \alpha \circ t$ is the desired injective homomorphism.

We first prove that the range of t_A is included in the domain of $\mathcal{C}_\Phi$, that is, that all the elements that appear in the range are coloured. Let $a_1, \ldots, a_6 \in A$. Since the range of s is included in the domain of $\mathcal{C}_\Phi$, there is an $M \in \sigma$ such that $s(a_1, \ldots, a_6) \in M^{\mathcal{C}_\Phi}$. By Lemma 5.30, the element $t(a_1, \ldots, a_6) \in M^{\mathcal{B}_{\mathcal{N}}^{\mathrm{ind}}}$, so that $h(\alpha(t(a_1, \ldots, a_6))) \in M^{\mathcal{C}_\Phi}$ and hence lies in $\mathcal{C}_\Phi$.

We now show that $t_A\colon \mathcal{A}^6 \to \mathcal{C}_\Phi^\tau$ is pseudo-Siggers modulo $\mathrm{Aut}(\mathcal{C}_\Phi, <)$. Note that since $(\alpha t)(\overline{a})$ and $(\alpha t)(\overline{b})$ lie in the same orbit in $\mathrm{Aut}(\mathcal{B}_{\mathfrak{F}}^{\mathrm{ind}}, <)$, the tuples $t_A(\overline{a})$ and $t_A(\overline{b})$ lie in the same orbit in $\mathrm{Aut}(\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, <)$ by the canonicity of h. Therefore, there exists $\beta \in \mathrm{Aut}(\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, <)$ such that $\beta t_A(\overline{a}) = t_A(\overline{b})$. Since the domain of $\mathcal{C}_\Phi$ is preserved by automorphisms of $(\mathcal{B}_{\mathfrak{F}}^{\mathrm{hom}}, <)$ the restriction of β to the domain of $\mathcal{C}_\Phi$ is an automorphism of $(\mathcal{C}_\Phi, <)$. In conclusion, t_A is pseudo-Siggers modulo $\mathrm{Aut}(\mathcal{C}_\Phi, <)$.

A standard compactness argument now shows that there exists $t'\colon (\mathcal{C}_\Phi)^6 \to \mathcal{C}_\Phi$ that is on every finite subset pseudo-Siggers modulo $\mathrm{Aut}(\mathcal{C}_\Phi, <)$. By Lemma 3.5, t' is pseudo-Siggers modulo $\overline{\mathrm{Aut}(\mathcal{C}_\Phi, <)}$. □

Theorem 5.50. *Let Φ be an MMSNP sentence in strong normal form. Let $\mathscr{C}$ be the clone of polymorphisms of $\mathcal{C}_\Phi^\tau$ that are canonical with respect to $(\mathcal{C}_\Phi, <)$. Then either*

- *there is a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{C}_\Phi^\tau) \to \mathscr{P}$ and $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$ is NP-complete, or*
- $\mathrm{Pol}(\mathcal{C}_\Phi^\tau)$ *contains a pseudo-Siggers operation modulo $\overline{\mathrm{Aut}(\mathcal{C}_\Phi)}$ and $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$ is in P.*

In particular, Conjecture 1 holds for all CSPs in MMSNP.

Proof. If there is a uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{C}_\Phi^\tau) \to \mathscr{P}$, then the NP-hardness of $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$ follows from Corollary 2.15. Otherwise, let Ψ be the standard precolouration of Φ with input signature $\rho \subseteq \tau$. By Theorem 5.31 there is no uniformly continuous clonoid homomorphism from $\mathrm{Pol}(\mathcal{C}_\Psi^\rho) \to \mathscr{P}$. Then Theorem 5.50 above states that $\mathrm{Pol}(\mathcal{C}_\Psi^\rho)$ contains a pseudo-Siggers operation modulo $\overline{\mathrm{Aut}(\mathcal{C}_\Psi)}$ that is canonical with respect to $\mathcal{C}_\Psi$, and $\mathrm{CSP}(\mathcal{C}_\Psi^\rho)$ is in P. By Theorem 5.31 the structure $\mathcal{C}_\Phi^\tau$ is isomorphic to a reduct of $\mathcal{C}_\Psi^\rho$, so it also has a pseudo-Siggers operation modulo $\overline{\mathrm{Aut}(\mathcal{C}_\Phi)}$ that is canonical with respect to $\mathcal{C}_\Phi$, and $\mathrm{CSP}(\mathcal{C}_\Phi^\tau)$ is also in P.

To show that the two cases are mutually exclusive, suppose that $\mathrm{Pol}(\mathcal{B})$ contains a pseudo-Siggers operation g. Then $\mathrm{Pol}(\mathcal{B}, \neq)$ has a pseudo-Siggers by Proposition 5.49. Since $(\mathcal{B}, \neq)$ is a model-complete core, Theorem 2.21 implies that there is no uniformly continuous clonoid homomorphism from $\mathrm{Pol}(\mathcal{B}, \neq) \xrightarrow{u.c.c.h.} \mathscr{P}$. By Proposition 5.29, there is no uniformly continuous clonoid homomorphism $\mathrm{Pol}(\mathcal{B}) \xrightarrow{u.c.c.h.} \mathscr{P}$.

Finally, we show that the above implies Conjecture 1 for CSPs in MMSNP. Suppose that $\mathcal{B}$ is an ω-categorical structure such that Φ describes $\mathrm{CSP}(\mathcal{B})$. Since $\mathcal{B}$ and $\mathcal{C}_\Phi^\tau$ are ω-categorical and have the same CSP, they are homomorphically equivalent. Theorem 2.14 then implies that there are uniformly continuous clonoid homomorphisms $\mathrm{Pol}(\mathcal{B}) \xrightarrow{u.c.c.h.} \mathrm{Pol}(\mathcal{C}_\Phi^\tau)$ and $\mathrm{Pol}(\mathcal{C}_\Phi^\tau) \xrightarrow{u.c.c.h.} \mathrm{Pol}(\mathcal{B})$. □

The proof of Theorem 5.50 shows that in order to decide for a given MMSNP sentence Φ in strong normal form which of the cases holds, it suffices to test whether $(\mathcal{C}_\Psi^\tau, <)$ has a polymorphism f that is canonical with respect to $(\mathcal{C}_\Phi, <)$ such that $\xi_1^{\mathrm{typ}}(f)$ is a Siggers operation (see item (a) in Theorem 5.48).

We can finally prove Theorem 5.6 from Section 5.1.3.

Proof. By Proposition 5.1, the sentence Φ is logically equivalent to a finite disjunction $\Phi_1 \vee \cdots \vee \Phi_k$ of connected MMSNP sentences. By Theorem 5.24, we can assume that each of the Φ_i is in strong normal form. The sentence Φ_i describes $\mathrm{CSP}(\mathcal{C}_{\Phi_i}^\tau)$. Theorem 5.50 above states that either $\mathrm{Pol}(\mathcal{C}_{\Phi_i}^\tau)$ has a uniformly continuous clonoid homomorphism to $\mathscr{P}$, and Φ_i is NP-complete, or $\mathrm{Pol}(\mathcal{C}_{\Phi_i}^\tau)$ contains a pseudo-Siggers polymorphism. Then Proposition 5.2 states that Φ is in Pif the second case applies for all $i \leq k$, and is NP-hard otherwise. □

Again, it is clear from the proof that given an MMSNP sentence Φ, the two cases in Theorem 6.5 can be distinguished algorithmically. The reason is that the connected MMSNP sentences $\Phi_1, \ldots, \Phi_k$ can be computed from Φ (Proposition 5.1), and also each of the Φ_i can be effectively rewritten into strong normal form (Theorem 5.24), and so the claim follows from our observations above.

We close with a consequence of Theorem 5.48 concerning the existence of pseudo-cyclic polymorphisms of $\mathcal{C}_\Phi^\tau$ for precoloured MMSNP sentences Φ. Recall that for finite structures $\mathcal{C}$, the existence of a Siggers polymorphisms is equivalent to the existence of a cyclic polymorphism. However, there are ω-categorical structures that have a pseudo-Siggers polymorphism but no pseudo-cyclic polymorphism, for example the structure $(\mathbb{Q}; <, \{(x, y, u, v) \mid x = y \Rightarrow u = v\})$. But the CSP for this structure cannot be expressed by MMSNP (a proof can be found in [9]). So it is natural to ask whether tractability of MMSNP sentences can also be characterised by pseudo-cyclic polymorphisms. The proof of Proposition 5.39 cannot be modified straightforwardly to produce a pseudo-cyclic polymorphism instead of a pseudo-Siggers polymorphism. However, the existence of a pseudo-cyclic polymorphism of $\mathcal{C}_\Phi^\tau$ can be deduced from Theorem 5.48 and the mentioned result about the existence of cyclic polymorphisms in the finite.

Theorem 5.51. *Let Φ be a precoloured MMSNP sentence. Then* $\mathrm{Pol}(\mathcal{C}_\Phi^\tau)$ *has a pseudo-Siggers polymorphism if and only if it has a pseudo-cyclic polymorphism.*

Proof. By Proposition 5.40 there exists an $m \in \mathbb{N}$ such that $\mathcal{C}_\Phi$ has a homogeneous expansion $\mathcal{C}_\Phi^*$ by primitive positive definable relations of maximal arity m. For the forward implication, the existence of a pseudo-Siggers polymorphism of $\mathcal{C}_\Phi^\tau$ implies by Theorem 5.48 that $\mathcal{C}_\Phi^\tau$ has a pseudo-Siggers operation modulo $\overline{\mathrm{Aut}(\mathcal{C}_\Phi, <)}$ which is canonical with respect to $(\mathcal{C}_\Phi, <)$, and hence $\mathscr{C}_m^{\mathrm{typ}}$ has a Siggers polymorphism. By Theorem 2.17 and Proposition 3.6, it follows that $\mathcal{C}_\Phi^\tau$ has a pseudo-cyclic polymorphism.

Now suppose that $\mathcal{C}_\Phi^\tau$ has a pseudo-cyclic polymorphism. Then $\mathscr{C}_1^{\mathrm{typ}}$ has a cyclic polymorphism, and hence $\mathcal{C}_\Phi^\tau$ has a pseudo-Siggers operation modulo $\overline{\mathrm{Aut}(\mathcal{C}_\Phi, <)}$ by Theorem 5.48. $\square$

Part II

Numeric CSPs

Chapter 6

The Complexity of Discrete Temporal CSPs

A famous CSP over the integers is *feasibility of systems of linear inequalities.* It is of great importance in practice and theory of computing, and NP-complete. In order to obtain a systematic understanding of polynomial-time solvable restrictions and variations of this computational problem, Jonsson and Lööw [64] proposed to study the class of CSPs where the constraint language $\mathcal{B}$ is definable in *Presburger arithmetic*; that is, it consists of relations that have a first-order definition over $(\mathbb{Z}; \leq, +)$. Equivalently, each relation $R(x_1, \ldots, x_n)$ in $\mathcal{B}$ can be defined by a disjunction of conjunctions of the atomic formulas of the form $p \geq 0$ where p is a linear polynomial with integer coefficients and variables from $\{x_1, \ldots, x_n\}$. Several constraint languages in this class are known where the CSP can be solved in polynomial time; an example of such a CSP is the problem of deciding whether a system of linear diophantine equations has a solution (a polynomial-time algorithm is given in [49]). However, a complete complexity classification for the CSPs of Jonsson-Lööw languages appears to be a very ambitious goal.

One of the most basic classes of constraint languages that falls into the framework of Jonsson and Lööw is the class of *distance constraint satisfaction problems* [15]. A distance constraint satisfaction problem is a CSP for a constraint language over the integers whose relations have a first-order definition over $(\mathbb{Z}; \mathrm{succ})$ where succ is the successor relation. It has been shown previously that distance CSPs for *locally finite* constraint languages, that is, constraint languages whose relations have *bounded Gaifman degree*, are NP-complete, in P, or can be formulated with a constraint language over a finite domain [15]. Another class of problems which can be expressed as Jonsson-Lööw constraint satisfaction problems is the class of *temporal CSPs* [23]. This is the class of problems whose constraint languages are over the rational numbers with relations definable over $(\mathbb{Q}; <)$. While the order of the rationals is not isomorphic to the order of the integers because of its density, this density is not witnessed by finite structures (i.e., $\mathrm{Age}(\mathbb{Q}; <) = \mathrm{Age}(\mathbb{Z}; <)$). It follows that for every structure $\mathcal{B}$ whose relations are first-order definable in $(\mathbb{Q}; <)$, there exists a structure $\mathcal{C}$ that is definable in $(\mathbb{Z}; <)$ and such that $\mathcal{B}$ and $\mathcal{C}$ have the same CSP. The converse is not true, since the structure $(\mathbb{Z}; \mathrm{succ})$ is a first-order reduct of $(\mathbb{Z}; <)$ that does not have the same CSP as any first-order reduct of $(\mathbb{Q}; <)$.

Our main result shows that the class of discrete temporal CSPs exhibits a P/NP-complete dichotomy. This result properly extends the results mentioned above for locally finite distance CSPs, since succ is first-order definable over $(\mathbb{Z}; <)$. By the comments of the previous paragraph it also extends the classification of temporal CSPs. A cornerstone

of our proof is the characterization of those problems that are discrete temporal CSPs but that are not temporal CSPs; the corresponding constraint languages have an interesting notion of *rank* which we use in the following to obtain a strong pre-classification of those languages up to homomorphic equivalence. The notion of rank is central to reduce the classification to the natural special case where the binary successor relation is part of the language.

Our proof relies on a modification of the universal-algebraic approach discovered by Bodirsky, Hils, and Martin [17]. In this work, the authors pointed out the relevance of the notion of saturation for the universal-algebraic approach, and one of the ideas developed in this chapter is that in order to use polymorphisms when the constraint language is not ω-categorical, we have to pass to the countable saturated model of the first-order theory of $(\mathbb{Z};<)$. Our classification has a particularly simple form when the constraint language $\mathcal{B}$ not only contains the binary successor relation, but also the relation $<$: if $\mathcal{B}$ has the polymorphism $(x,y) \mapsto \max(x,y)$ or $(x,y) \mapsto \min(x,y)$, then $\mathrm{CSP}(\mathcal{B})$ is in P, and $\mathrm{CSP}(\mathcal{B})$ is NP-hard otherwise. The results in this chapter have been published in [27, 28].

6.1 Discrete Temporal Constraint Satisfaction Problems

Definition 6.1 (Discrete Temporal CSP)**.** A *discrete temporal CSP* is a constraint satisfaction problem where the constraint language is a first-order reduct of $(\mathbb{Z};<)$ with finite signature.

Example 11. We present some concrete examples first-order reducts of $(\mathbb{Z};<)$; some of the relations from these examples will re-appear in later sections to illustrate important phenomena for reducts of $(\mathbb{Z};<)$.

1. $(\mathbb{Z};\mathrm{succ}^p)$, where $\mathrm{succ}^p = \{(x,y) \in \mathbb{Z}^2 \mid y = x+p\}$ for $p \in \mathbb{Z}$. Note that this structure is not connected, and that it has the same CSP as $(\mathbb{Z};\mathrm{succ})$. This example and example (3) will be considered again in Example 14.

2. $(\mathbb{Z};\mathrm{Diff}_S)$, where $\mathrm{Diff}_S := \{(x,y) \in \mathbb{Z}^2 \mid y - x \in S\}$ for a finite set $S \subset \mathbb{Z}$.

3. $(\mathbb{Z};\mathrm{succ}^2, \mathrm{Diff}_{\{-2,-1,0,1,2\}})$.

4. $(\mathbb{Z};F)$ where F is the 4-ary relation $\{(x,y,u,v) : y = x+1 \Leftrightarrow v = u+1\}$. This example and the following example have unbounded Gaifman degree (see Section 6.4.1), so they do not fall into the scope of [15].

5. $(\mathbb{Z};\neq, \mathrm{Dist}_i)$ where $i \in \mathbb{N}$ and $\mathrm{Dist}_i := \{(x,y) : |x-y| = i\}$.

6. $(\mathbb{Z};\{(x,y,z) \in \mathbb{Z}^3 \mid z+1 \leq \max(x,y)\})$. This structure is not a first-order reduct of $(\mathbb{Z};\mathrm{succ})$. Neither does it have the same CSP as a first-order reduct of $(\mathbb{Q};<)$, so we have a discrete temporal CSP that is not a temporal CSP and does not fall into the scope of [23]. The CSP for this structure is closely related to the so-called Max-Atom problem; the connection is explained in Section 6.7.

The structure $(\mathbb{Z};<)$ admits quantifier elimination in the language consisting of the binary relations $R_c = \{(x,y) \in \mathbb{Z}^2 \mid y \leq x+c\}$ for $c \in \mathbb{Z}$. This means that every first-order

formula $\phi(x_1, \ldots, x_n)$ in the expanded language $\{R_c \mid c \in \mathbb{Z}\}$ is equivalent to a quantifier-free formula in the same language. To see this, note that it suffices to prove that one can eliminate the quantifiers in existential formulas rather than in general first-order formulas; in fact, by de Morgan and the equivalence between $\neg y < x+c$ and $x < y+(1-c)$ it suffices to prove that one can eliminate the quantifiers in pp-formulas. Seeing a pp-formula as a system of inequalities, one then performs Gaussian elimination to remove the variables that are existentially quantified. The result of this is a system of inequalities that can be translated back into a quantifier-free formula. Similarly, $(\mathbb{Z}; \mathrm{succ})$ admits quantifier elimination in the language consisting of the binary relations given by $y = x + c$ for $c \in \mathbb{Z}$. Whenever we write that ϕ is a quantifier-free formula, we mean that ϕ is written in one of those two languages; which one will always be clear from the context. The empty relation, $\mathbb{Z}^2$, and the binary relations defined by $y = x + c$ for $c \in \mathbb{Z}$ are called *basic* relations. The following is easy to see.

Proposition 6.1. *All discrete temporal CSPs are in* NP.

Proof. Let q be the size of the biggest integer that appears in the quantifier-free formulas that define the relations in $\mathcal{B}$ over $(\mathbb{Z}; <)$; that is, for any atomic formula $x \leq y + k$ in those formulas, $k \in \mathbb{Z}$, we have $|k| \leq q$. For an instance Φ of $\mathrm{CSP}(\mathcal{B})$ with n variables, it is clear that $\mathcal{B} \models \Phi$ if and only if Φ is true on $\mathcal{B}[\{1, \ldots, (q+1)n\}]$. We may guess a satisfying assignment of values from $\{1, \ldots, (q+1)n\}$ to the variables of Φ, and verify in polynomial time that all the constraints are satisfied. □

The main result of this chapter (Theorem 6.5) immediately implies the following.

Theorem 6.2. *Every discrete temporal CSP is in* P *or* NP*-complete.*

6.2 Model-Theoretic Considerations

The structures that we consider in this chapter will in general not be ω-categorical; however, following the philosophy in [17], one can refine the universal-algebraic approach to apply it also in our situation. We will describe these refinements in the rest of this section.

The *(first-order) theory* of a structure $\mathcal{B}$, denoted by $\mathrm{Th}(\mathcal{B})$, is the set of all first-order sentences that are true in $\mathcal{B}$. We define some notation to conveniently work with models of $\mathrm{Th}(\mathcal{B})$ and their first-order reducts.

Definition 6.2 ($\kappa.\mathbb{Z}$)**.** Let κ be a linearly ordered set. We write $\kappa.\mathbb{Z}$ for κ copies of $\mathbb{Z}$ indexed by the elements of κ; formally, $\kappa.\mathbb{Z}$ is the set $\{(p, z) : p \in \kappa, z \in \mathbb{Z}\}$. Then $(\kappa.\mathbb{Z}; <)$ is the structure where $<$ denotes the lexicographic order on $\kappa.\mathbb{Z}$, that is, we define $(p, z) < (p', z')$ if $p < p'$ holds or if $p = p'$ and $z < z'$. If $p \in \kappa$, we write $p.\mathbb{Z}$ to denote the copy of $\mathbb{Z}$ indexed by p, instead of $\{p\} \times \mathbb{Z}$.

It is well known and easy to see that the models of $\mathrm{Th}(\mathbb{Z}; <)$ are precisely the structures isomorphic to $(\kappa.\mathbb{Z}; <)$, for some linear order κ. When $k \in \mathbb{Z}$ and $u = (p, z) \in \kappa.\mathbb{Z}$, we write $u + k$ for $(p, z + k)$.

Definition 6.3 ($\kappa.\mathcal{B}$)**.** Let $\mathcal{B}$ be a first-order reduct of $(\mathbb{Z}; <)$ with signature τ. Then $\kappa.\mathcal{B}$ denotes the 'corresponding' first-order reduct of $(\kappa.\mathbb{Z}; <)$ with signature τ. Formally, when $R \in \tau$ and ϕ_R is a formula that defines $R^{\mathcal{B}}$, then $R^{\kappa.\mathcal{B}}$ is the relation defined by ϕ_R over $(\kappa.\mathbb{Z}; <)$.

In the following, we identify $\mathbb{Z}$ with the copy of $\mathbb{Z}$ induced by $0.\mathbb{Z}$ in $\mathbb{Q}.\mathbb{Z}$. That is, we view $(\mathbb{Z};<)$ as a substructure of $(\mathbb{Q}.\mathbb{Z};<)$, and consequently $\mathcal{B}$ as a substructure of $\mathbb{Q}.\mathcal{B}$ for each first-order reduct $\mathcal{B}$ of $(\mathbb{Z};<)$. The structures $\mathcal{B}$ and $\mathbb{Q}.\mathcal{B}$ have the same first-order theory; in particular, they satisfy the same pp-sentences. It follows that $\mathcal{B}$ and $\mathbb{Q}.\mathcal{B}$ have the same CSP. Let $\phi(x_1,\dots,x_k)$ be a first-order formula in the language of $\mathcal{B}$. This formula defines a relation $R \subseteq \mathbb{Z}^k$ in $\mathcal{B}$ and a relation $R' \subseteq (\mathbb{Q}.\mathbb{Z})^k$. One sees (for example using quantifier elimination) that $R = R' \cap \mathbb{Z}^k$, i.e., the relations definable in $\mathcal{B}$ are precisely the intersections of $\mathbb{Z}$ with relations defined in $\mathbb{Q}.\mathcal{B}$. The link between endomorphisms of $\mathcal{B}$ and of $\mathbb{Q}.\mathcal{B}$ is more complicated, and is covered in Section 5.

A *partial type* of a structure $\mathcal{C}$ is a set p of formulas with free variables $x_1,\dots,x_n$ such that $p \cup \mathrm{Th}(\mathcal{C})$ is satisfiable (that is, $\{\phi(c_1,\dots,c_n) : \phi \in p\} \cup \mathrm{Th}(\mathcal{C})$, for new constant symbols $c_1,\dots,c_n$, has a model). A countable τ-structure $\mathcal{B}$ is *saturated* if for all choices of finitely many elements $a_1,\dots,a_n$ in $\mathcal{B}$, and every unary partial type p of $(\mathcal{B},a_1,\dots,a_n)$, there exists an element b of $\mathcal{B}$ such that $(\mathcal{B},a_1,\dots,a_n) \models \phi(b)$ for all $\phi \in p$. When $\mathcal{B}$ and $\mathcal{C}$ are two countable saturated structures with the same first-order theory, then $\mathcal{B}$ and $\mathcal{C}$ are isomorphic [56, Theorem 8.1.8]. Note that $(\mathbb{Q}.\mathbb{Z};<)$ is saturated. More generally, $\mathbb{Q}.\mathcal{B}$ is saturated for every first-order reduct $\mathcal{B}$ of $(\mathbb{Z};<)$.

We define the function $-\colon (\kappa.\mathbb{Z})^2 \to (\mathbb{Z} \cup \{\infty\})$ for $a,b \in \kappa.\mathbb{Z}$ by

$$\begin{aligned} a - b &:= k \in \mathbb{Z} && \text{if } a = b + k \\ a - b &:= \infty && \text{otherwise.} \end{aligned}$$

Lemma 6.3 (See Lemma 2.1 in [17])**.** *Let $\mathcal{B}$ be a countable saturated structure, let $\mathcal{C}$ be countable, let $d_1,\dots,d_k$ be elements of $\mathcal{C}$, and let $c_1,\dots,c_k$ be elements of $\mathcal{B}$. Suppose that for all pp-formulas ϕ such that $\mathcal{C} \models \phi(d_1,\dots,d_k)$ we have $\mathcal{B} \models \phi(c_1,\dots,c_k)$. Then there exists a homomorphism from $\mathcal{C}$ to $\mathcal{B}$ that maps d_i to c_i for all $i \leq k$.*

To classify the computational complexity of the CSP for all first-order reducts of a structure $\mathcal{B}$, it often turns out to be important to study the possible endomorphisms of those reducts first, before studying the polymorphisms. This has for instance been the case for the first-order reducts of $(\mathbb{Q};<)$ in [23] and the first-order reducts of the countably infinite random graph in [34].

We are now in the position to state a general fact, Theorem 6.4, whose proof might explain the importance of saturated models for the universal-algebraic approach. Let $\mathcal{B}$ be a structure with domain D. A relation $R \subseteq D^k$ is said to be *n-generated* under $\mathrm{End}(\mathcal{B})$ if there exist tuples $t_1,\dots,t_n \in R$ such that for every $t \in R$, there exist $e \in \mathrm{End}(\mathcal{B})$ and $i \in \{1,\dots,n\}$ such that $e(t_i) = t$. A *universal negative formula* is a first-order formula without existential quantifiers where the negation symbol only appears before an atom, and where all the atoms are negated.

Theorem 6.4. *Let $\mathcal{B}$ be a countable saturated structure, let $\mathcal{C}$ be a first-order reduct of $\mathcal{B}$, and R a relation with a first-order definition in $\mathcal{B}$. Then*

- *R has a first-order definition in $\mathcal{C}$ if and only if R is preserved by the automorphisms of $\mathcal{C}$;*
- *R has an existential positive definition in $\mathcal{C}$ if and only if R is preserved by all the endomorphisms of $\mathcal{C}$;*

- *if R is n-generated under* $\mathrm{End}(\mathcal{C})$*, then R has a pp-definition in $\mathcal{C}$ if and only if R is preserved by all polymorphisms of $\mathcal{C}$ of arity n.*

Proof. Suppose that R is k-ary, and let ϕ be the first-order definition of R in $\mathcal{B}$. We know from Lemma 2.11 that first-order formulas are preserved by automorphisms of $\mathcal{C}$, that existential positive formulas are preserved by endomorphisms of $\mathcal{C}$, and that pp-formulas are preserved by polymorphisms of $\mathcal{C}$. Therefore, only one implication is to be proved for every item.

Suppose first that R is preserved by all automorphisms of $\mathcal{C}$. Let ϕ be a first-order definition of R in $\mathcal{B}$. Let Ψ be the set of all first-order formulas in the language of $\mathcal{C}$ that are consequences of R. Formally,

$$\Psi = \{\psi(x_1, \ldots, x_k) \mid \forall (a_1, \ldots, a_k) \in R, \mathcal{C} \models \psi(a_1, \ldots, a_k)\}.$$

We prove that if a tuple $\overline{a}$ satisfies every formula in Ψ then $\overline{a}$ is in R. Let $\overline{a}$ be such a tuple. Let p be the type of $\overline{a}$ in $\mathcal{C}$. By replacing in p every relation symbol of the signature of $\mathcal{C}$ by a first-order definition of the corresponding relation in $\mathcal{B}$, we obtain a set q of formulas in the language of $\mathcal{B}$. If we can find some tuple $\overline{b}$ that satisfies $\{\phi\} \cup q$ in $\mathcal{B}$, then we are done. Indeed, we have that $\overline{b}$ is in R, and $\overline{b}$ has the same type as $\overline{a}$ in $\mathcal{C}$. The fact that $\overline{a}$ and $\overline{b}$ have the same type is equivalent to the fact that the structures $(\mathcal{C}, \overline{a})$ and $(\mathcal{C}, \overline{b})$ have the same first-order theory. We stated above that two countable saturated structures with the same first-order theory are isomorphic. Therefore, there exists an isomorphism $\alpha\colon (\mathcal{C}, \overline{b}) \to (\mathcal{C}, \overline{a})$. This isomorphism is an automorphism of $\mathcal{C}$ that maps $\overline{b}$ to $\overline{a}$, so that $\overline{a}$ is in R. So let us assume that $\{\phi\} \cup q$ is not satisfiable in $\mathcal{B}$. Since $\mathcal{B}$ is saturated, the set $\{\phi\} \cup q$ cannot possibly be a type. It follows that $\mathrm{Th}(\mathcal{B}) \cup q \cup \{\phi\}$ is not satisfiable. By the compactness theorem of first-order logic, there exists a finite subset q' of q such that $\mathrm{Th}(\mathcal{B}) \cup q' \cup \{\phi\}$ is not satisfiable. Note that q is closed under conjunctions of formulas, so that the conjunction of all the formulas of q' is a formula ψ in q. Therefore, $\mathrm{Th}(\mathcal{B}) \cup \{\phi, \psi\}$ is not satisfiable, i.e., we have $\mathrm{Th}(\mathcal{B}) \models \forall x_1, \ldots, x_k(\phi(x_1, \ldots, x_k) \Rightarrow \neg\psi(x_1, \ldots, x_k))$. By construction, the formula ψ corresponds to a formula θ in the language of $\mathcal{C}$. We obtain that $\neg\theta$ is in Ψ, so $\neg\theta$ is in p. But $\theta \in p$, a contradiction.

Suppose now that R is preserved by all endomorphisms of $\mathcal{C}$. In particular R is preserved by all the automorphisms of $\mathcal{C}$, so that there exists a first-order definition ϕ of R in $\mathcal{C}$. Let Ψ be the set of all universal negative consequences of R in $\mathcal{C}$. Formally,

$$\Psi = \{\psi(x_1, \ldots, x_k) \text{ universal negative formula } \mid \forall (a_1, \ldots, a_k) \in R, \mathcal{C} \models \psi(a_1, \ldots, a_k)\}.$$

As above, we aim to prove that if $\overline{a}$ satisfies all the formulas in Ψ, then $\overline{a}$ is in R. Let $\overline{a}$ be such a tuple, and let now p be the ep-type of $\overline{a}$, that is, the set of all the existential positive formulas ψ such that $\mathcal{C} \models \psi(\overline{a})$. If $p \cup \{\phi\}$ is satisfiable in $\mathcal{C}$, then we are done: there exists a tuple $\overline{b} \in R$ that has the same ep-type as $\overline{a}$. Lemma 6.3 implies that there exists an endomorphism of $\mathcal{C}$ that maps $\overline{b}$ to $\overline{a}$, so that $\overline{a}$ is in R. If $p \cup \{\phi\}$ is not satisfiable in $\mathcal{C}$, there exists a single formula $\psi \in p$ such that $\mathcal{B} \models \forall x_1, \ldots, x_k(\phi(x_1, \ldots, x_k) \Rightarrow \neg\psi(x_1, \ldots, x_k))$. To ψ corresponds an existential positive formula θ in the language of $\mathcal{C}$. We obtain that $\neg\theta$ is equivalent to a formula in Ψ, so that $\overline{a}$ must satisfy $\neg\theta$, contradicting the fact that $\overline{a}$ already satisfies θ.

Finally, suppose that R is n-generated under $\mathrm{End}(\mathcal{C})$, and that R is preserved by all polymorphisms of $\mathcal{C}$ of arity n. Let $(b_1^1, \ldots, b_k^1), \ldots, (b_1^n, \ldots, b_k^n)$ be n tuples of length k

generating the relation R under $\mathrm{End}(\mathcal{C})$. Let Ψ be the set of all pp-formulas with free variables $x_1, \ldots, x_k$ that hold on all these tuples, i.e.

$$\Psi = \{\psi(x_1, \ldots, x_k) \text{ pp-formula} \mid \forall i \in \{1, \ldots, n\}, \mathcal{C} \models \psi(\bar{b}^i)\}.$$

If $\overline{a}$ is in R, there exists by assumption an endomorphism e of $\mathcal{C}$ and an $i \in \{1, \ldots, n\}$ such that $e(\bar{b}^i) = \overline{a}$. Since pp-formulas are preserved by endomorphisms, the tuple $\overline{a}$ satisfies every pp-formula that $\bar{b}^i$ satisfies, so that in particular $\overline{a}$ satisfies Ψ. We now prove the converse. If $\overline{a}$ satisfies Ψ, we have that every pp-formula that holds on $(b_1^1, \ldots, b_k^1), \ldots, (b_1^n, \ldots, b_k^n)$ in $\mathcal{C}^n$ also holds on $\overline{a}$. By Lemma 6.3 and saturation of $\mathcal{C}$, there exists a homomorphism from $\mathcal{C}^n$ to $\mathcal{C}$ that maps $(b_i^1, \ldots, b_i^n)$ to a_i for all $i \in \{1, \ldots, k\}$. This map is a polymorphism of $\mathcal{C}$, and since R is preserved by polymorphisms of arity n, $(a_1, \ldots, a_k) \in R$. Therefore, $\overline{a}$ satisfies Ψ if and only if $\overline{a} \in R$. Similarly as before, a compactness argument for first-order logic over $\mathcal{B}$ shows that Ψ is equivalent to a single pp-formula that is equivalent to ϕ. □

6.3 Detailed Statement of the Results

In this section, we describe the border between the NP-complete and the polynomial-time tractable discrete temporal CSPs.

Definition 6.4. Let d be a positive integer. The *d-modular max*, $\max_d \colon \mathbb{Z}^2 \to \mathbb{Z}$, is defined by $\max_d(x, y) := \max(x, y)$ if $x = y \mod d$ and $\max_d(x, y) := x$ otherwise. The *d-modular min* is defined analogously, with $\min_d(x, y) = \min(x, y)$ if $x = y \mod d$ and $\min_d(x, y) = x$ otherwise.

Note that $\max_d$ and $\min_d$ are not commutative when $d > 1$. Also note that $\max_1 = \max$ and $\min_1 = \min$ are the usual maximum and minimum operations. Examples of relations which are preserved by max and which are definable over $(\mathbb{Z}; <)$ are the relations appearing in the last item of Example 11. An example of a relation which is preserved by $\max_d$ is the ternary relation containing the triples of the form

$$(a + d, a, a), (a + d, a + d, a), (a, a + d, a)$$

for all $a \in \mathbb{Z}$. Note that for a fixed d, this relation is preserved by $\max_d$ but not by $\max_{d'}$ for any other d'.

Theorem 6.5. *Let $\mathcal{B}$ be a first-order reduct of $(\mathbb{Z}; <)$ with finite signature. Then there exists a structure $\mathcal{C}$ such that $\mathrm{CSP}(\mathcal{C})$ equals $\mathrm{CSP}(\mathcal{B})$ and at least one of the following cases applies.*

1. *$\mathcal{C}$ has a finite domain, and the CSP for $\mathcal{B}$ is in P or NP-complete (Theorem 2.5).*
2. *$\mathcal{C}$ is a reduct of $(\mathbb{Q}; <)$, and the complexity of $\mathrm{CSP}(\mathcal{C})$ has been classified in [23].*
3. *$\mathcal{C}$ is a reduct of $(\mathbb{Z}; <)$ and preserved by max or by min. In this case, $\mathrm{CSP}(\mathcal{C})$ is in P.*
4. *$\mathcal{C}$ is a reduct of $(\mathbb{Z}; \mathrm{succ})$ such that $\mathcal{C}$ is preserved by a modular max or a modular min, or such that $\mathbb{Q}.\mathcal{C}$ is preserved by a binary injective function preserving succ. In this case, $\mathrm{CSP}(\mathcal{C})$ is in P.*

5. $\mathrm{CSP}(\mathcal{B})$ *is* NP*-complete.*

As an illustration of the algorithmic consequences of our main result, we give examples of computational problems that can be formulated as discrete temporal CSPs and are in P.

Example 12. Fix positive integers $d, C \geq 1$.

Input: A system of constraints of the form ($x = y \bmod d$ and $a \leq x - y \leq b$) where $a, b \in \mathbb{Z}$ are such that $|a|, |b| \leq C$.

Question: Is the system satisfiable in $\mathbb{Z}^n$?

This problem can be seen as $\mathrm{CSP}(\mathbb{Z}; \mathrm{Diff}_{S_1}, \ldots, \mathrm{Diff}_{S_k})$ where $S_1, \ldots, S_k$ are all the sets of the form $\{a, a+d, \ldots, b\}$ for $a, b \in \mathbb{Z}$, $|a|, |b| \leq C$, and $d|(b-a)$. All the relations are preserved by the d-modular maximum function, and thus Theorem 6.5 implies that this CSP is in P.

Example 13. Consider the reduct $(\mathbb{Z}; R, \mathrm{succ})$ of $(\mathbb{Z}; <)$ where

$$R := \{(x, y, z) \in \mathbb{Z} \mid x \leq \max(y, z)\}$$

The relations R and succ are preserved by the (regular) maximum function, and thus Theorem 6.5 implies that this CSP is in P. The problem $\mathrm{CSP}(\mathbb{Z}; R, \mathrm{succ})$ is easily seen to be equivalent to the so-called *Max-Atom problem* [8] where numbers are represented in unary, which is known to be in P; see Section 6.7.

6.4 Definability of Successor and Order

The goal of this section is a proof that the CSPs for first-order reducts of $(\mathbb{Z}; <)$ fall into five classes. This will allow us to focus in later sections on first-order reducts of $(\mathbb{Z}; <)$ where succ is pp-definable.

Theorem 6.6. *Let $\mathcal{B}$ be a first-order reduct of $(\mathbb{Z}; <)$ with finite signature. Then* $\mathrm{CSP}(\mathcal{B})$ *equals* $\mathrm{CSP}(\mathcal{C})$ *where $\mathcal{C}$ is one of the following:*

1. *a finite structure;*
2. *a first-order reduct of $(\mathbb{Q}; <)$;*
3. *a first-order reduct of $(\mathbb{Z}; <)$ where* Dist_k *is pp-definable for all $k \geq 1$;*
4. *a first-order reduct of $(\mathbb{Z}; <)$ where* succ *and $<$ are pp-definable;*
5. *a first-order reduct of $(\mathbb{Z}; \mathrm{succ})$ where* succ *is pp-definable.*

The proof of this result requires some effort and spreads over the following subsections. Before we go into this, we explain the significance of the five classes for the CSP.

The first class is known to exhibit a complexity dichotomy (Theorem 2.5). The CSPs for first-order reducts of $(\mathbb{Q}; <)$ have been studied by Bodirsky and Kára [23]; they are either in P or NP-complete. Hence, we are done if there exists a first-order reduct $\mathcal{C}$ of $(\mathbb{Q}; <)$ such that $\mathrm{CSP}(\mathcal{C}) = \mathrm{CSP}(\mathcal{B})$. Several equivalent characterisations of those first-order reducts $\mathcal{B}$ will be given in Section 6.4.4. This is essential for proving Theorem 6.6.

When $\mathcal{B}$ is a first-order reduct of $(\mathbb{Z}; <)$ where for all $k \geq 1$ the relation Dist_k is pp-definable, then $\mathrm{CSP}(\mathcal{B})$ is NP-complete; this is a consequence of Proposition 27 from [15], restated here.

Proposition 6.7. *Suppose that the relations* Dist_1 *and* Dist_5 *are pp-definable in* $\mathcal{B}$*. Then* $\mathrm{CSP}(\mathcal{B})$ *is NP-hard.*

The previous paragraphs explain why Theorem 6.6 indeed reduces the complexity classification of CSPs for finite-signature first-order reducts $\mathcal{B}$ of $(\mathbb{Z};<)$ to the case where succ is pp-definable in $\mathcal{B}$, which corresponds to the classes (4) and (5) of Theorem 6.6.

6.4.1 Degrees

We consider three notions of *degree* for relations R that are first-order definable in $(\mathbb{Z};<)$:

- For $x \in \mathbb{Z}$, we consider the number of $y \in \mathbb{Z}$ that appear together with x in a tuple from R; this number is the same for all $x \in \mathbb{Z}$, and called the *Gaifman-degree* of R (it is the degree of the Gaifman graph of $(\mathbb{Z};R)$).
- The *distance degree* of R is the supremum of d such that there are $x, y \in \mathbb{Z}$ that occur together in a tuple of R and $|x-y| = d$.
- The *quantifier-elimination degree (qe-degree)* of R is the minimal q so that there is a quantifier-free definition ϕ of R, such that for every literal $x \leq y + c$ in ϕ, we have $|c| \leq q$.

The degree of a first-order reduct of $(\mathbb{Z};<)$ is the supremum of the degrees of its relations, for any of the three notions of degree. The article [15] considered first-order reducts of $(\mathbb{Z};\mathrm{succ})$ with finite Gaifman-degree. Note that the Gaifman-degree is finite if and only if the distance degree is finite. In this chapter, qe-degree will play the central role, as any first-order reduct of $(\mathbb{Z};<)$ with finite relational signature has finite qe-degree.

6.4.2 Compactness

In this section we present some results, based on applications of König's tree lemma, that show how properties of finite substructures of finite-signature first-order reducts $\mathcal{B}$ of $(\mathbb{Z};<)$ correspond to the existence of certain homomorphisms from $\mathcal{B}$ to $\mathbb{Q}.\mathcal{B}$.

Let $(\kappa.\mathbb{Z};<)$ be a model of $\mathrm{Th}(\mathbb{Z};<)$, let S be any set, let $s \in \mathbb{N}$, and $f\colon S \to \kappa.\mathbb{Z}$. We say that $x, y \in S$ are (f,s)*-connected* if there is a sequence $x = u_1, \ldots, u_k = y \in S$ so that $0 \leq |f(u_i) - f(u_{i+1})| \leq s$ for all $i \in \{1, \ldots, k-1\}$. Note that this notion of connectivity defines an equivalence relation on S whose equivalence classes are naturally ordered. We define an equivalence relation $\sim_s$ on functions $f, g\colon S \to \kappa.\mathbb{Z}$ as follows: $f \sim_s g$ when the following conditions are met:

- $x, y \in S$ are (f,s)-connected if and only if they are (g,s)-connected,
- if $x, y \in S$ are (f,s)-connected (and therefore (g,s)-connected) then $f(x) - f(y) = g(x) - g(y)$,
- if $x, y \in S$ are not (f,s)-connected then $f(x) < f(y) \Leftrightarrow g(x) < g(y)$.

In other words, $f \sim_s g$ iff the equivalence relations defined by (f,s)-connectivity and (g,s)-connectivity have the same equivalence classes, are such that within each equivalence class the pairwise distances are the same, and the order of the equivalence classes is the same. This implies that if S is a finite set, there are only finitely many $\sim_s$-equivalence classes of functions $S \to \kappa.\mathbb{Z}$. Note that if $f \sim_s g$ and $s' \leq s$ then we also have $f \sim_{s'} g$.

Lemma 6.8 (Substitution Lemma)**.** *Let $\mathcal{B}$ be a first-order reduct of $(\mathbb{Z};<)$ with qe-degree q, and let $\mathcal{C}$ be a structure with the same signature as $\mathcal{B}$ and domain D. Let κ be a linearly ordered set. Let $f, g\colon D \to \kappa.\mathbb{Z}$ be such that $f \sim_q g$. Then f is a homomorphism from $\mathcal{C}$ to $\kappa.\mathcal{B}$ if and only if g is such a homomorphism.*

Proof. Suppose that f is a homomorphism from $\mathcal{C}$ to $\kappa.\mathcal{B}$. To prove that g is a homomorphism, it suffices to prove that $g(a) < g(b) + c$ if and only if $f(a) \leq f(b) + c$ for all $a, b \in D$ and $|c| \leq q$. This follows from the fact that every relation of $\mathcal{B}$ can be defined from literals of the form $x \leq y + c$ with $|c| \leq q$ using conjunctions and disjunctions. Let $a, b \in D$ and suppose that $f(a) \leq f(b) + c$. If a, b are (f, q)-connected, we have $g(b) - g(a) = f(b) - f(a) \geq c$ whence $g(a) \leq g(b) + c$. If a, b are not (f, q)-connected, we have in particular $|f(a) - f(b)| > q$ and $|g(a) - g(b)| > q$. This implies that if $f(a) < f(b)$ then $g(a) < g(b) - q \leq g(b) - |c| \leq g(b) + c$, so $g(a) \leq g(b) + c$. On the other hand, if $f(b) < f(a)$ then $f(b) + q < f(a)$. This gives $q < f(a) - f(b) \leq c$, a contradiction to $|c| \leq q$. □

Lemma 6.9. *Let S be a subset of $\mathbb{Q}.\mathbb{Z}$ and let $(a_i)_{i\in\mathbb{N}}$ be an enumeration of S. Let $(F_i)_{i\in\mathbb{N}}$ be a sequence of $\sim_s$-equivalence classes of functions from $\{a_0, \ldots, a_i\} \to \mathbb{Q}.\mathbb{Z}$, for some $s \in \mathbb{N}$, such that $g \in F_j$ and $i < j$ imply that $g|_{\{a_0,\ldots,a_i\}} \in F_i$. Then there exists a function $h\colon S \to \mathbb{Q}.\mathbb{Z}$ such that $h|_{\{a_0,\ldots,a_i\}} \in F_i$ for all i and if $x, y \in S$ are not (g, s)-connected for any $g \in \bigcup_i F_i$, then $h(x) - h(y) = \infty$.*

Proof. We first outline the strategy of the proof. We build the function h as a set-theoretic union of functions $h_i\colon \{a_0, \ldots, a_i\} \to \mathbb{Q}.\mathbb{Z}$. We force that at each step i, the function h_i is in F_i and satisfies $h_i(a_k) - h_i(a_l) < \infty$ if and only if (a_k, a_l) are (g, s)-connected for some $j \geq i$ and some $g \in F_j$. The technicality of the proof comes from the fact that although we build the functions h_i by induction, we have to look ahead before choosing whether two points have to be mapped to different copies of $\mathbb{Z}$ in $\mathbb{Q}.\mathbb{Z}$ and to which copy they can be mapped.

We define the function h by induction. We require that at each step, the function $h_i\colon \{a_0, \ldots, a_i\} \to \mathbb{Q}.\mathbb{Z}$ that we define is in F_i and

- whenever $a, b \in \{a_0, \ldots, a_i\}$ are not (g, s)-connected for *any* function g in some F_j, then $h_i(a) - h_i(b) = \infty$, and
- if $a, b \in \{a_0, \ldots, a_i\}$ are (g, s)-connected for *some* function $g \in F_j$, then $h_i(a) - h_i(b) = g(a) - g(b)$.

For $i = 0$, let h_0 be any function in F_0. Suppose now that h_i has been defined, and let $h_{i+1}(a_k) := h_i(a_k)$ for $k \in \{0, \ldots, i\}$. Let $g \in F_j$ be such that for every pair $a_k, a_l \in \{a_0, \ldots, a_{i+1}\}$, if there exist $j' \geq 0$ and $g' \in F_{j'}$ such that (a_k, a_l) are (g', s)-connected, then (a_k, a_l) are (g, s)-connected: such a function exists, by taking j sufficiently large so that $\{a_0, \ldots, a_j\}$ contains all the elements that witness that a_k, a_l are (g', s)-connected for some g'. From the induction hypothesis and the assumptions, we know that $h_i \sim_s g|_{\{a_0,\ldots,a_i\}}$. Define $h_{i+1}(a_{i+1})$ as follows:

1. If there exists $k \in \{0, \ldots, i\}$ such that a_{i+1} and a_k are (g, s)-connected. Define $h_{i+1}(a_{i+1}) := h_i(a_k) - g(a_k) + g(a_{i+1})$. This first case is depicted in Figure 6.1.

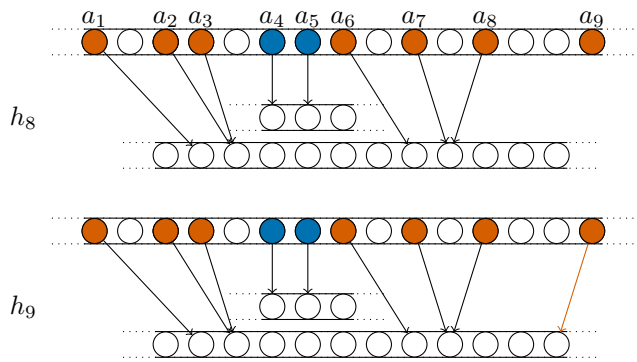

Figure 6.1: Illustration for item (1) of the proof of Lemma 6.9. We consider the case $i = 8$. The domain of h_8 is depicted above, and the copies of $\mathbb{Z}$ intersecting the image of h_8 are depicted below. Here, the image of h_8 intersects two copies of $\mathbb{Z}$. The colours represent equivalence classes of (g, s)-connectedness, where g is a function of some F_j that connects all the points of $\{a_1, \dots, a_9\}$ that eventually become connected. Since a_9 is in the same class as some previous point, we are in case (1) of the construction. Supposing that $g(a_9) = g(a_8) + 3$, we build h_9 by setting $h_9(a_9) = h_9(a_8) + 3$.

2. Otherwise consider the sets

$$U := \{u \in \mathbb{Q} \mid \exists k \in \{0, \dots, i\} : g(a_k) < g(a_{i+1}) \text{ and } h_i(a_k) \in u.\mathbb{Z}\}$$

and

$$V := \{v \in \mathbb{Q} \mid \exists k \in \{0, \dots, i\} : g(a_{i+1}) < g(a_k) \text{ and } h_i(a_k) \in v.\mathbb{Z}\}.$$

We have $U < V$. Indeed, let $u \in U$, $v \in V$, and let $k, l \in \{0, \dots, i\}$ be such that $h_i(a_k) \in u.\mathbb{Z}$ with $g(a_k) < g(a_{i+1})$ and $h_i(a_l) \in v.\mathbb{Z}$ with $g(a_{i+1}) < g(a_l)$. Since a_{i+1} is not (g, s)-connected to some element of $\{a_0, \dots, a_i\}$, we have that a_k and a_l are not (g, s)-connected. By construction, we therefore have that $h_i(a_k) - h_i(a_l) = \infty$. Since a_k and a_l are not (g, s)-connected and since $g(a_k) < g(a_l)$, we have that $h_i(a_k) < h_i(a_l)$. It follows that $u < v$. Thus, there exists $r \in \mathbb{Q}$ such that $U < r < V$. Define $h_{i+1}(a_{i+1}) := (r, 0)$. The situation is depicted in Figure 6.2.

We now prove that the induction hypothesis remains true for h_{i+1}. We claim that $h_{i+1} \sim_s g|_{\{a_0,\dots,a_{i+1}\}}$. Remember that we already know that $h_i \sim_s g|_{\{a_0,\dots,a_i\}}$ since $h_i \in F_i$ by induction and $g \in F_j$ for $j > i$. Let $a_j \in \{a_0, \dots, a_i\}$. If $h_{i+1}(a_{i+1})$ is at finite distance from $h_{i+1}(a_j)$, then by definition a_j, a_{i+1} are (g, s)-connected. Let $k \in \{0, \dots, i\}$ be the index used in the definition of h_{i+1}. We then have

$$\begin{aligned}
&h_{i+1}(a_{i+1}) - h_{i+1}(a_j) \\
&= h_i(a_k) - g(a_k) + g(a_{i+1}) - h_i(a_j) \\
&= g(a_k) - g(a_j) - g(a_k) + g(a_{i+1}) \qquad \text{(since } h_i(a_k) - h_i(a_j) = g(a_k) - g(a_j)\text{)} \\
&= g(a_{i+1}) - g(a_j).
\end{aligned}$$

It follows that a_{i+1}, a_j are (h_{i+1}, s)-connected iff they are $(g|_{\{a_0,\dots,a_{i+1}\}}, s)$-connected. If $h_{i+1}(a_{i+1})$ and $h_{i+1}(a_j)$ are at infinite distance, then a_{i+1}, a_j are neither (h_{i+1}, s)-connected nor (g, s)-connected. Then $h_{i+1}(a_{i+1}) < h_{i+1}(a_j) \Leftrightarrow g(a_{i+1}) < g(a_j)$ from

the construction of U and V. It follows that $h_{i+1} \sim_s g|_{\{a_0,\ldots,a_{i+1}\}}$. Moreover, h_{i+1} indeed separates integers that are never (g,s)-connected for any $g \in F_j$. Finally, if $g' \in F_{j'}$ is such that a, b are (g',s)-connected then a and b are also (g,s)-connected and $g'(a) - g'(b) = g(a) - g(b)$. This proves that h_{i+1} satisfies the induction hypothesis. Then $h := \bigcup_{i \geq 0} h_i$ satisfies the conclusion of the statement. □

The two previous lemmas will be applied frequently; one application is in the proof of the following proposition. Note that this makes essential use of the saturated model.

Proposition 6.10. *Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z};<)$. Then for all $a_1, a_2 \in \mathbb{Z}$ either*

- *there is an $r \geq 0$ and a finite $S \subseteq \mathbb{Z}$ that contains $\{a_1, a_2\}$ such that for all homomorphisms f from $\mathcal{B}[S]$ to $\mathcal{B}$ we have $|f(a_1) - f(a_2)| \leq r$, or*
- *there is a homomorphism h from $\mathcal{B}$ to $\mathbb{Q}.\mathcal{B}$ such that $h(a_1) - h(a_2) = \infty$.*

Proof. Let $a_1, a_2 \in \mathbb{Z}$ be arbitrary. Suppose that for all $r \geq 0$ and all finite $S \subset \mathbb{Z}$ containing $\{a_1, a_2\}$ there is a homomorphism f from $\mathcal{B}[S]$ to $\mathcal{B}$ such that $|f(a_1) - f(a_2)| > r$. We will describe how to construct the desired homomorphism h.

Let $a_1, a_2, a_3, \ldots$ be an enumeration of $\mathbb{Z}$, and let q be the qe-degree of $\mathcal{B}$. Consider the following infinite tree $\mathcal{T}$ whose vertices lie on levels $1, 2, \ldots$ The vertices at the n-th level are the $\sim_q$-equivalence classes of homomorphisms f from $\mathcal{B}[\{a_1, \ldots, a_{n+1}\}] \to \mathbb{Q}.\mathcal{B}$ such that a_1, a_2 are not (f,q)-connected (note that by Lemma 6.8, every element in the equivalence class of such a homomorphism is also a homomorphism). We have an arc in $\mathcal{T}$ from an equivalence class F on level n to an equivalence class G on level $n+1$ if there are $f \in F$, $g \in G$ such that f is the restriction of g. By assumption, $\mathcal{T}$ has vertices on each level n: indeed, at level n it suffices to take an f such that $|f(a_1) - f(a_2)| > qn$, and

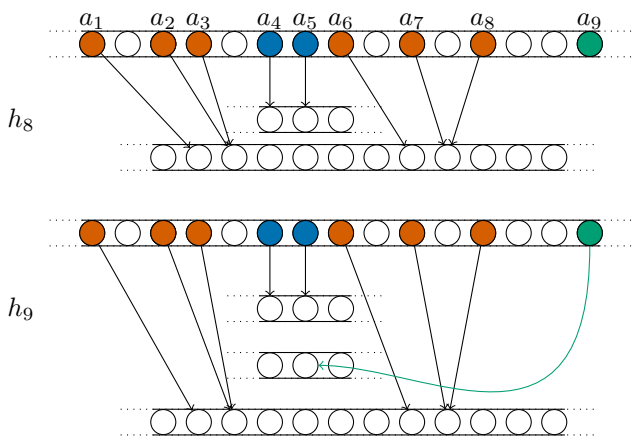

Figure 6.2: Illustration for item (2) of the proof of Lemma 6.9. Here, a_9 is not in the same equivalence class as any of the previous points. Assume that $g(a_8) < g(a_9) < g(a_4)$. We then find a copy of $\mathbb{Z}$ between the copies containing $h_8(a_8)$ and $h_8(a_4)$ and not containing any points of the image of h_8. We set $h_9(a_8)$ to be an arbitrary point in this new copy.

such an f exists by assumption. The tree $\mathcal{T}$ has finitely many vertices on each level, since the number of $\sim_q$-equivalence classes of homomorphisms from $\mathcal{B}[\{a_1, \ldots, a_n\}] \to \mathbb{Q}.\mathcal{B}$ is finite.

It follows by König's lemma that there is an infinite branch $\mathcal{B}$ of $\mathcal{T}$. By Lemma 6.9 applied with $S := \mathbb{Z}$ and $\ell := q$ and using the elements of $\mathcal{B}$ for the sequence $(F_i)_{i\in\mathbb{N}}$, there exists a function $h\colon \mathbb{Z} \to \mathbb{Q}.\mathbb{Z}$ such that $h|_{\{a_1,\ldots,a_i\}}$ is in the branch $\mathcal{B}$ for every $i \in \mathbb{N}$, and $h(a_1) - h(a_2) = \infty$ (since a_1, a_2 are not connected by any function in the branch $\mathcal{B}$). Finally, h is a homomorphism $\mathcal{B} \to \mathbb{Q}.\mathcal{B}$ by Lemma 6.8. □

Definition 6.5. A mapping $h\colon \kappa_1.\mathbb{Z} \to \kappa_2.\mathbb{Z}$ is called *isometric* if $|h(x) - h(y)| = |x - y|$ for all $x, y \in \kappa_1.\mathbb{Z}$.

The following proposition can be shown by straightforward modifications of the proof of Proposition 6.10.

Proposition 6.11. *Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z}; <)$. Then either*

- *for every $r \in \mathbb{N}$ there is a finite $S \subseteq \mathbb{Z}$ containing $\{0, r\}$ such that for all homomorphisms f from $\mathcal{B}[S]$ to $\mathcal{B}$ we have $|f(0) - f(r)| = r$, or*
- *there is a homomorphism h from $\mathcal{B}$ to $\mathbb{Q}.\mathcal{B}$ which is not isometric.*

6.4.3 Finite-range Endomorphisms

In this section we present a lemma that gives a useful sufficient condition for $\mathcal{B}$ to have endomorphisms with finite range. Note that $\mathcal{B}$ has a finite-range endomorphism if and only if there exists a finite structure $\mathcal{C}$ such that $\mathrm{CSP}(\mathcal{B}) = \mathrm{CSP}(\mathcal{C})$. We need the following combinatorial definitions and lemmas about the integers.

We say that $T \subseteq \mathbb{Z}$ *contains arbitrarily long intervals* when for every $m \in \mathbb{N}$ there exists $z \in \mathbb{Z}$ so that $[z, z + m] \subset T$. A sequence $u_1, \ldots, u_r$ is called a *($\leq m$)-progression* if $1 \leq u_{i+1} - u_i \leq m$ for all $i < r$. We say that T has *arbitrarily long ($\leq m$)-progressions* if for every $r \in \mathbb{N}$ the set T contains a ($\leq m$)-progression $u_1, \ldots, u_r$. Clearly, if $\mathbb{Z} \setminus T$ does not have arbitrarily long intervals then there exists an $m \in \mathbb{N}$ so that T has arbitrarily long ($\leq m$)-progressions.

Lemma 6.12. *Let $T \subseteq \mathbb{Z}$ contain arbitrarily long ($\leq m$)-progressions, and let $T = T_1 \cup \cdots \cup T_k$ be a partition of T into finitely many sets. Then there exists an $i \leq k$ and an $m' \in \mathbb{N}$ such that T_i contains arbitrarily long ($\leq m'$)-progressions.*

Proof. If there exists an $m' \in \mathbb{N}$ such that T_1 contains arbitrarily long ($\leq m'$)-progressions, then there is nothing to show. So suppose that this is not the case.

We will show that $T' := T \setminus T_1$ contains arbitrarily long ($\leq m$)-progressions; the statement then follows by induction. Let $s \in \mathbb{N}$ be arbitrary. We want to find a ($\leq m$)-progression $u_1, \ldots, u_s$ in T'. By the above assumption, T_1 does not contain arbitrarily long ($\leq ms$)-progressions, and hence there exists an r such that T_1 does not contain a ($\leq ms$)-progression of length r.

Since T contains arbitrarily long ($\leq m$)-progressions, it contains in particular an ($\leq m$)-progression ρ of length msr. Consider the first s elements of ρ. If all those elements are in T' we have found the desired ($\leq m$)-progression of length s, and are done. So suppose

otherwise; that is, at least one of those first s elements must be from T_1. We apply the same argument to the next s elements of ρ, and can again assume that at least one of those elements must be from T_1. Continuing like this, we find a subsequence of ρ of elements of T_1 which form a $(\leq ms)$-progression. The length of this subsequence is $msr/ms = r$. But this contradicts our assumption that T_1 does not contain $(\leq ms)$-progression of length r. □

Lemma 6.13. *Let $m \in \mathbb{N}$ and let $T \subseteq \mathbb{Z}$ be with arbitrarily long $(\leq m)$-progressions. Then for all $S \subset \mathbb{Z}$ of cardinality $m+1$ there are $x_1, x_2 \in S$ and $y_1, y_2 \in T$ such that $x_1 - x_2 = y_1 - y_2$.*

Proof. Let r be greater than $\max(S) - \min(S)$. Then there exists an $(\leq m)$-progression $w_1, \ldots, w_r$ in T. Define $T_i := \{z - w_1 + \min(S) + i \mid z \in T\}$. Then $T_0 \cup \cdots \cup T_{m-1}$ includes the entire interval $[\min(S), \max(S)]$. By the pigeon-hole principle there is an i such that $|T_i \cap S| \geq 2$, which clearly implies the statement. □

Lemma 6.14. *Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z}; <)$ and h a homomorphism from $\mathcal{B} \to \mathbb{Q}.\mathcal{B}$. Let $S \subseteq \mathbb{Z}$ be finite and $z_0 \in \mathbb{Z}$. If $(\mathbb{Z} \setminus h^{-1}(S)) \cap \{z \in \mathbb{Z} : z \geq z_0\}$ does not contain arbitrarily long intervals then $\mathcal{B}$ has a finite-range endomorphism.*

Proof. Since $(\mathbb{Z} \setminus h^{-1}(S)) \cap \{z \in \mathbb{Z} : z \geq z_0\}$ does not contain arbitrarily long intervals, there exists an $m' \in \mathbb{N}$ such that $T := h^{-1}(S)$ contains arbitrarily long $(\leq m')$-progressions. Suppose that $S = \{s_1, \ldots, s_k\}$, and define $T_i := h^{-1}(s_i)$ for $i \in \{1, \ldots, k\}$. Then by Lemma 6.12 there exists an $m \in \mathbb{N}$ and an $i \leq k$ such that T_i contains arbitrarily long $(\leq m)$-progressions.

Our argument is based on König's tree lemma, involving the finitely branching infinite tree $\mathcal{T}$ defined similarly as in the proof of Proposition 6.10. Let $a_0, a_1, \ldots$ be an enumeration of $\mathbb{Z}$, and let q be the qe-degree of $\mathcal{B}$. The vertices of $\mathcal{T}$ on the n-th level are the $\sim_q$-equivalence classes of homomorphisms g from $\mathcal{B}[\{a_0, \ldots, a_n\}]$ to $\mathcal{B}$ such that $|g(\{a_0, \ldots, a_n\})| \leq m$. Adjacency is defined by restriction, and $\mathcal{T}$ is finitely branching, as in the proof of Proposition 6.10.

We show that $\mathcal{T}$ has vertices on all levels n by induction on n. We prove that for any finite $X \subset \mathbb{Z}$ there exists a homomorphism $g \colon \mathcal{B}[X] \to \mathcal{B}$ whose range has size at most m. For $|X| \leq m$, this is witnessed by the restriction of the identity function to X. Now let $|X| = n+1$ for $n \geq m$. By Lemma 6.13, there are $x_1, x_2 \in X$ and $y_1, y_2 \in T_i$ such that $x_1 - x_2 = y_1 - y_2$. We therefore have that $f \colon x \mapsto h(x - x_1 + y_1)$ is a homomorphism $\mathcal{B}[X] \to \mathbb{Q}.\mathcal{B}$ whose range has size at most n. Indeed, we have $f(x_1) = h(y_1) = h(y_2) = h(x_2 - x_1 + y_1) = f(x_2)$. Let g be given by the induction hypothesis applied to the image of f. We then have that $g \circ f$ is a homomorphism $\mathcal{B}[X] \to \mathcal{B}$ whose range has size at most m, and the claim is proved.

Hence, $\mathcal{T}$ has vertices on all levels, and therefore an infinite branch $\mathcal{B}$ by König's lemma. By Proposition 6.9 applied to this infinite branch, $S := \mathbb{Z}$, and $\ell := q$ there exists a function $h \colon \mathbb{Z} \to \mathbb{Q}.\mathbb{Z}$ such that $h|_{\{a_0, \ldots, a_i\}} \in \mathcal{B}$ for all $i \in \mathbb{N}$. In particular, the range of h has size at most m. Up to $\sim_q$-equivalence, we can assume that the image of h lies in one copy of $\mathbb{Z}$ in $\mathbb{Q}.\mathbb{Z}$, say in $\mathbb{Z}$. Then Lemma 6.8 implies that h is a homomorphism $e \colon \mathcal{B} \to \mathcal{B}$ whose range has cardinality at most m, concluding the proof. □

The next lemma is an important consequence of Lemma 6.14.

Lemma 6.15. *Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z};<)$ without finite-range endomorphisms, $\ell \in \mathbb{N}$, and h a homomorphism from $\mathcal{B}$ to $\mathbb{Q}.\mathcal{B}$. Then there exists an $e \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$ such that for all $x, y \in \mathbb{Q}.\mathbb{Z}$ with $x - y = \infty$ we have $e(x) - e(y) = \infty$, and such that $h \sim_\ell e|_\mathbb{Z}$.*

Proof. We first give an idea about the proof. Since $\mathcal{B}$ does not have finite-range endomorphisms, we know from the previous lemma that the preimage of any finite subset of $\mathbb{Q}.\mathbb{Z}$ under h leaves arbitrarily large gaps in $\mathbb{Z}$. It follows that for every finite subset S of $\mathbb{Q}.\mathbb{Z}$, there exists a homomorphism $p\colon \mathbb{Q}.\mathcal{B}[S] \to \mathcal{B}$ such that $h \circ p$ does not connect any pair of integers that sit in different copies. Since we have such homomorphisms for arbitrarily large finite subsets $S \subset \mathbb{Q}.\mathbb{Z}$, an application of König's lemma and Lemma 6.9 give the desired endomorphism of $\mathbb{Q}.\mathcal{B}$.

We now give the detailed argument. Note that if $h \sim_{\ell'} g$ and $\ell < \ell'$, then $h \sim_\ell g$. It follows that without loss of generality, we can assume that ℓ is greater than the qe-degree of $\mathcal{B}$. As in the proof of Proposition 6.10, we build e through an argument involving König's lemma and an infinite tree $\mathcal{T}$. Let $a_1, a_2, \dots$ be an enumeration of $\mathbb{Q}.\mathbb{Z}$. For the n-th level of $\mathcal{T}$ we will consider $\sim_\ell$-classes of homomorphisms f from $\mathbb{Q}.\mathcal{B}[\{a_1, \dots, a_n\}]$ to $\mathbb{Q}.\mathcal{B}$ with the property that

- for all $x, y \in \{a_1, \dots, a_n\}$ with $x - y = \infty$ the elements x, y are not (f, ℓ)-connected, and
- $f|_{\{a_1,\dots,a_n\}} \sim_\ell h|_{\{a_1,\dots,a_n\}}$.

Adjacency is defined by restriction as in the proof of Proposition 6.10.

The only difficulty of the proof is to show that $\mathcal{T}$ has vertices on all levels n. We will first construct a homomorphism p from $\mathbb{Q}.\mathcal{B}[\{a_1, \dots, a_n\}]$ to $\mathcal{B}$ with the property that $p(a_i) = a_i$ for a_i in the domain of h, and if $a_i - a_j = \infty$ for $i, j \leq n$, then $p(a_i)$ and $p(a_j)$ are not (h, ℓ)-connected. Let S be the set of points that are at distance at most ℓ from some $a_1, \dots, a_n$. Let $S_1 \cup \dots \cup S_k$ be the partition of S induced by the copies of $\mathbb{Z}$ in $\mathbb{Q}.\mathbb{Z}$, that is, $S_1, \dots, S_k$ are pairwise disjoint and each S_i only contains points that lie in the same copy of $\mathbb{Z}$ in $\mathbb{Q}.\mathbb{Z}$. Suppose without loss of generality that $S_1 < \dots < S_{m-1} < S_m < S_{m+1} < \dots < S_k$ and that $S_m \subset \mathbb{Z}$, the standard copy in $\mathbb{Q}.\mathbb{Z}$. For every $i \in \{1, \dots, k\}$, let s_i and t_i be the minimal and the maximal element of S_i, respectively. The situation is represented in Figure 6.3.

For the elements $x \in S_m$ we set $p(x) := x$. Let $Q_m = \{z \in \mathbb{Q}.\mathbb{Z} \mid \exists z' \in S_m : |h(z') - z| \leq \ell\}$. Write S'_m for $h^{-1}(Q_m)$. If $\mathbb{Z} \setminus S'_m \cap \{z \mid z > t_m\}$ does not contain arbitrarily long intervals, then $\mathcal{B}$ has a finite-range endomorphism by Lemma 6.14, contrary to our assumptions. So there exists a $z_m \in \mathbb{Z}$ greater than t_m such that $[z_m, z_m + t_{m+1} - s_{m+1} + 2\ell] \cap S'_m = \emptyset$. For $x \in S_{m+1}$, we set $p(x) := x - s_{m+1} + z_m + \ell$. The mapping is illustrated in Figure 6.4. As above, set Q_{m+1} to be the set of points that are at distance at most ℓ from a point in $h(p(S_m \cup S_{m+1}))$. Now, set $S'_{m+1} := h^{-1}(Q_{m+1})$. Then there exists a $z_{m+1} \in \mathbb{Z}$ such that $[z_{m+1}, z_{m+1} + t_{m+2} - s_{m+2} + 2\ell] \cap S'_{m+1} = \emptyset$. For $x \in S_{m+2}$, we set $p(x) := x - s_{m+2} + z_{m+1} + \ell$. Continuing in this way, we define p for all $x \in \{a_1, \dots, a_n\}$ (the construction for $i < m$ is symmetric). We have that p is a homomorphism $\mathbb{Q}.\mathcal{B}[\{a_1, \dots, a_n\}] \to \mathcal{B}$ since it is $\sim_\ell$-equivalent to the identity function

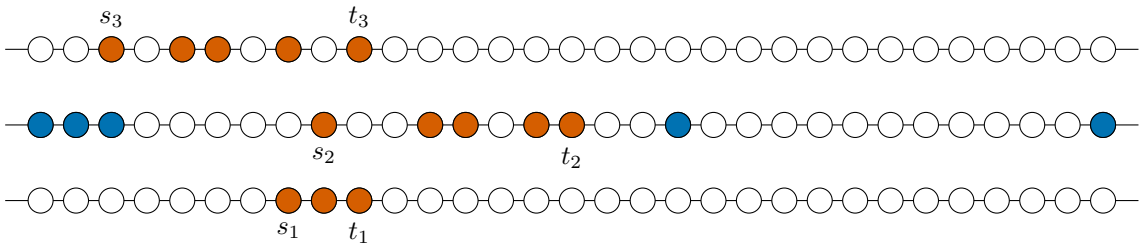

Figure 6.3: Illustration of the proof of Lemma 6.15. Here, $k = 3$, $\ell = 1$ and $m = 2$. The nodes coloured in red (light grey) are the integers in S_1, S_2, S_3. The nodes coloured in blue (dark grey) are the integers in $S_2' \setminus S_2$, that is, the integers that are mapped under h to integers near $h(S_2)$. The assumption that $\mathcal{B}$ does not have finite-range endomorphisms guarantees that there are arbitrarily long intervals of white nodes in the middle line, both on the left of s_2 and the right of t_2.

on $\mathbb{Q}.\mathbb{Z}[\{a_1, \ldots, a_n\}]$. Observe that by construction of p, when $a_i - a_j = \infty$, then a_i, a_j are not $(h \circ p, \ell)$-connected. Therefore the $\sim_q$-equivalence class of $h \circ p$ is a vertex of $\mathcal{T}$ on level n.

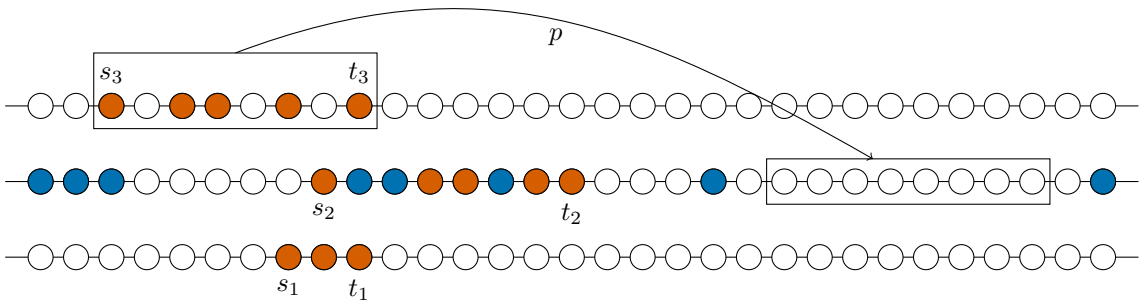

Figure 6.4: Illustration of the proof of Lemma 6.15, after the first step of the construction. The blue nodes (light grey) are now the integers in S_3' that are not in S_2 or in $p(S_3)$, that is, the integers that are mapped by h to integers near $h(S_2 \cup p(S_3))$.

The tree $\mathcal{T}$ is finitely branching, and by König's lemma it contains an infinite branch $\mathcal{B}$. By Lemma 6.9 applied to this branch, $S := \mathbb{Q}.\mathbb{Z}$, and ℓ as in the statement of Lemma 6.15 there exists a function $e \colon \mathbb{Q}.\mathbb{Z} \to \mathbb{Q}.\mathbb{Z}$ such that $e|_{\{a_1,\ldots,a_i\}} \in \mathcal{B}$ for all $i \in \mathbb{N}$ and if $x - y = \infty$ then $e(x) - e(y) = \infty$. By Lemma 6.8, e is an endomorphism of $\mathbb{Q}.\mathcal{B}$. We also have that $e|_{\mathbb{Z}} \sim_\ell h$ and hence e has the required properties. $\square$

6.4.4 Petrus

The following theorem is the rock upon which we build our church.

Theorem 6.16 (Petrus ordinis). *Let $\mathcal{B}$ be a first-order reduct of $(\mathbb{Z};<)$ with finite relational signature and without an endomorphism of finite range. Then the following are equivalent:*

(1) there exists a first-order reduct $\mathcal{C}$ of $(\mathbb{Q};<)$ such that $\mathrm{CSP}(\mathcal{C})$ equals $\mathrm{CSP}(\mathcal{B})$;

(2) for all $t \geq 1$, there is an $e \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$ and $z \in \mathbb{Q}.\mathbb{Z}$ such that $|e(z+t) - e(z)| > t$;

(3) all binary relations with a pp-definition in $\mathbb{Q}.\mathcal{B}$ are either empty, the equality relation, or have unbounded distance degree;

(4) for all distinct $z_1, z_2 \in \mathbb{Z}$ there is a homomorphism $h\colon \mathcal{B} \to \mathbb{Q}.\mathcal{B}$ such that $h(z_1) - h(z_2) = \infty$;

(5) for all distinct $z_1, z_2 \in \mathbb{Z}$ there is an $e \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$ such that $e(z_1) - e(z_2) = \infty$; and for all $z_1', z_2' \in \mathbb{Q}.\mathbb{Z}$ with $z_1' - z_2' = \infty$ we have $e(z_1') - e(z_2') = \infty$;

(6) there exists an $e \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$ with infinite range such that $e(x) - e(y) = \infty$ or $e(x) = e(y)$ for any two distinct $x, y \in \mathbb{Q}.\mathcal{B}$.

Proof. Throughout the proof, let q be the qe-degree of $\mathcal{B}$, which is finite since $\mathcal{B}$ has a finite signature.

(1) $\Rightarrow$ (2). Since $\mathcal{C}$ has the same CSP as $\mathcal{B}$, and $\mathcal{C}$ is ω-categorical, Lemma 3.1.5 in [11]) states that there is a homomorphism f from the countable structure $\mathbb{Q}.\mathcal{B}$ to $\mathcal{C}$. Lemma 6.3 asserts the existence of a homomorphism g from $\mathcal{C}$ to $\mathbb{Q}.\mathcal{B}$, because every pp-sentence that is true in $\mathcal{C}$ is also true in $\mathbb{Q}.\mathcal{B}$, and $\mathbb{Q}.\mathcal{B}$ is saturated.

Let $t \geq 1$. It is not possible that $f(z) = f(z+t)$ for all $z \in \mathbb{Q}.\mathbb{Z}$, for otherwise $\mathcal{B}$ would have a finite-range endomorphism. Indeed, we can restrict $g \circ f$ to a homomorphism $\mathcal{B} \to \mathbb{Q}.\mathcal{B}$ whose range is finite. We can then construct a function $e\colon \mathbb{Z} \to \mathbb{Q}.\mathbb{Z}$ such that $g \circ f \sim_q e$ and such that the range of e is contained in $\mathbb{Z}$. This e would then be an endomorphism of $\mathcal{B}$ by Lemma 6.8, a contradiction. Pick a $z \in \mathbb{Q}.\mathbb{Z}$ such that $f(z) \neq f(z+t)$. The range of g is infinite, for otherwise the range of $g \circ f$ would be finite. Thus, there are two rationals $p \neq p'$ such that $|g(p) - g(p')| > t$. Let α be an automorphism of $\mathcal{C}$ that maps $\{f(z), f(z+t)\}$ to $\{p, p'\}$. We now have $|(g \circ \alpha \circ f)(z+t) - (g \circ \alpha \circ f)(z)| = |g(p) - g(p')| > t$.

(2) $\Rightarrow$ (3). Let R be a binary relation with a pp-definition in $\mathbb{Q}.\mathcal{B}$. Suppose that R is not empty and is not the equality relation. Let k be the supremum of the integers t such that there exists $(z_1, z_2) \in R$ with $|z_1 - z_2| = t$. Since R is neither empty nor the equality relation, it follows that k is positive. If k is ∞, then R has infinite distance degree. Otherwise let (z_1, z_2) be a pair in R such that $|z_1 - z_2| = k$. Let e be an endomorphism of $\mathbb{Q}.\mathcal{B}$ and z be such that $|e(z+k) - e(z)| > k$. Let α be an automorphism of $\mathbb{Q}.\mathcal{B}$ that maps $\{z_1, z_2\}$ to $\{z, z+k\}$. Then $(e \circ \alpha)(z_1, z_2)$ is in R since R is preserved by the endomorphisms of $\mathbb{Q}.\mathcal{B}$ and by construction $|(e \circ \alpha)(z_1) - (e \circ \alpha)(z_2)| > k$, a contradiction to the choice of k.

(3) $\Rightarrow$ (4). Suppose that (4) does not hold, that is, there are distinct $a_1, a_2 \in \mathbb{Z}$ such that for all homomorphisms h from $\mathcal{B}$ to $\mathbb{Q}.\mathcal{B}$ we have that $h(a_1) - h(a_2) < \infty$. Then by Proposition 6.10 there is an $r \geq 0$ and a finite $S \subseteq \mathbb{Z}$ containing $\{a_1, a_2\}$ such that for all homomorphisms $f\colon \mathcal{B}[S] \to \mathcal{B}$ we have $|f(a_1) - f(a_2)| \leq r$. Now consider the following pp-formula ϕ: the variables of ϕ are the elements of S, all existentially quantified except a_1 and a_2, which are free. The formula ϕ contains the conjunct $R(x_1, \ldots, x_n)$ for a relation

R from $\mathcal{B}$ if and only if $\mathcal{B}[S] \models R(x_1, \ldots, x_n)$. Then ϕ defines a binary relation, which has bounded distance degree by the previous discussion, and which is not the equality relation since it contains the pair (a_1, a_2).

(4) $\Rightarrow$ (5). Let $z_1, z_2 \in \mathbb{Z}$ be distinct, let h be given by item (4), and let e be given by Lemma 6.15 applied to h for $\ell := q$. Pick any function $p\colon e(\mathbb{Q}.\mathbb{Z}) \to \mathbb{Q}.\mathbb{Z}$ such that if $x, y \in \mathbb{Q}.\mathbb{Z}$ are not (e, q)-connected then $(p \circ e)(x) - (p \circ e)(y) = \infty$ and such that $p \sim_q \mathrm{id}$. It is clear that such a function exists because $(\mathbb{Q}; <)$ embeds all countable linear orders. Indeed, consider the equivalence relation on $e(\mathbb{Q}.\mathbb{Z})$ where $x \sim y$ if there are $x := u_1, \ldots, u_k =: y \in e(\mathbb{Q}.\mathbb{Z})$ such that $|u_i - u_{i+1}| \leq q$ for all $i \in \{1, \ldots, k-1\}$. The equivalence classes induced by this relation are naturally ordered by setting $\rho < \pi$ if for all $x \in \rho, y \in \pi$, we have $x < y$. There are at most countably many equivalence classes, hence there exists an increasing function f from the set of equivalence classes to $\mathbb{Q}$. We let $p(a, z) := (f(\rho), z)$ where ρ is the equivalence class of (a, z). Then we have that $p \sim_q \mathrm{id}$, and this implies that $p \circ e \sim_q e$ so that $p \circ e$ is an endomorphism of $\mathbb{Q}.\mathcal{B}$ by the substitution lemma. Moreover, p is such that $x - y = \infty \Rightarrow (p \circ e)(x) - (p \circ e)(y) = \infty$. Finally, z_1 and z_2 are not (e, q)-connected because $e|_{\mathbb{Z}} \sim_q h$, so that $(p \circ e)(z_1) - (p \circ e)(z_2) = \infty$.

(5) $\Rightarrow$ (6). Again an argument based on König's tree lemma. Let $a_1, a_2, \ldots$ be an enumeration of $\mathbb{Q}.\mathbb{Z}$. Let $\mathcal{T}$ be a tree whose vertices on the n-th level are the $\sim_q$-equivalence classes of homomorphisms g from $\mathbb{Q}.\mathcal{B}[\{a_1, \ldots, a_n\}]$ to $\mathbb{Q}.\mathcal{B}$ such that for all $i, j \leq n$ either a_i and a_j are not (g, q)-connected or $g(a_i) = g(a_j)$. Adjacency of vertices is defined by restriction between representatives. We have to show that the tree has vertices on all levels. Let $\{u_1, v_1\}, \ldots, \{u_k, v_k\}$ be an enumeration of all 2-element subsets of $\{a_1, \ldots, a_n\}$. We will show by induction on $i \geq 0$ that there exists an endomorphism f_i such that $f_j(u_j) - f_j(v_j) = \infty$ or $f_j(u_j) = f_j(v_j)$ for all $j \leq i$. The statement is trivial for $i = 0$. So suppose we have already found f_i for some $i \geq 0$, and want to find f_{i+1}. If $f_i(u_{i+1}) - f_i(v_{i+1}) = \infty$ or $f_i(u_{i+1}) = f_i(v_{i+1})$ then there is nothing to show. Otherwise, let α be an automorphism of $\mathbb{Q}.\mathcal{B}$ that maps $f_i(u_{i+1})$ and $f_i(v_{i+1})$ to $\mathbb{Z}$. By (5), there exists an $e \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$ such that $e(\alpha(f_i(u_{i+1}))) - e(\alpha(f_i(v_{i+1}))) = \infty$, and such that for all $x, y \in \mathbb{Q}.\mathbb{Z}$ with $x - y = \infty$ we have that $e(x) - e(y) = \infty$. Hence, $f_{i+1} := e \circ \alpha \circ f_i$ has the desired property. The tree $\mathcal{T}$ has finitely many vertices on each level and hence must contain an infinite branch, which gives rise to an endomorphism of $\mathbb{Q}.\mathcal{B}$ by Lemmas 6.9 and 6.8.

(6) $\Rightarrow$ (1). Let $\mathcal{C}$ be the structure induced by $\mathbb{Q}.\mathcal{B}$ on the image of the endomorphism e whose existence has been asserted in (6). The structures $\mathcal{C}$ and $\mathcal{B}$ have the same CSP. Note that a literal $x \leq y + k$ for $k \in \mathbb{Z}$ is true in $\mathcal{C}$ iff $x \leq y$ is true. Therefore the relations of $\mathcal{C}$ are definable with quantifier-free formulas using only $x < y$ and $x = y$. It follows that $\mathcal{C}$ has the same CSP as a first-order reduct of $(\mathbb{Q}; <)$. $\square$

6.4.5 Boundedness and rank

Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z}; <)$ without a finite-range endomorphism. Theorem 6.16 (Petrus) characterized the "degenerate case" when $\mathrm{CSP}(\mathcal{B})$ is the CSP for a first-order reduct of $(\mathbb{Q}; <)$. For such $\mathcal{B}$, as we have mentioned before, the complexity of the CSP has already been classified. In the following we will therefore assume that the equivalent items of Theorem 6.16, and in particular, item (2), do *not* apply. To make the best use of those findings, we introduce the following terminology.

Definition 6.6. Let $k \in \mathbb{N}^+, c \in \mathbb{N}$. A function $e\colon \kappa_1.\mathbb{Z} \to \kappa_2.\mathbb{Z}$ is (k,c)-*bounded* if for all $u \in \kappa_1.\mathbb{Z}$ we have $|e(u+k) - e(u)| \leq c$.

We say that e is *tightly-k-bounded* if it is (k,k)-bounded, and *k-bounded* if it is (k,c)-bounded for some $c \in \mathbb{N}$. For given k, c, we say that $\kappa.\mathcal{B}$ is (k,c)-*bounded* (resp. *k-bounded*, tightly-k-bounded) if all its endomorphisms are. We call the smallest t such that $\kappa.\mathcal{B}$ is tightly-t-bounded the *tight rank* of $\kappa.\mathcal{B}$. Similarly, we call the smallest r such that $\kappa.\mathcal{B}$ is r-bounded the *rank* of $\kappa.\mathcal{B}$.

The negation of item (2) in Theorem 6.16 says that there exists a $t \in \mathbb{N}$ such that $\mathbb{Q}.\mathcal{B}$ is tightly-t-bounded. Clearly, being tightly-t-bounded implies being t-bounded. Hence, the negation of item (2) in Theorem 6.16 also implies that $\mathbb{Q}.\mathcal{B}$ has finite rank $r \leq t$.

Example 14. For $p > 0$, the structure $(\mathbb{Z}; \mathrm{succ}^p)$ of Example 11 (1) has rank and tight rank equal to p. The structure $(\mathbb{Z}; \mathrm{succ}^2, \mathrm{Diff}_{\{-2,-1,0,1,2\}})$ of Example 11 (3) is an example whose rank is 1 and whose tight rank is greater (it is equal to 2).

Sections 6.4.5 and 6.4.5 are devoted to proving that one can replace $\mathcal{B}$ by another first-order reduct $\mathcal{C}$ of $(\mathbb{Z}; <)$ which has the same CSP and such that $\mathbb{Q}.\mathcal{C}$ has both rank one and tight rank one.

Example 15. There are rank one first-order reducts of $(\mathbb{Z}; <)$ which do have non-injective endomorphisms, but no finite-range endomorphisms. Consider the third structure in Example 11:

$$\mathcal{B} := (\mathbb{Z}; \mathrm{succ}^2, \mathrm{Diff}_{\{-2,-1,0,1,2\}}) \,.$$

Note that $\mathcal{B}$ has rank one: as every endomorphism e preserves the relation $\mathrm{Diff}_{\{-2,-1,0,1,2\}}$ we have $|e(x+1) - e(x)| \leq 2$. Also note that $\mathcal{B}$ has the non-injective endomorphism e defined by $e(x) = x$ for even x, and $e(x) = x+1$ for odd x.

Corollary 6.17. *Let $\mathcal{B}$ be a finite-signature reduct of $(\mathbb{Z}; <)$ without finite-range endomorphisms. Then $\mathbb{Q}.\mathcal{B}$ has finite rank if and only if $\mathbb{Q}.\mathcal{B}$ has finite tight rank.*

Proof. We have just seen that having finite tight rank implies having finite rank. Conversely, when $\mathbb{Q}.\mathcal{B}$ has finite rank, then item (5) in Theorem 6.16 is false. Then Theorem 6.16 implies that item (2) is false, too, which is to say that $\mathbb{Q}.\mathcal{B}$ has finite tight rank. □

We also make the following important observation.

Lemma 6.18. *Let $\mathcal{B}$ be a finite-signature reduct of $(\mathbb{Z}; <)$ without finite-range endomorphisms and such that $\mathbb{Q}.\mathcal{B}$ has finite rank r. Then there exists a $c \geq 0$ such that every $e \in \mathrm{End}(\mathcal{B})$ is (r,c)-bounded.*

Proof. Let $a_1 < a_2$ be two integers at distance r. We know from the negation of item (4) in Theorem 6.16 that every homomorphism $h\colon \mathcal{B} \to \mathbb{Q}.\mathcal{B}$ satisfies $h(a_1) - h(a_2) < \infty$. Proposition 6.10 gives a $c \geq 0$ and a finite $S \subset \mathbb{Z}$ containing a_1, a_2 such that every homomorphism $f\colon \mathcal{B}[S] \to \mathcal{B}$ satisfies $|f(a_1) - f(a_2)| \leq c$. In particular, every endomorphism f of $\mathcal{B}$ also satisfies this.

To prove that every endomorphism of $\mathcal{B}$ is (r,c)-bounded, let now $f \in \mathrm{End}(\mathcal{B})$ and $a \in \mathbb{Z}$. Let α be the automorphism of $(\mathbb{Z}; <)$ that maps a_1 to a. By the paragraph above applied to the endomorphism $f \circ \alpha$, we have $|(f \circ \alpha)(a_1) - (f \circ \alpha)(a_2)| \leq c$, i.e., $|f(a) - f(a+r)| \leq c$. This proves that f is (r,c)-bounded. □

The next lemma connects the rank and the tight rank of a structure and its countable saturated extension.

Lemma 6.19. *Let $\mathcal{B}$ be a first-order reduct of $(\mathbb{Z};<)$ with finite relational signature such that $\mathbb{Q}.\mathcal{B}$ has rank r and tight rank t. Then $\mathcal{B}$ has rank $r' \leq r$ and tight rank $t' \leq t$.*

Proof. Let f be an endomorphism of $\mathcal{B}$, and let $a \in \mathbb{Z}$. Let $\ell = \max\{|f(a+r) - f(a)|, q\}$. We view f as a homomorphism $\mathcal{B} \to \mathbb{Q}.\mathcal{B}$ and find an endomorphism e of $\mathbb{Q}.\mathcal{B}$ such that $e|_{\mathbb{Z}} \sim_\ell f$ by Lemma 6.15. There exists a $c > 0$ such that the endomorphism e is (r,c)-bounded, by assumption on $\mathbb{Q}.\mathcal{B}$. This gives $|f(a+r) - f(a)| = |e(a+r) - e(a)| \leq c$, i.e., f is (r,c)-bounded. Therefore, every endomorphism of $\mathcal{B}$ is r-bounded and $\mathcal{B}$ has finite rank $r' \leq r$. We prove similarly that every endomorphism of $\mathcal{B}$ is tightly-t-bounded, which implies that $\mathcal{B}$ has finite tight rank $t' \leq t$. □

The rank one case

The main result of this section, Theorem 6.28, implies that for each rank one first-order reduct $\mathcal{B}$ of $(\mathbb{Z};<)$ without finite range endomorphisms there exists a first-order reduct $\mathcal{C}$ of $(\mathbb{Z};<)$ which has the same CSP as $\mathcal{B}$ and where succ is pp-definable, or for all $k \geq 1$ the relation Dist_k is pp-definable. By Theorem 6.4, it suffices to show that the endomorphisms of $\mathbb{Q}.\mathcal{C}$ preserve succ, or that the endomorphisms of $\mathbb{Q}.\mathcal{C}$ preserve Dist_k and Dist_k is 1-generated under $\mathrm{End}(\mathbb{Q}.\mathcal{C})$. The endomorphisms of $\mathcal{B}$ are better behaved than the endomorphisms of $\mathbb{Q}.\mathcal{B}$, as the latter endomorphisms can exhibit different behaviours in each copy of $\mathbb{Z}$, and can collapse copies, whereas the former endomorphisms are more uniform, as we will show below. Theorem 6.27 is the first milestone in our strategy, as it allows us to replace $\mathcal{B}$ with a first-order reduct $\mathcal{C}$ of $(\mathbb{Z};<)$ such that $\mathbb{Q}.\mathcal{C}$ has tight rank one.

Lemma 6.20. *Let $e\colon \mathbb{Z} \to \mathbb{Z}$ be tightly-t-bounded and $(1,c)$-bounded for some $c, t \in \mathbb{N}$. Then for all $n \in \mathbb{N}$, and $z \in \mathbb{Z}$, $|e(z+n) - e(z)| \leq n + ct$.*

Proof. Let $n = pt + k$ for $0 \leq k < t$. We have $|e(z+pt+k) - e(z+pt)| \leq kc$ by k applications of $(1,c)$-boundedness, and $|e(z+pt) - e(z)| \leq pt$ by p applications of tight rank t. We obtain

$$\begin{aligned} |e(z+n) - e(z)| &\leq |e(z+pt+k) - e(z+pt)| + |e(z+pt) - e(z)| \\ &\leq kc + pt = n + c(k-1) \leq n + ct \end{aligned}$$

by the triangle inequality. □

The following can be shown by the same proof as the proof of Lemma 6 in [15]; since our statement is more general, and since we use rank and tight rank instead of bounded distance degree, we still give the proof here for the convenience of the reader.

Lemma 6.21. *Let $e\colon \mathbb{Z} \to \mathbb{Z}$ be tightly-t-bounded and $(1,c)$-bounded. Then either $\{e\} \cup \mathrm{Aut}(\mathbb{Z};<)$ locally generates a function with finite range, or there exists $k > ct + 1$ such that for all $x, y \in \mathbb{Z}$ with $|x - y| = k$ we have $|e(x) - e(y)| \geq k$.*

Proof. Assume for all $k > ct + 1$ there are $x, y \in \mathbb{Z}$ with $|x - y| = k$ and $|e(x) - e(y)| < k$. We will prove that e locally generates a function with range of size at most $2ct + 1$. We again use an argument based on König's tree lemma, albeit with a different flavour than in the previous proofs. Enumerate $\mathbb{Z}$ as $a_1, a_2, \ldots$. The vertices of the tree on level n are the functions $h\colon \{a_1, \ldots, a_n\} \to \mathbb{Z}$ generated by $\{e\} \cup \mathrm{Aut}(\mathbb{Z}; <)$ such that the diameter of the image of h is bounded above by $2ct+1$ and such that $h(a_1) = 0$. The edges of the tree between the levels n and $n + 1$ are defined by function restriction. The condition on the diameter of the image of h implies that the tree is finitely branching, and we now prove that the tree is infinite.

Let $A \subseteq \mathbb{Z}$ be a finite set. Enumerate the pairs $(x, y) \in A^2$ with $x < y$ by $(x_1, y_1), \ldots, (x_r, y_r)$. Let m be the smallest number with the property that $\mathcal{F} := \{e\} \cup \mathrm{Aut}(\mathbb{Z}; <)$ generates a function h_1 such that $|h_1(x_1) - h_1(y_1)| = m$. We claim that $m \leq ct + 1$. Otherwise, by assumption there are $x, y \in \mathbb{Z}$ with $|x - y| = m$ and $|e(x) - e(y)| < m$. Let a be the automorphism of $(\mathbb{Z}; <)$ such that $a(\{h_1(x), h_1(y)\}) = \{x_1, y_1\}$. Then $\mathcal{F}$ also generates $h_1' := e \circ a \circ h_1$, but $|h_1'(x_1) - h_1'(y_1)| < m$ in contradiction to the choice of m. We conclude that $\mathcal{F}$ generates a function h_1 such that $|h_1(x_1) - h_1(y_1)| \leq ct + 1$.

Similarly, there exists h_2 generated by $\mathcal{F}$ such that $|h_2(h_1(x_2)) - h_2(h_1(y_2))| \leq ct + 1$. Continuing like this we arrive at a function h_r generated by $\mathcal{F}$ such that

$$|h_r h_{r-1} \cdots h_1(x_r) - h_r h_{r-1} \cdots h_1(y_r)| \leq ct + 1.$$

Now consider $h := h_r \circ \cdots \circ h_1$. Set $f_j := h_r \circ \cdots \circ h_{j+1}$ and $g_j := h_j \circ \cdots \circ h_1$, for all $1 \leq j \leq r$; so $h = f_j \circ g_j$. Then, since by construction $|g_j(x_j) - g_j(y_j)| \leq ct + 1$, we have that for all $j \in \mathbb{Z}$ with $1 \leq j \leq r$,

$$\begin{aligned} |h(x_j) - h(y_j)| &= |f_j(g_j(x_j)) - f_j(g_j(y_j))| \\ &\leq |g_j(x_j) - g_j(y_j)| + ct && \text{(by Lemma 6.20)} \\ &\leq 2ct + 1\,, \end{aligned}$$

and our claim follows. □

Definition 6.7. For $e\colon \kappa_1.\mathbb{Z} \to \kappa_2.\mathbb{Z}$, we call $s \in \mathbb{N}^+$ *stable for* e if for every $p \in \kappa_1$, one of the following applies:

- $e(z + s) = e(z) + s$ for all $z \in p.\mathbb{Z}$,
- $e(z + s) = e(z) - s$ for all $z \in p.\mathbb{Z}$.

Note that if a function e has a stable number, it does not generate a function with finite range. Indeed, it follows from the definition that for all $k \in \mathbb{Z}$ we have $|e(z+kt)-e(z)| = kt$.

Lemma 6.22. *Let $e\colon \mathbb{Z} \to \mathbb{Z}$ be tightly-t-bounded and 1-bounded. Then t is stable for e, or $\{e\} \cup \mathrm{Aut}(\mathbb{Z}; <)$ locally generates a function with finite range.*

Proof. Let $c \in \mathbb{N}$ be such that e is $(1, c)$-bounded, and assume that e does not locally generate a function with finite range. By Lemma 6.21, there exists $k > ct + 1$ such that for all z we have $|e(z + k) - e(z)| \geq k$, and hence either $e(z + k) \geq e(z) + k$ or $e(z + k) \leq e(z) - k$ for each $z \in \mathbb{Z}$. We will first show that either $e(z + k) \geq e(z) + k$ for all $z \in \mathbb{Z}$, or $e(z + k) \leq e(k) - k$ for all $z \in \mathbb{Z}$. Suppose otherwise that there are $z_1, z_2 \in \mathbb{Z}$

such that $e(z_1 + k) \geq e(z_1) + k$ and $e(z_2 + k) \leq e(z_2) - k$. Clearly, we can choose z_1, z_2 such that $|z_1 - z_2| = 1$. We only treat the case that $z_2 = z_1 + 1$, since the other case is symmetric. Then

$$\begin{aligned} e(z_2) - e(z_2 + k) &\geq k && \text{by assumption,} \\ -e(z_2) + e(z_1) &\geq -c && \text{by 1-boundedness,} \\ e(z_2 + k) - e(z_1 + k) &\geq -c && \text{by 1-boundedness,} \\ e(z_1 + k) - e(z_1) &\geq k && \text{by assumption.} \end{aligned}$$

Summing over those inequalities yields $0 \geq 2k - 2c$, a contradiction since $k > c$.

In the following we assume without loss of generality that $e(z + k) \geq e(z) + k$ for all $z \in \mathbb{Z}$. Recall that $|e(z + t) - e(z)| \leq t$ for all $z \in \mathbb{Z}$ because e is tightly-t-bounded. We next claim that $e(z + kt) = e(z) + kt$ for all $z \in \mathbb{Z}$. Since points at distance t cannot be mapped to points at larger distance, we get that $e(z+kt)-e(z) \leq kt$. On the other hand, since $e(z + k) \geq e(z) + k$ for all $z \in \mathbb{Z}$, we obtain that $e(z + kt) \geq e(z) + kt$, proving the claim.

We now show that $e(z + t) \geq e(z) + t$ for all $z \in \mathbb{Z}$. Note that

$$\begin{aligned} e(z) + kt &= e(z + kt) \\ &= e(z + t + (k - 1)t) \\ &\leq e(z + t) + (k - 1)t \end{aligned}$$

the latter inequality holding since $e(z + mt) - e(z) \leq mt$ for each $m \in \mathbb{N}$. Subtracting $(k - 1)t + e(z)$ on both sides, our claim follows. Since $|e(z + t) - e(z)| \leq t$ for all $z \in \mathbb{Z}$, we obtain that $e(z + t) = e(z) + t$ and have proved the lemma. □

Corollary 6.23. *Let $\mathcal{B}$ be a finite-signature reduct of $(\mathbb{Z}; <)$ without finite range endomorphism such that $\mathbb{Q}.\mathcal{B}$ has rank one. Then $\mathcal{B}$ has finite tight rank t and t is stable for every $e \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$.*

Proof. By Corollary 6.17, $\mathbb{Q}.\mathcal{B}$ has finite tight rank t', and by Lemma 6.19, $\mathcal{B}$ has tight rank $t \leq t'$ and rank one. Let $e \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$. Since $\mathbb{Q}.\mathcal{B}$ has rank one, we have $e(z+k)-e(z) < \infty$ for all $z \in \mathbb{Q}.\mathbb{Z}$ and $k \in \mathbb{Z}$. As a consequence, for any $p \in \mathbb{Q}$, the function e induces an endomorphism $e' \colon \mathcal{B} \to \mathcal{B}$ by restricting e to $p.\mathbb{Z}$. By Lemma 6.22, t is stable for e', and we conclude that t is stable for e. □

Lemma 6.24. *Every stable number of a function $e \colon \mathbb{Z} \to \mathbb{Z}$ is divisible by the smallest stable number of e.*

Proof. Suppose that p is stable but not divisible by s. Write $p = ms + r$ where m, r are positive integers and $0 < r < s$. Since r is not stable there exists $z \in \mathbb{Z}$ such that $e(z + r) - e(z) \neq r$. But this is impossible since

$$\begin{aligned} e(z + r) - e(z) &= e(z + p - ms) - e(z) \\ &= e(z - ms) + p - e(z) \\ &= e(z) - ms + p - e(z) = r \ . \end{aligned}$$

□

Lemma 6.25. *Let $\mathcal{B}$ be a finite-signature reduct of $(\mathbb{Z};<)$ without finite-range endomorphisms and such that $\mathbb{Q}.\mathcal{B}$ has rank one. Let e be an endomorphism of $\mathcal{B}$, and let s be the smallest stable number for e. Then $\{e\} \cup \mathrm{Aut}(\mathbb{Z};<)$ generates a function f such that $f(\mathbb{Z}) = \{s \cdot z : z \in \mathbb{Z}\}$.*

Proof. We prove by induction on $i \in \{0, \dots, s-1\}$ that there exists a function f_i, generated by $\{e\} \cup \mathrm{Aut}(\mathbb{Z};<)$, such that $f_i(j) \in \{s \cdot z : z \in \mathbb{Z}\}$ for all $j \in \{0, \dots, i\}$, and $f_i(0) = 0$. Without loss of generality, assume that $e(0) = 0$. The base case $i = 0$ is trivial: the identity function on $\mathbb{Z}$ satisfies the requirements. Let f_{i-1} be given. If $f_{i-1}(i)$ is a multiple of s there is nothing to do. Otherwise, $f_{i-1}(i)$ is not stable for e by Lemma 6.24. Since e has a stable number, it does not generate a function with finite range, so by Lemma 6.22 it is not tightly-$f_{i-1}(i)$-bounded. It follows that there exist $x_0, y_0 \in \mathbb{Z}$ with $x_0 - y_0 = f_{i-1}(i)$ and $|e(x_0) - e(y_0)| > |f_{i-1}(i)|$. Write $r_1 := |e(x_0) - e(y_0)|$. If r_1 is a multiple of s, then we are done: let α_0 be the automorphism of $(\mathbb{Z};<)$ that maps $\{0, f_{i-1}(i)\}$ to $\{x_0, y_0\}$, let β be the automorphism of $(\mathbb{Z};<)$ that maps $(e \circ \alpha_0 \circ f_{i-1})(0)$ to 0, and let $f_i = \beta \circ e \circ \alpha_0 \circ f_{i-1}$. Since s is stable for e and α_0, we have that $f_i(j) \in \{s \cdot z : z \in \mathbb{Z}\}$ for $j \in \{0, \dots, i-1\}$ and $|f_i(i)| = |e(y_0) - e(x_0)|$ is a multiple of s by hypothesis. Otherwise, using again Lemma 6.22 and Lemma 6.24, we know that e is not tightly-r_1-bounded. Therefore there exist $x_1, y_1 \in \mathbb{Z}$ with $|x_1 - y_1| = r_1$ and $|e(x_1) - e(y_1)| =: r_2 > r_1$. Continuing this way, we obtain a sequence of pairs $(x_0, y_0), (x_1, y_1), \dots$ such that $r_j = |x_j - y_j|$, and $r_{j+1} > r_j$. Up to exchanging x_j and y_j, we can assume that $e(x_j) < e(y_j)$ iff $x_{j+1} < y_{j+1}$. Since $\mathbb{Q}.\mathcal{B}$ has rank one, Lemma 6.18 gives a $c \geq 0$ such that every endomorphism of $\mathcal{B}$ is $(1, c)$-bounded. This implies that the sequence built above must stop in at most c steps. By construction, this can only happen when r_k is a multiple of s. For $j \in \{1, \dots, k-1\}$, set α_j an automorphism of $(\mathbb{Z};<)$ such that $\alpha_{j+1}(e(x_j)) = x_{j+1}$ and $\alpha_{j+1}(e(y_j)) = y_{j+1}$. Let β be the translation that maps x_k to 0. Finally, set $f_i := \beta \circ \alpha_k \circ e \circ \alpha_{k-1} \circ e \circ \cdots \circ \alpha_1 \circ e \circ \alpha_0 \circ f_{i-1}$. Since s is stable for e and automorphisms of $(\mathbb{Z};<)$, we have that $f_i(j)$ is a multiple of s for every $j \in \{0, \dots, i-1\}$. Finally we have $f_i(i) = y_k - x_k$ which is a multiple of s by construction. This finishes the inductive proof.

The function f whose existence is claimed in the statement is then f_{s-1}. Indeed, s is stable for f as f is obtained as the composition of e and automorphisms of $(\mathbb{Z};<)$. Therefore $f(\mathbb{Z})$ contains the set $\{s \cdot z : z \in \mathbb{Z}\}$. For the other inclusion, let $v \in \mathbb{Z}$ be arbitrary, and write $v = s \cdot z + r$, where $z \in \mathbb{Z}$ and $0 \leq r < s$. Then $f(s \cdot z + r) - f(r)$ is a multiple of s since s is stable for f. By construction, $f(r)$ is a multiple of s as well, so that $f(v) \in \{s \cdot z \mid z \in \mathbb{Z}\}$. □

The following definition arises naturally from the statement of Lemma 6.25.

Definition 6.8. Let $\mathcal{B}$ be a structure over $\mathbb{Z}$ and let $k \in \mathbb{N}^+$. Then we write $\mathcal{B}/k$ for the substructure of $\mathcal{B}$ induced by the set $\{z \in \mathbb{Z} : z = 0 \bmod k\}$.

Lemma 6.26. *For all first-order reducts $\mathcal{B}$ of $(\mathbb{Z};<)$ and $k \in \mathbb{N}^+$, the structure $\mathcal{B}/k$ is isomorphic to a first-order reduct of $(\mathbb{Z};<)$, the isomorphism being the function $x \mapsto x/k$.*

Proof. Let R be an n-ary relation of $\mathcal{B}$, and let ϕ be a quantifier-free formula defining R. Construct a formula ϕ' as follows: For all $i \in \mathbb{Z}$, replace every atomic formula of the form $x \leq y + i$ by $x \leq y + \lfloor i/k \rfloor$. We prove by structural induction on ϕ that for all $z_1, \dots, z_n \in \mathcal{B}/k$ we have $(\mathbb{Z},<) \models \phi(z_1, \dots, z_n) \Leftrightarrow (\mathbb{Z};<) \models \phi'(z_1/k, \dots, z_n/k)$. If

ϕ is $x \leq y + i$ for some $i \in \mathbb{Z}$, then $\mathcal{B}/k \models \phi(x, y)$ iff $x \leq y + i$ iff $x/k \leq y/k + \lfloor i/k \rfloor$. The cases of conjunction, disjunction, and negation follow immediately from the induction hypothesis. $\square$

For instance, in Example 15 the structure $\mathcal{B}/2$ is isomorphic to $(\mathbb{Z}; \mathrm{succ}, \{(x, y) : |x - y| \leq 1\})$ which has tight rank one.

Theorem 6.27. *Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z}; <)$ without finite range endomorphisms and such that $\mathbb{Q}.\mathcal{B}$ has rank one. Then $\mathcal{B}$ has an endomorphism that maps $\mathcal{B}$ to $\mathcal{B}/k$ for some $k \in \mathbb{N}^+$, which is isomorphic to a reduct $\mathcal{C}$ of $(\mathbb{Z}; <)$ such that $\mathbb{Q}.\mathcal{C}$ has tight rank one.*

Proof. Let t be the tight rank of $\mathbb{Q}.\mathcal{B}$, and let c be such that $\mathbb{Q}.\mathcal{B}$ is $(1, c)$-bounded (which exists by Corollary 6.17). By Lemma 6.19, $\mathcal{B}$ has tight rank t', with $t' \leq t$. By Corollary 6.23, every endomorphism of $\mathbb{Q}.\mathcal{B}$ has a stable number, and in particular each endomorphism has a minimal one. If the minimal stable number of every endomorphism is 1, then $\mathbb{Q}.\mathcal{B}$ has tight rank one and we are done, choosing $k = 1$. Otherwise there exists an $e \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$ such that 1 is not stable for e. So there exists a copy of $\mathbb{Z}$ and some integer $s > 1$ such that s is stable for the restriction of e to that copy, which we call $\hat{e}$, and so that no s' with $s' < s$ is stable for $\hat{e}$. Since $\mathbb{Q}.\mathcal{B}$ has rank one, e sends copies of $\mathbb{Z}$ to copies of $\mathbb{Z}$. By composing $\hat{e}$ with an automorphism of $(\mathbb{Q}.\mathbb{Z}; <)$ we can assume that $\hat{e} \in \mathrm{End}(\mathcal{B})$. By Lemma 6.25, there exists a function f generated by $\{\hat{e}\} \cup \mathrm{Aut}(\mathbb{Z}; <)$ such that $f(\mathbb{Z}) = \{s \cdot z \mid z \in \mathbb{Z}\}$. By Lemma 6.22, t' is stable for f, and t' is divisible by s since $|f(z + t') - f(z)| = t'$ and $f(z + t'), f(z) \in \{s \cdot z \mid z \in \mathbb{Z}\}$. Also note that s is stable for f since f is generated by $\hat{e}$.

Observe that $\mathcal{B}/s$ cannot have a finite range endomorphism: if g were such an endomorphism, then $g \circ f$ would be a finite range endomorphism for $\mathcal{B}$, contrary to our assumption. By Lemma 6.26, $\mathcal{B}/s$ is isomorphic to a first-order reduct $\mathcal{C}$ of $(\mathbb{Z}; <)$ via the function $x \mapsto x/s$. It is also clear that the function $(a, z) \mapsto (a, sz)$ from $\mathbb{Q}.\mathbb{Z}$ to $\mathbb{Q}.\mathbb{Z}$ is a homomorphism from $\mathbb{Q}.\mathcal{C}$ to $\mathbb{Q}.\mathcal{B}$. We claim that $\mathbb{Q}.\mathcal{C}$ has rank one and tight rank at most t'/s.

Let $e \in \mathrm{End}(\mathbb{Q}.\mathcal{C})$. Let $x \in \mathbb{Z} \subset \mathbb{Q}.\mathbb{Z}$. Define $e'(z) = s \cdot e(f(z)/s)$ which is a homomorphism $\mathcal{B} \to \mathbb{Q}.\mathcal{B}$. Note that every homomorphism from $\mathcal{B}$ to $\mathbb{Q}.\mathcal{B}$ is $(1, c)$-bounded, since otherwise by Lemma 6.15 we can find an endomorphism of $\mathbb{Q}.\mathcal{B}$ which is not $(1, c)$-bounded. Since f is surjective as a function $\mathbb{Z} \to \{s \cdot z \mid z \in \mathbb{Z}\}$, there exists $y \in \mathbb{Z}$ such that $f(y)/s = x$. Since s is stable for f, we have either $f(y + s) = f(y) + s$ or $f(y - s) = f(y) + s$. If $f(y + s) = s + f(y)$, then $e(\frac{f(y)+s}{s}) = \frac{1}{s} \cdot e'(y + s)$. In the other case, $e(\frac{f(y)+s}{s}) = \frac{1}{s} \cdot e'(y - s)$. In any case, by applying s times the $(1, c)$-boundedness of e', we obtain that

$$|e(x + 1) - e(x)| = \left| e\left(\frac{f(y)}{s} + 1\right) - e\left(\frac{f(y)}{s}\right) \right| \leq c.$$

The same argument works for all $x \in \mathbb{Q}.\mathbb{Z}$, so all the endomorphisms of $\mathbb{Q}.\mathcal{C}$ are $(1, c)$-bounded and $\mathbb{Q}.\mathcal{C}$ has rank one. Similarly, we have

$$\left| e\left(x + \frac{t'}{s}\right) - e(x) \right| = \left| e\left(\frac{f(y)}{s} + \frac{t'}{s}\right) - e\left(\frac{f(y)}{s}\right) \right| \leq \frac{t'}{s}$$

i.e., e is tightly-t'/s-bounded and $\mathbb{Q}.\mathcal{C}$ has tight rank at most t'/s.

Since $\mathcal{C}$ satisfies all the assumptions that we had on $\mathcal{B}$, we may repeat the argument. If $\mathbb{Q}.\mathcal{C}$ has tight rank 1, then we are done. This process terminates, since the tight rank of $\mathbb{Q}.\mathcal{C}$ is bounded above by t'/s, which is strictly smaller than the tight rank of $\mathbb{Q}.\mathcal{B}$. Observe furthermore that if $\mathcal{C}'$ is the first-order reduct of $(\mathbb{Z};<)$ that is isomorphic to $\mathcal{C}/s'$, then $\mathcal{C}'$ is isomorphic to $\mathcal{B}/ss'$ by the obvious composition of isomorphisms, so that the resulting structure at termination is indeed of the form $\mathcal{B}/k$ for some $k \in \mathbb{N}$. □

Theorem 6.28. *Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z};<)$ such that $\mathbb{Q}.\mathcal{B}$ has rank one. Then* $\mathrm{CSP}(\mathcal{B})$ *equals* $\mathrm{CSP}(\mathcal{C})$ *where $\mathcal{C}$ is one of the following:*

1. *a finite structure;*
2. *a first-order reduct of $(\mathbb{Z};<)$ where* Dist_k *is pp-definable for all $k \geq 1$;*
3. *a first-order reduct of $(\mathbb{Z};<)$ where* succ *is pp-definable.*

Proof. If $\mathcal{B}$ has a finite-range endomorphism f, then the image of the endomorphism induces a finite structure with the same CSP as $\mathcal{B}$, thus we are in case (1) of the statement and done. So assume that this is not the case. Then by Theorem 6.27, $\mathcal{B}$ has an endomorphism g that maps $\mathcal{B}$ to $\mathcal{B}/k$ which is isomorphic to a reduct $\mathcal{C}$ of $(\mathbb{Z};<)$ such that $\mathbb{Q}.\mathcal{C}$ has tight rank one. Lemma 6.19 implies that $\mathcal{C}$ has tight rank one, too. The structure $\mathcal{B}/k$ cannot have finite-range endomorphisms f since otherwise $f \circ g$ would be a finite-range endomorphism for $\mathcal{B}$. Hence, $\mathcal{C}$ does not have finite-range endomorphisms. Since $\mathbb{Q}.\mathcal{C}$ has rank one, Corollary 6.23 is applicable, and implies that 1 is stable for every endomorphism of $\mathbb{Q}.\mathcal{C}$. Hence all endomorphisms of $\mathbb{Q}.\mathcal{C}$ are isometries and the relation Dist_k is preserved by the endomorphisms of $\mathbb{Q}.\mathcal{C}$.

If succ is preserved by all the endomorphisms of $\mathbb{Q}.\mathcal{C}$, then Theorem 6.4 implies that succ is pp-definable in $\mathbb{Q}.\mathcal{C}$ since succ is 1-generated under $\mathrm{End}(\mathbb{Q}.\mathcal{C})$. In this case, succ is pp-definable in $\mathcal{C}$, too, and we are in case (3) of the statement.

Otherwise, there exists an endomorphism e of $\mathbb{Q}.\mathcal{C}$ that does not preserve succ. Therefore, there exists an $x \in \mathbb{Q}.\mathbb{Z}$ such that $e(x+k) = e(x)-k$ for all $k \geq 1$. For each $k \geq 1$, the relation Dist_k is then 1-generated under $\mathrm{End}(\mathbb{Q}.\mathcal{C})$, the pair $(x, x+k)$ being a generator. Since Dist_k is preserved by all endomorphisms of $\mathbb{Q}.\mathcal{C}$, it follows from Theorem 6.4 that Dist_k is pp-definable in $\mathbb{Q}.\mathcal{C}$ for all $k \geq 1$. Finally, this implies that Dist_k is pp-definable in $\mathcal{C}$ for all $k \geq 1$ and we are in case (2) of the statement. □

Arbitrary rank

In this section we study first-order reducts of $(\mathbb{Z};<)$ of arbitrary finite rank. The goal is to reduce this to the rank one situation (in Proposition 6.32). For this, we need the following proposition, which is quite similar, but formally unrelated, to the implication from item (2) to item (4) in Theorem 6.16.

Lemma 6.29. *Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z};<)$ and $k \in \mathbb{N}$ such that $\mathbb{Q}.\mathcal{B}$ is not k-bounded. Then for all $x, y \in \mathbb{Z}$ such that $x - y = k$ there exists an endomorphism h of $\mathbb{Q}.\mathcal{B}$ such that $h(x) - h(y) = \infty$ and for all z, z' with $z - z' = \infty$ we have $h(z) - h(z') = \infty$.*

Proof. Since $\mathbb{Q}.\mathcal{B}$ is not k-bounded, for any $r \geq 0$ there exist $x_0, y_0 \in \mathbb{Q}.\mathbb{Z}$ with $|x_0 - y_0| = k$ and an endomorphism $e\colon \mathbb{Q}.\mathcal{B} \to \mathbb{Q}.\mathcal{B}$ such that $e(x_0) - e(y_0) > r$. Composing e with an automorphism we can take $\{x_0, y_0\} = \{x, y\}$. For every finite set $S \subset \mathbb{Z}$, we then have a homomorphism $e\colon \mathbb{Q}.\mathcal{B}[S] \to \mathbb{Q}.\mathcal{B}$ such that $e(x) - e(y) > r$. It follows from an analog of Proposition 6.10 that there exists a homomorphism $h\colon \mathbb{Q}.\mathcal{B} \to \mathbb{Q}.\mathcal{B}$ such that $h(x) - h(y) = \infty$ and for all z, z' with $z - z' = \infty$ we have $h(z) - h(z') = \infty$. □

Proposition 6.30. *Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z}; <)$ such that $\mathbb{Q}.\mathcal{B}$ has rank r, and let e be an endomorphism of $\mathbb{Q}.\mathcal{B}$. Then $e(z_1) = e(z_2) \mod r$ for all $z_1, z_2 \in \mathbb{Q}.\mathbb{Z}$ such that $z_1 = z_2 \mod r$.*

Proof. Suppose that $e \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$ is (r, c)-bounded and $z_1, z_2 \in \mathbb{Q}.\mathbb{Z}$ contradict the statement of the proposition. Choose z_1, z_2 such that $z_1 > z_2$ and $z_1 - z_2$ is minimal.

Claim 1. $z_1 - z_2 = r$.

Suppose otherwise; then there are $p_1, \ldots, p_k$ for $k > 2$ such that $p_1 = z_1$, $p_k = z_2$, and $p_i - p_{i+1} = r$ for all $i \in \{1, \ldots, k-1\}$ because r divides $z_1 - z_2$. By the choice of z_1, z_2 we have $e(p_i) = e(p_j) \bmod r$. But then $e(p_1) = e(p_k) \bmod r$, a contradiction to the assumption that $e(z_1) \neq e(z_2) \bmod r$.

Let $w, v \in \mathbb{N}$ be such that $|e(z_1) - e(z_2)| = wr + v$ and $v < r$. Note that $v > 0$ because $e(z_1) \neq e(z_2) \bmod r$. Assume that $e(z_1) > e(z_2)$; the proof when $e(z_2) > e(z_1)$ is analogous. Let $e' \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$ be arbitrary, and $u_1, u_2 \in \mathbb{Z}$ be arbitrary such that $u_1 - u_2 = v$.

Claim 2. $|e'(u_1) - e'(u_2)| \leq (w+1)c + 1$.

To prove the claim, suppose the contrary. Let $\alpha \in \mathrm{Aut}(\mathbb{Z}; <)$ be such that $\alpha(e(z_1)) = u_1$. Note that $\alpha(e(z_2) + wr) = u_2$. Set $e'' := e' \circ \alpha \circ e$. Then

$$\begin{aligned}
|e''(z_1) - e''(z_2))| &\geq |e''(z_1) - e'(u_2)| - |e'(u_2) - e''(z_2))| \\
&= |e'(u_1) - e'(u_2)| - |e'(\alpha(e(z_2) + wr)) - e'(\alpha(e(z_2)))| \\
&\geq (w+1)c + 1 - wc = c + 1
\end{aligned}$$

where the first inequality is the triangle inequality, and the second inequality is by assumption and (r, c)-boundedness. But $|e''(z_1)) - e''(z_2))| > c$ contradicts the assumption that $\mathbb{Q}.\mathcal{B}$ is (r, c)-bounded, and this finishes the proof of Claim 2.

Since e' was chosen arbitrarily, we obtain that $\mathbb{Q}.\mathcal{B}$ is $(v, w(c+1)+1)$-bounded, and hence has rank $v < r$, a contradiction. □

Lemma 6.31. *Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z}; <)$ such that $\mathbb{Q}.\mathcal{B}$ has rank $r \in \mathbb{N}$. Then there exists an endomorphism e of $\mathbb{Q}.\mathcal{B}$ with the property that for all $x, y \in \mathbb{Q}.\mathbb{Z}$,*

$$\begin{aligned}
&\textit{either } e(y) - e(x) = \infty \\
&\quad \textit{or } e(y) - e(x) = 0 \mod r.
\end{aligned}$$

Proof. We construct e by an application of König's tree lemma as follows. Let $a_1, a_2, \ldots$ be an enumeration of the elements of $\mathbb{Q}.\mathbb{Z}$. Given a partial function $f\colon \{a_1, \ldots, a_n\} \to \mathbb{Q}.\mathbb{Z}$, we say that f has property (†) if for all $x, y \in \{a_1, \ldots, a_n\}$, either $f(y) - f(x) = \infty$ or $f(x) = f(y) \mod r$. The vertices on level n of the tree are $\sim_q$-equivalence classes of

homomorphisms h from $\mathbb{Q}.\mathcal{B}[\{a_1, \ldots, a_n\}]$ to $\mathbb{Q}.\mathcal{B}$ that satisfy property (†). Adjacency between vertices is defined by restriction of representatives.

The interesting part of the proof is to show that the tree has vertices on all levels. Let g be a homomorphism from $\mathbb{Q}.\mathcal{B}[\{a_1, \ldots, a_n\}]$ to $\mathbb{Q}.\mathcal{B}$ such that the number m of pairs $i, j \in \{1, \ldots, n\}$ with $g(a_i) - g(a_j) = \infty$ or $g(a_i) = g(a_j) \mod r$ is maximal. If $m = \binom{n}{2}$ then we are done; so suppose that there are $p, q \in \{1, \ldots, n\}$ such that $g(a_p) - g(a_q) \in \mathbb{Z}$ is not divisible by r. Let $k \in \{1, \ldots, r-1\}$ and $l \in \mathbb{Z}$ be such that $g(a_p) - g(a_q) = lr + k$, $0 < k < r$. Since $\mathbb{Q}.\mathcal{B}$ is not k-bounded, by Lemma 6.29 there exists an endomorphism f of $\mathbb{Q}.\mathcal{B}$ such that $f(g(a_p) + lr) - f(g(a_q)) = \infty$. By Proposition 6.30 we have $f(g(a_p)) = f(g(a_p) + lr) \mod r$, and hence $f(g(a_p)) - f(g(a_q)) = \infty$. We claim that the number m' of pairs $i, j \in \{1, \ldots, n\}$ such that $f(g(a_i)) - f(g(a_j)) = \infty$ or $f(g(a_i)) = f(g(a_j)) \mod r$ is larger than m. If $g(a_i) - g(a_j) = \infty$ then $f(g(a_i)) - f(g(a_j)) = \infty$; if $g(a_i) = g(a_j) \mod r$ then $f(g(a_i)) = f(g(a_j)) \mod r$. Therefore, $m' \geq m$. Moreover, we have $f(g(a_p)) - f(g(a_q)) = \infty$, and hence $m' > m$. Then $f \circ g$ is a homomorphism from $\mathbb{Q}.\mathcal{B}[a_1, \ldots, a_n] \to \mathbb{Q}.\mathcal{B}$, contradicting the maximality of m.

By Lemma 6.9, we obtain an endomorphism $e \colon \mathbb{Q}.\mathbb{Z} \to \mathbb{Q}.\mathbb{Z}$ such that for every n, $e|_{\{a_1,\ldots,a_n\}}$ is $\sim_q$-equivalent to some function g_n satisfying (†). Let $x, y \in \mathbb{Q}.\mathbb{Z}$. If x, y are (e, q)-connected, then they are $(e|_{\{a_1,\ldots,a_n\}}, q)$-connected for some n, so that they are (g_n, q)-connected. It follows that $e(x) - e(y) = g_n(x) - g_n(y) = 0 \bmod r$. If x, y are not (e, q)-connected, they are not (g_n, q)-connected for any function g_n in the tree, and we have $e(x) - e(y) = \infty$ by Lemma 6.9. Therefore, e satisfies (†). □

Proposition 6.32. *Let $\mathcal{B}$ be a finite-signature reduct of $(\mathbb{Z}; <)$ without finite-range endomorphism and such that $\mathbb{Q}.\mathcal{B}$ has rank $r \in \mathbb{N}$ and tight rank $t \in \mathbb{N}$. Then $\mathcal{B}/r$ has the same CSP as $\mathcal{B}$, and is isomorphic to a first-order reduct $\mathcal{C}$ of $(\mathbb{Z}; <)$ such that $\mathbb{Q}.\mathcal{C}$ has tight rank at most t/r.*

Proof. By Lemma 6.26, there is a first-order reduct $\mathcal{C}$ of $(\mathbb{Z}; <)$ such that $x \mapsto r \cdot x$ is an isomorphism between $\mathcal{C}$ and $\mathcal{B}/r$. Let e be the endomorphism of $\mathbb{Q}.\mathcal{B}$ constructed in Lemma 6.31. Replacing e by $\alpha \circ e$ for an appropriate automorphism α of $(\mathbb{Q}.\mathbb{Z}; <)$, we can assume that the range of e lies within $S := \{r \cdot z : z \in \mathbb{Q}.\mathbb{Z}\}$. Since $x \mapsto r \cdot x$ is an isomorphism between $\mathbb{Q}.\mathcal{C}$ and the structure induced by S in $\mathbb{Q}.\mathcal{B}$, we obtain that $\mathcal{B}$, $\mathbb{Q}.\mathcal{B}$, $\mathbb{Q}.\mathcal{C}$, and $\mathcal{C}$ all have the same CSP.

It remains to be shown that $\mathbb{Q}.\mathcal{C}$ has rank at most t/r. For an arbitrary $e \in \mathrm{End}(\mathbb{Q}.\mathcal{B})$, the quantity $\delta(e) := \max_{z \in \mathbb{Q}.\mathbb{Z}} |e(z + t) - e(z)|$ is well-defined and finite, since $\mathbb{Q}.\mathcal{B}$ has tight-rank t. Let e be an endomorphism of $\mathbb{Q}.\mathcal{B}$ as in Lemma 6.31 such that $\delta(e)$ is maximal among all endomorphisms satisfying the conclusion of Lemma 6.31. Let $z_0 \in \mathbb{Q}.\mathcal{B}$ be a witness for the maximum taken in $\delta(e)$. If $e(z_0 + t) = e(z_0)$, then for all $z \in \mathbb{Q}.\mathbb{Z}$ we have $e(z + t) = e(z)$. As in the proof of (1) $\Rightarrow$ (2) in Theorem 6.16, this would imply that $\mathcal{B}$ has a finite-range endomorphism, a contradiction to the assumption. So we have $e(z_0 + t) \neq e(z_0)$. Suppose that $e(z_0 + t) > e(z_0)$, the other case being treated similarly. Since e satisfies the property of Lemma 6.31, the distance $e(z_0 + t) - e(z_0)$ is equal to kr for some $k \leq t/r$. We prove that $\mathbb{Q}.\mathcal{C}$ has tight rank k. Let f be an endomorphism of $\mathbb{Q}.\mathcal{C}$, and suppose that there exists a $y \in \mathbb{Q}.\mathcal{C}$ such that $|f(y + k) - f(y)| > k$. Up to composition of f with an automorphism of $\mathbb{Q}.\mathcal{C}$, we can assume that $y = \frac{e(z_0)}{r}$. Note that $y + k = \frac{e(z_0)}{r} + k = \frac{e(z_0)+kr}{r} = \frac{e(z_0+t)}{r}$. Let $e' \colon \mathbb{Q}.\mathcal{B} \to \mathbb{Q}.\mathcal{B}$ be defined by $e'(x) = r \cdot f(\frac{e(x)}{r})$. Note that e' satisfies the property of Lemma 6.31. We have furthermore

$|e'(z_0+t)-e'(z_0)| = r \cdot |f(y+k)-f(y)| > kr$. This contradicts the fact that e was chosen to maximise the distance $|e(z_0+t)-e(z_0)|$. □

Iterating the previous proposition, we finally obtain a reduction to the rank one case.

Corollary 6.33. *Let $\mathcal{B}$ be a finite-signature reduct of $(\mathbb{Z};<)$ such that $\mathbb{Q}.\mathcal{B}$ has rank $r \in \mathbb{N}$. Then there exists a $k \in \mathbb{N}$ such that $\mathcal{B}/k$ has the same CSP as $\mathcal{B}$, and is isomorphic to a reduct $\mathcal{C}$ of $(\mathbb{Z};<)$ such that $\mathbb{Q}.\mathcal{C}$ has rank one.*

Proof. If $\mathbb{Q}.\mathcal{B}$ has rank one there is nothing to prove, so assume that $r > 1$. By Proposition 6.32, $\mathcal{B}/r$ has the same CSP as $\mathcal{B}$, is isomorphic to a reduct $\mathcal{C}_1$ of $(\mathbb{Z};<)$, and the tight rank t_1 of $\mathbb{Q}.\mathcal{C}_1$ is strictly smaller than that of $\mathbb{Q}.\mathcal{B}$. Write $\mathcal{C}_0 := \mathcal{B}$. We iterate this construction, obtaining reducts $\mathcal{C}_0, \mathcal{C}_1, \ldots, \mathcal{C}_{n+1}$ of $(\mathbb{Z};<)$ with ranks $r_0, r_1, \ldots, r_{n+1}$ and tight ranks $t_0 > t_1 > \cdots > t_n = t_{n+1}$ until the sequence of tight ranks stabilises, which can only happen if the rank of $\mathbb{Q}.\mathcal{C}_n$ is one. The structure $\mathcal{C}_n$ is isomorphic to $\mathcal{B}/(r_0 \ldots r_{n-1})$, which proves the corollary. □

6.4.6 Defining succ and $<$

In the remainder of this section, we prove the following dichotomy: a first-order reduct of $(\mathbb{Z};<)$ that pp-defines succ either pp-defines $<$, or is a first-order reduct of $(\mathbb{Z};\mathrm{succ})$. Call a binary relation R with a first-order definition over $(\mathbb{Z};<)$ *one-sided infinite* if there exist $c, d \in \mathbb{Z}$ with $c \leq d$ so that

- $R(x, x+z)$ holds for no $z < c$,
- $R(x, x+z)$ holds for all $z \geq d$.

Note that this definition does not depend on $x \in \mathbb{Z}$, since R is first-order definable over $(\mathbb{Z};<)$.

Lemma 6.34. *Let $\mathcal{B}$ be a first-order reduct of $(\mathbb{Q}.\mathbb{Z};<)$ such that* succ *is pp-definable in $\mathcal{B}$. Then $<$ is pp-definable in $\mathcal{B}$ if and only if some one-sided infinite binary relation is pp-definable in $\mathcal{B}$.*

Proof. Since $<$ is one-sided infinite we only have to show the reverse implication. Choose a binary one-sided infinite relation R with a pp-definition in $\mathcal{B}$ such $d-c$ is minimal, where c and d are as in the definition of one-sided infinity of R. If $c = d$ then R is a relation of the form $x < y + k$ for $k \in \mathbb{Z}$, and using succ we can pp-define $<$ in $\mathcal{B}$. We now show that $c \neq d$ is impossible. Replace R by the relation T defined by the formula $R(x,y) \wedge R(x, y+d-c-1)$, which is equivalent to a pp-formula over $\mathcal{B}$. Then $(0,z)$ is in T for all $z \geq d$. On the other hand, for $z < c+1$, we have $(0,z) \notin T$. Indeed, if $z < c$ then $(0,z)$ is not in R, so not in T. If $z = c$, then $(0, d-1)$ is not in R by the minimality of d, so that $(0,c)$ is not in T. Therefore, the integers c', d' as defined for T in place of R have a smaller difference than $d-c$, contradicting the choice of R such that $d-c$ is minimal. □

If R is a relation of arity n, and $i_1, \ldots, i_k \in \{1, \ldots, n\}$ are distinct indices, the projection of R onto $\{i_1, \ldots, i_k\}$, denoted by $\pi_{i_1,\ldots,i_k}(R)$, is the relation defined by

$$\exists x_{j_1}, \ldots, x_{j_{n-k}}.R(x_1, \ldots, x_n)$$

over $(\mathbb{Z};R)$ where $\{j_1, \ldots, j_{n-k}\} = \{1, \ldots, n\} \setminus \{i_1, \ldots, i_k\}$. A *binary projection* of R is a projection of R onto a set of size 2.

Lemma 6.35. *Let $\mathcal{B}$ be a first-order reduct of $(\mathbb{Z};<)$ in which* succ *is pp-definable. Then, either $\mathcal{B}$ pp-defines $<$ or $\mathcal{B}$ is a first-order reduct of $(\mathbb{Z};\mathrm{succ})$.*

Proof. Let R be a relation of $\mathcal{B}$ of arity k. If $E = \{a - b \mid (a,b) \in \pi_{i,j}(R)\}$, for distinct $i,j \in \{1,\dots,k\}$, is a finite or cofinite set, there is a definition of $R(x_1,\dots,x_k)$ over $(\mathbb{Z};<)$ without literals of the form $x_i < x_j + k$. Indeed, such a literal can be replaced by a disjunction of literals $\mathrm{succ}^p(x_i,x_j)$ for suitable integers p if E is finite, or a by a conjunction of literals $\neg\mathrm{succ}^p(x_i,x_j)$ if E is cofinite. Therefore, if $\mathcal{B}$ is not a first-order reduct of $(\mathbb{Z};\mathrm{succ})$ there exists a relation R of $\mathcal{B}$ and integers i,j such that the set $\{a-b \mid (a,b) \in R'\}$ for $R' := \pi_{i,j}(R)$ is neither finite nor cofinite. Let q be the quantifier-elimination degree of R'. It is clear that if $(a,b) \in R'$ with $a - b > q$, then $(a',b') \in R'$ whenever $a' - b' > q$. It follows that R' or $\{(b,a) \in \mathbb{Z}^2 \mid (a,b) \in R'\}$ is one-sided infinite. From Lemma 6.34 and the fact that R' is pp-definable in $\mathcal{B}$ follows that $<$ is pp-definable in $\mathcal{B}$. □

Combining the results of the preceding subsections, we can finally prove Theorem 6.6, which we restate here for the convenience of the reader.

Theorem 6.6. Let $\mathcal{B}$ be a first-order reduct of $(\mathbb{Z};<)$ with finite signature. Then $\mathrm{CSP}(\mathcal{B})$ equals $\mathrm{CSP}(\mathcal{C})$ where $\mathcal{C}$ is one of the following:

1. a finite structure;
2. a first-order reduct of $(\mathbb{Q};<)$;
3. a first-order reduct of $(\mathbb{Z};<)$ where Dist_k is pp-definable for all $k \geq 1$;
4. a first-order reduct of $(\mathbb{Z};<)$ where succ and $<$ are pp-definable;
5. a first-order reduct of $(\mathbb{Z};\mathrm{succ})$ where succ is pp-definable.

Proof. Let $\mathcal{B}$ be a first-order reduct of $(\mathbb{Z};<)$ with finite signature. If $\mathcal{B}$ has an endomorphism with finite range, then $\mathcal{B}$ is homomorphically equivalent to a finite structure; hence item (1) of Theorem 6.6 holds and we are done. So suppose that this is not the case. If there exists a first-order reduct of $(\mathbb{Q};<)$ with the same CSP, then item (2) of Theorem 6.6 holds and we are done. Otherwise, the equivalence of (2) and (1) in Theorem 6.16 implies that $\mathbb{Q}.\mathcal{B}$ has bounded tight rank t and bounded rank r. If $r > 1$, then by Proposition 6.32 we have that $\mathcal{B}$ has the same CSP as a first-order reduct $\mathcal{C}$ of $(\mathbb{Z};<)$ such that $\mathbb{Q}.\mathcal{C}$ has rank 1. It follows from Theorem 6.28 that there exists a first-order reduct $\mathcal{C}'$ of $(\mathbb{Z};<)$ that has the same CSP as $\mathcal{B}$ and such that Dist_k is pp-definable in $\mathcal{C}'$ for all $k \geq 1$ or succ is pp-definable in $\mathcal{C}'$. In the former case, item (3) of Theorem 6.6 holds. In the latter case, we finally have by Lemma 6.35 that $<$ has a pp-definition in $\mathcal{C}'$, in which case item (4) holds, or that $\mathcal{C}'$ is a first-order reduct of $(\mathbb{Z};\mathrm{succ})$, in which case item (5) holds. □

6.5 Tractable Classes

We treat the algorithmic part of our main result, that is, we prove that if $\mathcal{B}$ is a first-order reduct of $(\mathbb{Z};<)$ that is preserved by $\max_d$ or $\min_d$, or if $\mathcal{B}$ is a first-order reduct of $(\mathbb{Z};\mathrm{succ})$ such that $\mathbb{Q}.\mathcal{B}$ is preserved by a binary injective operation preserving succ, then $\mathrm{CSP}(\mathcal{B})$ is in P(items (3) and (4) in Theorem 6.5).

6.5.1 The Horn case

The two structures $(\mathbb{Q}.\mathbb{Z}, \mathrm{succ})^2$ and $(\mathbb{Q}.\mathbb{Z}, \mathrm{succ})$ are isomorphic. Let si be an isomorphism from $(\mathbb{Q}.\mathbb{Z}, \mathrm{succ})^2$ to $(\mathbb{Q}.\mathbb{Z}, \mathrm{succ})$. In the following we will also consider si as a binary operation on $\mathbb{Q}.\mathbb{Z}$ that preserves succ. Remember that relations that are first-order definable over $(\mathbb{Z}; \mathrm{succ})$ are also definable by quantifier-free formulas with (positive or negative) literals of the form $\mathrm{succ}^p(x, y)$ for $p \in \mathbb{Z}$ (see items (1)-(5) in Example 11). A quantifier-free formula in conjunctive normal form (CNF) is called *Horn* if each clause of the formula contains at most one positive literal. A relation is said to be *Horn-definable* if there exists a Horn formula that defines the relation.

We use the following characterisation of Horn definability, which is Proposition 5.9 in [17]: if $\mathcal{C}$ is a structure with an embedding e of $\mathcal{C}^2$ into $\mathcal{C}$ (such as for instance $\mathcal{C} = (\mathbb{Q}.\mathbb{Z}; \mathrm{succ})$) then a relation R with a quantifier-free definition in $\mathcal{C}$ is Horn-definable over $\mathcal{C}$ if and only if R is preserved by e. Applied to our situation, we obtain the following.

Proposition 6.36. *Let $\mathcal{B}$ be a first-order reduct of* $(\mathbb{Z}; \mathrm{succ})$. *Then the following are equivalent.*

- *every relation of $\mathcal{B}$ is Horn-definable over* $(\mathbb{Z}; \mathrm{succ})$*;*
- $\mathbb{Q}.\mathcal{B}$ *is preserved by si;*
- $\mathbb{Q}.\mathcal{B}$ *has a binary injective polymorphism that preserves* succ.

Proposition 6.37. *Let $\mathcal{B}$ be a finite-signature first-order reduct of* $(\mathbb{Z}; \mathrm{succ})$ *such that si is a polymorphism of $\mathbb{Q}.\mathcal{B}$. Then* $\mathrm{CSP}(\mathcal{B})$ *is in* P.

Proof. From Proposition 6.36 we know that the relations of $\mathcal{B}$ are definable with quantifier-free Horn formulas over $(\mathbb{Z}; \mathrm{succ})$. It is easy to see that there is a polynomial-time algorithm that decides whether a set of constraints of the form $\mathrm{succ}^{p_i}(x_i, y_i)$ is satisfiable. Moreover, we can also efficiently decide whether it implies another constraint of this form. Indeed, to see if the set of constraints is satisfiable, consider the graph whose vertices are the variables, and whose arcs consists of those pairs (x_i, y_i), labelled by p_i, such that there is a constraint $\mathrm{succ}^{p_i}(x_i, y_i)$ in the input. For each variable x, using a graph traversal we can check whether all the directed paths going from x to some other variable y have the same weight (which is given by the sum of the labels over the arcs); If this is not the case, the constraints are unsatisfiable. Otherwise, to decide whether the constraints imply $\mathrm{succ}^p(x, y)$, check whether there is a directed path from x to y where the sum of the labels equals p.

We view the instance of $\mathrm{CSP}(\mathcal{B})$ as a set of Horn-clauses over $(\mathbb{Z}; \mathrm{succ})$. We iterate the following algorithm: form the set U of clauses that consist of only one positive literal (these clauses are called *positive unit clauses*). For each negative literal $\neg\ell$ appearing in the instance, we can use the algorithm above to test whether U is consistent and whether it implies ℓ. If U is inconsistent, we reject the instance. If ℓ is implied by U, we remove every occurrence of $\neg\ell$ in the input. If we derive the empty clause, we reject the input. Otherwise, the resolution stabilises in a polynomial number of steps with a set of Horn clauses; in this case, accept the input. Since the resulting clauses are Horn, they are preserved by si. We apply si to show that in this case indeed there exists a solution. By assumption, for each Horn clause $\bigwedge_i \mathrm{succ}^{p_i}(x_i, y_i) \Rightarrow \mathrm{succ}^p(x, y)$ there exists an assignment

that falsifies some literal $\mathrm{succ}^{p_i}(x_i, y_i)$ and additionally satisfies all the positive unit clauses: otherwise the literal would have been removed by the resolution procedure. Let $s_1, \ldots, s_r$ be those assignments for the r clauses. Since si is an isomorphism, the assignment $s := si(s_1, \ldots, si(s_{r-1}, s_r) \ldots)$ simultaneously breaks all the equalities in the premises of all the clauses. Moreover, since si preserves succ, the resulting assignment s preserves the positive unit clauses, and hence is a valid assignment for the input. □

6.5.2 Modular minimum and modular maximum

Theorem 6.38. *Let $\mathcal{B}$ be a finite-signature first-order reduct of $(\mathbb{Z}; <)$ that admits a modular max or modular min polymorphism. Then* $\mathrm{CSP}(\mathcal{B})$ *is in* P.

Proof. Suppose that $\mathcal{B}$ is preserved by max, the regular maximum operation. Then $\mathrm{CSP}(\mathcal{B})$ is solvable in polynomial time as follows. Let q be the qe-degree of $\mathcal{B}$. Let ϕ be an instance of $\mathrm{CSP}(\mathcal{B})$ with n variables. We already noted in the proof of Proposition 6.1 that ϕ is satisfiable in $\mathcal{B}$ iff it is satisfiable in $\mathcal{B}[\{0, \ldots, (q+1)n\}]$, and the latter structure can be constructed in polynomial time, and is preserved by the maximum function on $\{0, \ldots, (q+1)n\}$. We can then decide whether $\mathcal{B}[\{0, \ldots, (q+1)n\}] \models \phi$ using the arc-consistency algorithm, noting that the arc-consistency procedure can be implemented in such a way that the running time is polynomial in both the size of the formula and of the structure [74].

Suppose now that $\mathcal{B}$ is preserved by $\max_d$ for $d \geq 2$. It follows that $<$ is not pp-definable in $\mathcal{B}$, as $\max_d$ does not preserve $<$. We can suppose that $\mathcal{B}$ pp-defines succ, because this only increases the complexity of $\mathrm{CSP}(\mathcal{B})$ and succ is preserved by $\max_d$. By Lemma 6.35, $\mathcal{B}$ is a first-order reduct of $(\mathbb{Z}; \mathrm{succ})$. In [15], the authors prove that the CSP of a first-order reduct of $(\mathbb{Z}; \mathrm{succ})$ with finite distance degree and which is preserved by a modular maximum or minimum is decidable in polynomial time. An inspection of the proof shows that the finite distance degree hypothesis is not necessary. Indeed, the critical idea of the algorithm is that if $\mathcal{B}$ is preserved by the d-modular maximum, then $\mathrm{CSP}(\mathcal{B})$ reduces in polynomial time to $\mathrm{CSP}(\mathcal{C})$, where $\mathcal{C}$ is a reduct of $(\mathbb{Z}; \mathrm{succ})$ which is preserved by the usual maximum or minimum. Then, arc-consistency can be used to solve $\mathrm{CSP}(\mathcal{C})$ in polynomial time (for the details, we have to refer to [15]). The reduction and the algorithm for $\mathrm{CSP}(\mathcal{C})$ do not rely on the distance degree of $\mathcal{B}$ being finite to work. □

6.6 The Classification

In this section we prove our complexity classification result, Theorem 6.5. By Theorem 6.6 and the comments before and after Proposition 6.7, we are left with the task to classify the CSP for finite-signature reducts $\mathcal{B}$ of $(\mathbb{Z}; <)$ where the binary relation succ is among the relations of $\mathcal{B}$ (that is, when we are in case (4) or (5) of Theorem 6.6).

An important case distinction in this section is whether the order relation $<$ is pp-definable in $\mathcal{B}$. The situation when this is the case is treated in Section 6.6.1. Otherwise, if succ is pp-definable in $\mathcal{B}$, but $<$ is not, then $\mathcal{B}$ is a first-order reduct of $(\mathbb{Z}; \mathrm{succ})$ by Lemma 6.35. In this case, we further distinguish whether $\mathcal{B}$ is *positive* in the sense that each of its relations can be defined over $(\mathbb{Z}; \mathrm{succ})$ with a *positive* quantifier-free formula, that is, a first-order formula without negation symbols. Positivity of reducts of $(\mathbb{Z}; \mathrm{succ})$ has several natural different characterisations, which is the topic of Section 6.6.2. We

first treat the case of non-positive reducts of $(\mathbb{Z}; \mathrm{succ})$ in Section 6.6.3, and then the case of positive reducts of $(\mathbb{Z}; \mathrm{succ})$ in Section 6.6.4. All the formulas considered here are quantifier-free unless stated otherwise.

6.6.1 First-order expansions of $(\mathbb{Z}; \mathrm{succ}, <)$

We have already seen that the CSP for first-order reducts of $(\mathbb{Z}; <)$ preserved by max or by min is in P. The following lemma provides the matching hardness result for first-order expansions of $(\mathbb{Z}; <, \mathrm{succ})$.

Definition 6.9. A *d-progression* is a set of the form $[a, b \mid d] := \{a, a+d, a+2d, \ldots, b\}$, for $a \leq b$ with $b - a$ divisible by d. A d-progression is *trivial* if it has cardinality one.

We need the following, which is Proposition 47 from [15]. Remember that a structure definable over $(\mathbb{Z}; \mathrm{succ})$ is *locally finite* if every relation has finite distance degree.

Proposition 6.39. *Let $\mathcal{B}$ be a locally finite first-order expansion of $(\mathbb{Z}; \mathrm{succ})$ such that Diff_S is pp-definable in $\mathcal{B}$ for a non-trivial 1-progression S. If $\mathcal{B}$ is neither preserved by max nor min then $\mathrm{CSP}(\mathcal{B})$ is NP-hard.*

Lemma 6.40. *Let $\mathcal{B}$ be a first-order expansion of $(\mathbb{Z}; <, \mathrm{succ})$. If $\mathcal{B}$ is preserved by neither* max *nor* min*, then $\mathrm{CSP}(\mathcal{B})$ is NP-hard.*

Proof. Let R be a relation of $\mathcal{B}$ which is not preserved by max, and let T be a relation of $\mathcal{B}$ which is not preserved by min. Then there are tuples $\overline{a}, \overline{b}$ in R such that $\max(\overline{a}, \overline{b}) \notin R$. Let m be $\max_{i,j}(|a_i - a_j|, |b_i - b_j|)$. Since the binary relation defined by $x \leq y + m$ has a pp-definition in $\mathcal{B}$, the relation R^* defined by

$$R(x_1, \ldots, x_n) \wedge \bigwedge_{i,j} x_i \leq x_j + m$$

is pp-definable in $\mathcal{B}$, too. Note that $\overline{a}$ and $\overline{b}$ are in R^*, and that $\max(\overline{a}, \overline{b}) \notin R^*$. Also note that R^* is first-order definable over succ and has finite distance degree. Dually, we find a pp-definition over $\mathcal{B}$ of a relation T^* which is not preserved by min, first-order definable over succ and with finite distance degree. The pp-formula $\exists u(u = \mathrm{succ}^3(x) \wedge x < y \wedge y < u)$ defines $\mathrm{Diff}_{\{1,2\}}$. It then follows from Proposition 6.39 that $\mathrm{CSP}(\mathbb{Z}; \mathrm{succ}, R^*, T^*)$ is NP-hard, and hence $\mathrm{CSP}(\mathcal{B})$ is NP-hard, too. □

6.6.2 Endomorphisms of and Definability in positive reducts

Positivity of reducts $\mathcal{B}$ of $(\mathbb{Z}; \mathrm{succ})$ can be characterised via the endomorphisms of $\mathbb{Q}.\mathcal{B}$, but also via the non-definability of certain binary relations with pp-formulas (Lemma 6.41). These binary relations then play an important role in the complexity classification of the non-positive reducts of $(\mathbb{Z}; \mathrm{succ})$.

Binary relations R with a first-order definition in $(\mathbb{Z}; \mathrm{succ})$ come in two flavours. Indeed, the set $\{x - y \mid (x, y) \in R\}$ is either finite or cofinite. This easily follows from the quantifier elimination in $(\mathbb{Z}; \mathrm{succ})$. Remember that a binary relation R that is first-order definable over $(\mathbb{Z}; \mathrm{succ})$ (or over $(\mathbb{Q}.\mathbb{Z}; \mathrm{succ})$) is called *basic* if it is empty, $\mathbb{Z}^2$, or defined by the formula $y = x + c$ for some $c \in \mathbb{Z}$, and *non-basic* otherwise.

In the following, we use expressions of the form $\mathrm{succ}^p(x, y)$ (see Example 11) as if they were atomic symbols of the language. Since they are pp-definable in a first-order expansion of $(\mathbb{Z}; \mathrm{succ})$, this is without loss of generality. Recall that a formula over succ is *positive* if it only includes literals of the form $\mathrm{succ}^p(x, y)$. A formula over the signature of $(\mathbb{Z}; \mathrm{succ})$ in disjunctive normal form (DNF) is called *reduced* when every formula obtained by removing literals or conjunctive clauses is not logically equivalent over $(\mathbb{Z}; \mathrm{succ})$. It is clear that every first-order formula on $(\mathbb{Z}; \mathrm{succ})$ is equivalent to a reduced formula in DNF.

Lemma 6.41. *For a first-order expansion $\mathcal{B}$ of $(\mathbb{Z}; \mathrm{succ})$, the following are equivalent:*

1. *Every formula in reduced DNF that defines a relation of $\mathcal{B}$ is positive;*
2. *$\mathbb{Q}.\mathcal{B}$ has an endomorphism that violates the binary relation given by $x - y = \infty$;*
3. *$\mathcal{B}$ does not pp-define a non-basic binary relation with infinite distance degree.*

Proof. (2) implies (1). Let e be an endomorphism of $\mathbb{Q}.\mathcal{B}$ that violates $x - y = \infty$, and let a, b be such that $a - b = \infty$ and $e(a) - e(b) < \infty$. Using automorphisms of $(\mathbb{Q}.\mathbb{Z}; \mathrm{succ})$, we may assume that $e(a) = e(b) = b$ without loss of generality. For contradiction, suppose that $\mathcal{B}$ has a relation with a reduced DNF definition $\phi(x_1, \ldots, x_n)$ which is not positive.

We now show that we can choose $s\colon \{x_1, \ldots, x_n\} \to \mathbb{Z}$ such that s is a satisfying assignment for ϕ but $e \circ s$ is not. For this, let us write one of the non-positive disjuncts ψ of ϕ as $\neg\mathrm{succ}^p(z_2, z_1) \wedge \phi'$ where ϕ' is a conjunction of literals, $z_1, z_2 \in \{x_1, \ldots, x_n\}$, and $p \in \mathbb{Z}$. Moreover, let $\psi_2, \ldots, \psi_m$ be the other disjuncts of ϕ. Suppose that all assignments that satisfy $\phi' \wedge \mathrm{succ}^p(z_2, z_1)$ also satisfy $\bigvee_{2 \leq i \leq m} \psi_i$. Then we could rewrite ϕ as $\phi' \vee \bigvee_{i \geq 2} \psi_i$, which is impossible since ϕ is reduced. Hence, there exists $t\colon \{x_1, \ldots, x_n\} \to \mathbb{Z}$ such that t is a satisfying assignment for $\phi' \wedge \mathrm{succ}^p(z_2, z_1)$ but not for ψ_i for every $i \geq 2$; in particular, t does not satisfy ϕ. Using an automorphism of $(\mathbb{Q}.\mathbb{Z}; \mathrm{succ})$, we can assume that $t(z_1) = b$. Moreover, we can assume that the image S of t lies in only one copy of $\mathbb{Z}$. To see this, let $g\colon S \to \mathbb{Q}.\mathbb{Z}$ be any function that maps S to the first copy of $\mathbb{Z}$ in such a way that if $t(x_i)$ and $t(x_j)$ are in different copies, then $g(t(x_i))$ and $g(t(x_j))$ are at distance at least $q + 1$, where q is the qe-degree of ϕ. We have that g is $\sim_q$-equivalent to an embedding of S into the first copy of $\mathbb{Z}$ in $\mathbb{Q}.\mathbb{Z}$. Therefore, by the substitution lemma (Lemma 6.8), the function $g \circ t$ is a satisfying assignment to the variables of ϕ that only occupies one copy of $\mathbb{Z}$.

We now derive from t an assignment s that satisfies $\neg\mathrm{succ}^p(z_2, z_1)$, that gives the same truth value as t to all the other literals of ψ, and such that $e \circ s = t$. If we consider ϕ' as a graph on $\{z_1, \ldots, z_k\}$ where edges represent positive literals, then z_1 and z_2 are in different connected components. Indeed, if there were a path from z_1 to z_2 in this graph we would have that ϕ' implies a statement of the form $\mathrm{succ}^q(z_2, z_1)$. But then the conjunction $\neg\mathrm{succ}^p(z_2, z_1) \wedge \mathrm{succ}^q(z_2, z_1)$ is either contradictory or is equivalent to $\mathrm{succ}^q(z_2, z_1)$, which is in contradiction to ϕ being reduced. Let V be the variables in the connected component of z_1. Define s on V by $s(v) := a - t(z_1) + t(v)$ (in particular $s(z_1) = a$) and define $s(v) := t(v)$ on the variables v that are not in V. We have that s satisfies $\neg\mathrm{succ}^p(z_2, z_1)$ and that the other literals in ϕ' are satisfied by s, too:

- The truth of positive literals is preserved since we performed a translation on variables that are connected by positive literals.

- Negative literals between the variables in V and the other variables are also true, since for $v \in V$ and $v' \notin V$ we have that $s(v) - s(v') = \infty$ ($s(v)$ lies in the same component as a and $s(v')$ lies in the same component as b.)
- Finally, negative literals between variables not in V are preserved: they are satisfied by t, and for $v \notin V$ we have $s(v) = t(v)$ by definition.

Hence, s is a satisfying assignment of ϕ. We have $e \circ s = t$: If v is a variable in V, then $e(s(v)) = e(a - t(z_1) + t(v)) = e(a) - t(z_1) + t(v) = t(v)$ since e preserves succ and $e(a) = b = t(z_1)$, and if $v \notin V$ we defined $s(v)$ to be $t(v)$, so that $e(s(v)) = e(t(v)) = t(v)$. Since t does not satisfy ϕ, this contradicts the assumption that e is an endomorphism of $\mathbb{Q}.\mathcal{B}$.

(1) implies (3). Let R be a binary relation with a pp-definition $\phi(x, y)$ in $\mathcal{B}$ of the form $\exists \overline{z} \bigwedge_i \phi_i$ where ϕ_i is for each i an atomic formula over $\mathcal{B}$. Let us replace ϕ_i by its definition ψ_i over $(\mathbb{Z}; \mathrm{succ})$ in quantifier-free reduced DNF. By assumption, all the literals in ψ_i are positive. The formula $\phi(x, y)$ is therefore equivalent to a formula $\psi(x, y) := \bigvee_j \exists \overline{z}.\psi_j$ where ψ_j is a conjunction of positive literals of the form $\mathrm{succ}^k(u, v)$. If one of the disjuncts of ψ is vacuously true, then ψ defines a basic binary relation. So let us assume that this is not the case. Since all the literals in ψ_j are positive, the relations defined by the disjuncts have finite distance degree. Their disjunction therefore also defines a binary relation of finite distance degree. In either case, ψ does not define a non-basic binary relation of infinite distance degree.

(3) implies (2). Suppose that all the endomorphisms of $\mathbb{Q}.\mathcal{B}$ preserve the binary relation defined by $x - y = \infty$. Then all the endomorphisms preserve the relation defined by $x \neq y$. Indeed, if $x - y < \infty$ then $e(x) - e(y) = x - y$ since e preserves succ, and hence $x \neq y$ implies $e(x) \neq e(y)$. On the other hand, if $x - y = \infty$, then $e(x) - e(y) = \infty$ by assumption. It follows from Theorem 6.4 that $x \neq y$ has an existential positive definition in $\mathbb{Q}.\mathcal{B}$ and in $\mathcal{B}$. Let $\bigvee \phi_i(x, y)$ be such a definition, where each ϕ_i is a pp-formula over $\mathcal{B}$. Since $\neq$ has infinite distance degree, one of the ϕ_i must define a binary relation with infinite distance degree. This relation is also distinct from $(\mathbb{Q}.\mathbb{Z})^2$ because it is contained in the relation defined by $x \neq y$, so the infinite distance degree implies that it is non-basic. Thus, item (3) does not hold. □

6.6.3 The non-positive case

Let $\mathcal{B}$ be a non-positive reduct of $(\mathbb{Z}; \mathrm{succ})$ such that $\mathbb{Q}.\mathcal{B}$ is not preserved by si. Our aim in this section is to show that $\mathcal{B}$ has an NP-hard CSP. Together with Proposition 6.37, this completes the complexity classification for the CSP of non-positive reducts of $(\mathbb{Z}; \mathrm{succ})$. Note that si is an arbitrary isomorphism $(\mathbb{Q}.\mathbb{Z}; \mathrm{succ})^2 \to (\mathbb{Q}.\mathbb{Z}; \mathrm{succ})$, but the discussion below does not depend on which function we take for si. Indeed, given two isomorphisms si, si' as above, there exists an automorphism α of $(\mathbb{Q}.\mathbb{Z}; \mathrm{succ})$ such that $si = \alpha \circ si'$.

In order to show that $\mathrm{CSP}(\mathcal{B})$ is NP-hard, we show in Proposition 6.44 that when $\mathbb{Q}.\mathcal{B}$ is not preserved by si then there is a non-basic binary relation with finite distance degree that is pp-definable in $\mathcal{B}$. This binary relation will serve to define the set of vertices of a certain finite undirected graph. The edge relation of that graph comes from the binary relation of Lemma 6.41 which provided an alternative characterisation of non-positivity of $\mathcal{B}$. We finally use the classification of the CSPs for finite undirected graphs [55] to conclude that $\mathrm{CSP}(\mathcal{B})$ is NP-hard.

A formula ϕ in CNF is called *reduced* when removing any literal in a clause yields a formula that is not equivalent to ϕ. This is equivalent to saying that for any literal ℓ in a clause ψ of ϕ, there exists an assignment that satisfies ϕ and that satisfies only ℓ in ψ. This assignment witnesses the fact that the given literal cannot be removed from the formula without changing the set of satisfying assignments.

Lemma 6.42. *Let ϕ be a quantifier-free formula over $(\mathbb{Z}; \mathrm{succ})$, and suppose that ϕ is equivalent to a Horn formula over $(\mathbb{Z}; \mathrm{succ})$. If ϕ is reduced, then it is Horn.*

Proof. Note that ϕ is equivalent to a Horn formula over $(\mathbb{Z}; \mathrm{succ})$ if and only if it is equivalent to a Horn formula over $(\mathbb{Q}.\mathbb{Z}; \mathrm{succ})$, since both structures have the same first-order theory. By Proposition 6.36 the formula ϕ is preserved by si.

Suppose for contradiction that ϕ is not Horn, that is, it contains a clause ψ of the form $(\mathrm{succ}^p(y,x) \vee \mathrm{succ}^q(v,u) \vee \dots)$. Since this formula is reduced, there exist satisfying assignments s, t of ϕ such that s satisfies only $\mathrm{succ}^p(y,x)$ in ψ, and t satisfies only $\mathrm{succ}^q(v,u)$ in ψ. The assignment (s,t) that maps a variable x_i of ϕ to the pair $(s(x_i), t(x_i))$ in $(\mathbb{Q}.\mathbb{Z})^2$ is not a satisfying assignment for ϕ. Since si is an isomorphism between $(\mathbb{Q}.\mathbb{Z}; \mathrm{succ})^2$ and $(\mathbb{Q}.\mathbb{Z}; \mathrm{succ})$, we have that the assignment $si(s,t)$ does not satisfy ψ, which contradicts the fact that ϕ is preserved by si. □

Clearly, every formula ϕ in CNF is equivalent to a reduced one, since we can repeatedly remove logically redundant literals until we obtain a reduced formula ϕ': in this case we say that we *obtain ϕ' from reducing ϕ*.

Lemma 6.43. *A binary relation $R \subseteq \mathbb{Z}^2$ is Horn definable over $(\mathbb{Z}; \mathrm{succ})$ if and only if it is basic or has infinite distance-degree.*

Proof. The backward implication is clear, since a binary relation with infinite distance-degree and different from $\mathbb{Z}^2$ can be defined by a conjunction of literals of the form $\neg\mathrm{succ}^p(x,y)$. Basic relations can be defined by a formula of the form $\mathrm{succ}(x,x)$, $x = x$, or $\mathrm{succ}^c(x,y)$, which are all Horn formulas.

Let us prove the forward implication. Let $\phi(x,y)$ be a reduced Horn quantifier-free formula. In every clause of ϕ, there is at most one positive literal. Note that two negative literals cannot appear in the same clause of ϕ, for the disjunction $\neg\mathrm{succ}^c(x,y) \vee \neg\mathrm{succ}^d(x,y)$ is either trivial if $c \neq d$ or equivalent to a single literal if $c = d$, and ϕ is assumed to be reduced. Similarly, a positive literal and a negative literal cannot appear in the same clause, because $\mathrm{succ}^c(x,y) \vee \neg\mathrm{succ}^d(x,y)$ is equivalent to $\neg\mathrm{succ}^d(x,y)$ if $c \neq d$, and is vacuously true if $c = d$. Therefore every clause of ϕ contains exactly one literal, so that ϕ is a conjunction of literals. If one of those literals is positive, ϕ is equivalent to $\mathrm{succ}^c(x,y)$ for some c or defines the empty relation, so that the relation that ϕ defines is basic. Otherwise all the literals in ϕ are negative, and ϕ has infinite distance-degree. □

Proposition 6.44. *Let $\mathcal{B}$ be a first-order expansion of $(\mathbb{Z}; \mathrm{succ})$, and suppose that $\mathcal{B}$ pp-defines a relation that is not Horn-definable over $(\mathbb{Z}; \mathrm{succ})$. Then $\mathcal{B}$ also pp-defines a binary relation that is not Horn-definable over $(\mathbb{Z}; \mathrm{succ})$.*

Proof. Let R be a relation with a pp-definition in $\mathcal{B}$ that is not Horn-definable over $(\mathbb{Z}; \mathrm{succ})$, and whose arity n is minimal among all the relations with the same property. We claim that for all $i, j \leq n$ and $p \in \mathbb{Z}$ the relation defined by the formula

$R(x_1, \ldots, x_n) \wedge \mathrm{succ}^p(x_j, x_i)$ is Horn-definable over $(\mathbb{Z}; \mathrm{succ})$. Otherwise, any reduced definition ϕ of this relation over $(\mathbb{Z}; \mathrm{succ})$ has a clause ψ with at least two positive literals ℓ_1 and ℓ_2. Hence, there are satisfying assignment s_1 and s_2 for ϕ such that s_1 only satisfies ℓ_1 in ψ and s_2 only satisfies ℓ_2 in ψ. Let ϕ' be the formula obtained from ϕ by replacing literals of the form $\mathrm{succ}^{p'}(x_j, x_k)$ or $\mathrm{succ}^{-p'}(x_k, x_j)$, for $p' \in \mathbb{Z}$, by $\mathrm{succ}^{p'-p}(x_i, x_k)$. Then the variable x_j no longer occurs in ϕ', and ϕ' is equivalent to $\exists x_j(R(x_1, \ldots, x_n) \wedge \mathrm{succ}^p(x_j, x_i))$. In particular, the restrictions of s_1 and s_2 to $\{x_1, \ldots, x_n\} \setminus \{x_j\}$ are satisfying assignments for ϕ', and they witness that the literals ℓ_1 and ℓ_2 of ϕ (or the literals that correspond to those literals in ϕ') cannot be removed from ϕ'. Lemma 6.42 implies that ϕ' is not equivalent to a Horn formula. Note that ϕ' defines a relation of arity $n-1$ that is not Horn and that is pp-definable in $\mathcal{B}$, a contradiction to the choice of R.

If a binary projection of R is non-basic and has finite distance-degree, then it is not Horn by Lemma 6.43 and we are done. If a binary projection of R is basic, then we have a contradiction to the minimality of n as we have seen above. So we can assume that the binary projections of R have infinite distance degree.

Suppose for contradiction that $n > 2$. Let $\phi(x_1, \ldots, x_n)$ be a reduced quantifier-free formula that defines R in $(\mathbb{Z}; \mathrm{succ})$ whose number of non-Horn clauses is minimal. We first prove that every non-Horn clause of ϕ is positive, i.e., consists of positive literals only. Pick a non-Horn clause ψ of ϕ with two positive literals ℓ_1, ℓ_2, and suppose ψ also contains the negative literal $\neg\mathrm{succ}^p(x_j, x_i)$ for some $i, j \in \{1, \ldots, n\}$ and $p \in \mathbb{Z}$. Since ϕ is reduced, there are satisfying assignment s_1 and s_2 for ϕ such that s_1 only satisfies ℓ_1 in ψ and s_2 only satisfies ℓ_2 in ψ; in particular, both s_1 and s_2 satisfy $\mathrm{succ}^p(x_j, x_i)$. Then these two assignments show that both ℓ_1 and ℓ_2 cannot be removed when reducing $\phi \wedge \mathrm{succ}^p(x_j, x_i)$; by Lemma 6.42, this contradicts the fact that $\phi \wedge \mathrm{succ}^p(x_j, x_i)$ is equivalent to a Horn formula, which was established in the first paragraph of the proof.

Therefore, there exists a positive non-Horn clause ψ in ϕ. Let ϕ' denote the rest of ϕ, and define

$$E_{i,j} := \{s(x_j) - s(x_i) \mid s\colon \{x_1, \ldots, x_n\} \to \mathbb{Z} \text{ satisfies } \phi' \wedge \neg\psi\}.$$

If $E_{i,j}$ is empty for some $i, j \in \{1, \ldots, n\}$, then the formulas ϕ and ϕ' are equivalent. But ϕ' contains fewer non-Horn clauses than ϕ, contradicting the choice of ϕ. By the first paragraph, for all distinct i, j and $p \in E_{i,j}$, the formula $\phi \wedge \mathrm{succ}^p(x_i, x_j)$ is equivalent to a Horn formula, and by Lemma 6.42, it even *reduces* to a Horn formula. Note that since ψ is a positive clause, the only way to reduce $\phi \wedge \mathrm{succ}^p(x_i, x_j)$ to a Horn formula is to remove all literals in ψ but one. Also note that at least one literal of ψ must remain when reducing $\phi \wedge \mathrm{succ}^p(x_i, x_j)$ because we chose p from $E_{i,j}$. This means that there exists a literal $\ell_p^{i,j}$ of ψ such that

$$\phi \wedge \mathrm{succ}^p(x_i, x_j) \models \ell_p^{i,j}\,.$$

Let q be the qe-degree of ϕ. If $p \in E_{i,j}$ is greater than nq, then we may take $\ell_p^{i,j}$ to be $\ell_{nq+1}^{i,j}$, by the substitution lemma.

First consider the case that there are distinct i, j such that $E_{i,j}$ is finite. Then ϕ is equivalent over $(\mathbb{Z}; \mathrm{succ})$ to the formula

$$\chi := \phi' \wedge \bigwedge_{p \in E_{i,j}} (\mathrm{succ}^p(x_i, x_j) \Rightarrow \ell_p^{i,j}).$$

Indeed, ϕ implies χ directly from the hypotheses we have. Conversely, if s satisfies χ one of two cases occur. Either some $\ell_p^{i,j}$, for $p \in E_{i,j}$, is satisfied by s, and then s satisfies ψ and ϕ. Or we must have $s(x_j) \neq s(x_i) + p$ for every $p \in E_{i,j}$, i.e., $s(x_j) - s(x_i) \notin E_{i,j}$. Since s is known to satisfy ϕ', by definition of $E_{i,j}$ it must also satisfy ψ, whence we get that s satisfies ϕ. Note that χ contains fewer non-Horn clauses than ϕ, which contradicts the choice of ϕ.

Therefore, $E_{i,j}$ is not finite, and thus cofinite for all distinct $i, j \leq n$. We claim that ϕ has a satisfying assignment s such that $|s(x_i) - s(x_j)| > 2(n+1)q$ for all distinct $i, j \in \{1, \ldots, n-1\}$. The binary projections of R all have infinite distance degree, so by the substitution lemma we find for each pair (i, j) such that $1 \leq i < j \leq n$ a satisfying assignment $s_{i,j} \colon \{x_1, \ldots, x_n\} \to \mathbb{Q}.\mathbb{Z}$ for ϕ such that $s_{i,j}(x_i) - s_{i,j}(x_j) = \infty$. Also note that the $(n-1)$-projection R' of R onto $\{1, \ldots, n-1\}$ is Horn, and hence preserved by si. Then for $s' \colon \{x_1, \ldots, x_{n-1}\} \to \mathbb{Q}.\mathbb{Z}$ defined by $s'(x) := si(s_{1,2}(x), \ldots si(s_{n-3,n-1}(x), s_{n-2,n-1}(x)) \ldots)$ we have that $s'(x_i) - s'(x_j) = \infty$ for all distinct i, j, and that $(s'(x_1), \ldots, s'(x_{n-1})) \in R'$. Since R' is a projection of R, we can extend s' to a satisfying assignment s'' for ϕ. Again using the substitution lemma, we obtain from s'' a satisfying assignment $s \colon \{x_1, \ldots, x_n\} \to \mathbb{Z}$ for ϕ such that $|s(x_i) - s(x_j)| > 2(n+1)q$ for distinct $i, j \in \{1, \ldots, n-1\}$, and this concludes the proof of the claim.

For ψ to be satisfied by s, there must exist an $i \in \{1, \ldots, n-1\}$ such that $|s(x_i) - s(x_n)| \leq q$, since ψ only contains positive literals of degree at most q. Let $k \in \{1, \ldots, n-1\}$ be different from i. Note that $|s(x_k) - s(x_i)| > nq$ and $|s(x_k) - s(x_n)| > nq$. Also note that s satisfies the literals $\ell_{nq+1}^{k,i}$ and $\ell_{nq+1}^{k,n}$ by the definition of $\ell_{nq+1}^{k,i}$ and $\ell_{nq+1}^{k,n}$. Then the literal $\ell_{nq+1}^{k,i}$ relates x_i and x_n, and so does the literal $\ell_{nq+1}^{k,n}$, because x_i and x_n are with respect to s the only variables that are able to satisfy a positive literal. Since all binary projections of R have infinite distance degree, there is a satisfying assignment t of ϕ such that $|t(x_i) - t(x_n)| > 2(n+1)q$. Either $|t(x_k) - t(x_i)| > nq$ or $|t(x_k) - t(x_n)| > nq$. In the first case, $\ell_{nq+1}^{k,i}$ must be satisfied by t. But $\ell_{nq+1}^{k,i}$ is a literal of the form $\mathrm{succ}^{p_1}(x_n, x_i)$ for $|p_1| \leq q$, and $|t(x_i) - t(x_n)| > nq$, so t cannot satisfy $\ell_{nq+1}^{k,i}$. Similarly, in the second case, t must satisfy $\ell_{nq+1}^{k,n}$, which is impossible since this literal is of the form $\mathrm{succ}^{p_2}(x_n, x_i)$ for $|p_2| \leq q$. We have reached a contradiction. Therefore, we must have $n = 2$, and R is the desired binary non-Horn relation with a pp-definition over $\mathcal{B}$. □

We can finally conclude the complexity classification for non-positive first-order expansions of $(\mathbb{Z}; \mathrm{succ})$.

Proposition 6.45. *Let $\mathcal{B}$ be a non-positive first-order expansion of $(\mathbb{Z}; \mathrm{succ})$. Then $\mathbb{Q}.\mathcal{B}$ is preserved by si and $\mathrm{CSP}(\mathcal{B})$ is in* P*, or $\mathrm{CSP}(\mathcal{B})$ is* NP*-hard.*

Proof. If $\mathbb{Q}.\mathcal{B}$ is preserved by si then $\mathrm{CSP}(\mathcal{B})$ is in Pby Proposition 6.37. Otherwise, Proposition 6.36 implies that $\mathcal{B}$ has a non-Horn relation. By Proposition 6.44, a binary non-Horn relation is pp-definable in $\mathcal{B}$. A binary relation which is definable over $(\mathbb{Z}; \mathrm{succ})$ but not Horn is non-basic and has finite distance degree, by Lemma 6.43. Hence, a non-basic binary relation T of finite distance degree is pp-definable in $\mathcal{B}$.

By Lemma 6.41, there exists a non-basic binary relation N pp-definable in $\mathcal{B}$ and which has infinite distance degree. The relation defined by $N(x, y) \wedge N(y, x)$ in $\mathcal{B}$ is symmetric and has infinite distance degree, and is again pp-definable in $\mathcal{B}$, so we will assume that N is already symmetric. Let a be the smallest positive integer such that

$(0, b)$ is in N for all $b \geq a$. With succ and pp-definition, we may assume that T contains $(0, 0)$ and that N does not contain $(0, 0)$. Then by repeatedly replacing T by the pp-definable relation $\{(x, y) \in \mathbb{Z}^2 \mid \exists z \in \mathbb{Z} : (x, z) \in T \wedge (z, y) \in T\}$ we may assume that $(0, b), (0, 2b) \in T$ with $b \geq a$. Let $\mathcal{G}$ be the undirected graph whose vertices are the integers v such that $(0, v) \in T$, and where v and w are adjacent if $(v, w) \in N$. This graph has no loop and contains the triangle $\{0, b\}, \{b, 2b\}, \{0, 2b\}$, so that G is not bipartite and $\mathrm{CSP}(\mathcal{G})$ is NP-hard [55]. Furthermore, $\mathrm{CSP}(\mathcal{G})$ is polynomial-time reducible to $\mathrm{CSP}(\mathcal{B})$: if $\exists x_1, \ldots, x_n.\phi$ is an instance of $\mathrm{CSP}(G)$, create an instance of $\mathrm{CSP}(\mathcal{B})$ by adding an existentially quantified variable z, and by adding the constraints $T(z, x_i)$ for all i. This instance is satisfiable if and only if the original instance is satisfiable in $\mathcal{G}$, using the fact that the automorphism group of $\mathcal{B}$ is transitive. This proves that $\mathrm{CSP}(\mathcal{B})$ is NP-hard. $\square$

6.6.4 The positive case

We prove in this section that a positive first-order expansion $\mathcal{B}$ of $(\mathbb{Z}; \mathrm{succ})$ which is not preserved by any d-modular maximum or minimum has an NP-hard CSP. As in the non-positive case, an important step of the classification is to show that there exists a non-basic binary relation with a pp-definition in $\mathcal{B}$.

Let R be a relation of arity n with a first-order definition ϕ over $\mathcal{B}$. We say that R is *r-decomposable* if $\phi(x_1, \ldots, x_n)$ is equivalent over $\mathcal{B}$ to

$$\bigwedge_{1 \leq i_1 < \cdots < i_{n-r} \leq n} \exists x_{i_1}, \ldots, x_{i_{n-r}}.\phi(x_1, \ldots, x_n)\,.$$

The following lemma states that a positive first-order expansion of $(\mathbb{Z}; \mathrm{succ})$ which is not preserved by a modular maximum or minimum pp-defines a non-basic binary relation. It is a positive variant of Lemma 38 in [15], and its proof is essentially the same. Intuitively this is because in both cases the binary relations that are pp-definable in $\mathcal{B}$ have either finite distance degree or are $\mathbb{Z}^2$ (if $\mathcal{B}$ has finite distance degree this is immediate, and when $\mathcal{B}$ is positively definable in $(\mathbb{Z}; \mathrm{succ})$ this is the content of Lemma 6.41). For the sake of completeness, we reproduce the proof with the necessary adjustments.

Lemma 6.46. *Let $\mathcal{B}$ be a positive first-order expansion of $(\mathbb{Z}; \mathrm{succ})$ without a modular max or a modular min as polymorphism. Then there is a non-basic binary relation pp-definable in $\mathcal{B}$ which has a finite distance degree.*

Proof. The binary relations pp-definable in $\mathcal{B}$ are either basic, or non-basic and of finite distance degree, by the fact that $\mathcal{B}$ is positive and Lemma 6.41. Suppose for contradiction that all the binary relations with a pp-definition in $\mathcal{B}$ are basic.

If every relation S of $\mathcal{B}$ were 2-decomposable then $\mathcal{B}$ would be invariant under max: indeed, we assumed that the binary relations pp-definable in $\mathcal{B}$ are already pp-definable in $(\mathbb{Z}; \mathrm{succ})$, so that a 2-decomposable relation S already has a pp-definition in $(\mathbb{Z}; \mathrm{succ})$, and is thus preserved by max. Therefore, $\mathcal{B}$ contains a relation that is not 2-decomposable. This implies that, by projecting out coordinates from S, we can obtain a relation R of arity $r \geq 3$ which is not $(r-1)$-decomposable.

This implies, in particular, that there exists a tuple $(a_1, \ldots, a_r) \notin R$ such that for all $i \in \{1, \ldots, r\}$, there exists some integer p_i such that $(a_1, \ldots, p_i, \ldots, a_r) \in R$. By replacing

R by the relation with the pp-definition

$$\exists y_1, \dots, y_r \big(\bigwedge_{i \in [r]} (y_i = x_i + a_i) \wedge R(y_1, \dots, y_r) \big)$$

we can further assume that $a_i = 0$ for all $i \in [r]$. We can also assume, w.l.o.g., that $p_1 \neq -p_2$ because $r \geq 3$.

Suppose that the arity of R is greater than 3, and consider the ternary relation $T(x_1, x_2, x_3)$ defined by $R(x_1, x_2, x_3, \dots, x_3)$. If there is a $z \in \mathbb{Z}$ so that $R(0, 0, z, \dots, z)$, then T would not be 2-decomposable, since $(0, 0, 0) \notin T$, although $(p_1, 0, 0), (0, p_2, 0)$, and $(0, 0, z)$ are all in T. This contradicts the minimality of the arity of R. If there is no such z then $\exists x_3.R(x_1, x_2, x_3, \dots, x_3)$ defines a binary relation omitting $(0, 0)$ and containing $(0, -p_1)$ and $(0, p_2)$. This relation is binary, pp-definable in $\mathcal{B}$, and non-basic, contradiction.

Suppose now that $r = 3$. We claim that every binary projection of R is $\mathbb{Z}^2$. Suppose otherwise that one such binary projection, say the one defined by $\exists x_1.R(x_1, x_2, x_3)$, is not of this form. By assumption, all binary relations with a pp-definition in $\mathcal{B}$ are basic, so this formula is equivalent to $x_3 = x_2 + p$ for some $p \in \mathbb{Z}$. Let $(a, b, c) \in \mathbb{Z}^3$ be such that

- (a, b) is in the projection of R onto $\{1, 2\}$,
- (a, c) is in the projection of R onto $\{1, 3\}$, and
- (b, c) is in the projection of R onto $\{2, 3\}$ (i.e., $c = b + p$).

Since (a, b) is in the projection of R onto $\{1, 2\}$, there exists $d \in \mathbb{Z}$ such that $(a, b, d) \in R$. Since the projection of R onto $\{2, 3\}$ is basic we have $d = b + p = c$, so that (a, b, c) is in R. Hence, R is 2-decomposable, contradicting our assumptions. This shows the claim.

Let $\phi(x_1, x_2, x_3)$ be a positive formula in reduced DNF defining R over $(\mathbb{Z}; \mathrm{succ})$. This formula has at least two disjuncts, otherwise R would be pp-definable over $(\mathbb{Z}; \mathrm{succ})$. Each disjunct contains at most two literals, because it suffices to describe only two distances between three variables to determine the type of a triple of integers. We claim that there is a disjunct in ϕ that consists of only one literal. If that was not the case, every disjunct would have two literals and would be equivalent to $\mathrm{succ}^{p_i}(x_2, x_1) \wedge \mathrm{succ}^{q_i}(x_3, x_1)$ for some $p_i, q_i \in \mathbb{Z}$. In this case, the formula $\exists x_2.\phi(x_1, x_2, x_3)$ defines a binary relation with finite distance degree, contradicting the claim established in the previous paragraph. Furthermore, there are at least two such disjuncts: if there were only one, say $\mathrm{succ}^p(x_2, x_1)$, the relation defined by $\exists x_3.\phi(x_1, x_2, x_3)$ is binary and has a finite distance degree, a contradiction. Hence, there are at least two disjuncts in ϕ that contain only one literal. One of x_1, x_2, x_3 must appear twice in those literals, and we may assume by permuting the variables that it is x_1. Let us write these literals as $\mathrm{succ}^p(x_2, x_1)$ and $\mathrm{succ}^q(x_3, x_1)$, for $p, q \in \mathbb{Z}$. Then the formula $\exists x_3 \big(\phi(x_1, x_2, x_3) \wedge \mathrm{succ}^{p-q+1}(x_2, x_3)\big)$ is equivalent to a binary DNF which is reduced and contains the two disjuncts $\mathrm{succ}^p(x_2, x_1)$ and $\mathrm{succ}^{p+1}(x_2, x_1)$. The relation defined by this formula has finite distance degree, again contradicting our assumptions.

It follows that there exists a non-basic binary relation pp-definable in $\mathcal{B}$, and this relation has finite distance degree by positivity of $\mathcal{B}$. □

The following is Lemma 43 in [15].

Lemma 6.47. *Let $S \subset \mathbb{Z}$ be finite with $|S| > 1$, and let d be the greatest common divisor of all $a - a'$ for $a, a' \in S$. Then for any d-progression T, the relation Diff_T is pp-definable in $(\mathbb{Z}; \mathrm{succ}, \mathrm{Diff}_S)$.*

Proposition 6.48. *Let $\mathcal{B}$ be a first-order expansion of $(\mathbb{Z}; \mathrm{succ})$, and $S \subset \mathbb{Z}$ a 1-progression with $|S| > 1$, such that Diff_S is pp-definable in $\mathcal{B}$. Then $\mathcal{B}$ is preserved by max or min; or $\mathrm{CSP}(\mathcal{B})$ is NP-hard.*

Proof. Suppose that $\mathcal{B}$ is not preserved by max nor min. Therefore, there exist in $\mathcal{B}$ a relation $R \subseteq \mathbb{Z}^n$ that is not preserved by max and a relation $T \subseteq \mathbb{Z}^m$ which is not preserved by min. This means that there are tuples $\overline{a}, \overline{b}$ in R such that $\max(\overline{a}, \overline{b})$ is not in R and similarly for T. By hypothesis and Lemma 6.47, all the 1-progressions are definable in $\mathcal{B}$. Let M be $\max_{i,j}\{|a_i - a_j|, |b_i - b_j|\}$, and let ϕ be the pp-definition of $\mathrm{Dist}_{[0,M|1]}$ in $\mathcal{B}$. Define the relation R^* by $R(x_1, \ldots, x_n) \wedge \bigwedge_{i,j \leq n} \phi(x_i, x_j)$ and analogously define T^* from T. We have that $\overline{a}, \overline{b} \in R^*$ by construction, and still $\max(\overline{a}, \overline{b}) \notin R^*$ since $R^* \subseteq R$. Also note that R^* has finite distance degree. Analogously, S^* is not preserved by min and has finite distance degree. It follows from Proposition 6.39 that $\mathrm{CSP}(\mathbb{Z}; \mathrm{succ}, \mathrm{Diff}_S, R^*, T^*)$ is NP-hard. Therefore, $\mathrm{CSP}(\mathcal{B})$ is also NP-hard. □

We can now prove the complexity classification for positive first-order expansions of $(\mathbb{Z}; \mathrm{succ})$.

Proposition 6.49. *Let $\mathcal{B}$ be a positive first-order expansion of $(\mathbb{Z}; \mathrm{succ})$. Then $\mathcal{B}$ is preserved by a modular max or a modular min, and $\mathrm{CSP}(\mathcal{B})$ is in P, or $\mathrm{CSP}(\mathcal{B})$ is NP-hard.*

Proof. If $\mathcal{B}$ is preserved by a modular max or a modular min, then $\mathrm{CSP}(\mathcal{B})$ is in P by Theorem 6.38, so assume that this is not the case. Lemma 6.46 implies there exists a non-basic binary relation R with finite distance degree and a pp-definition in $\mathcal{B}$. Lemma 44 in [15] states that if $S \subset \mathbb{Z}$ is finite, but not a d-progression, for any $d > 0$, then $\mathrm{CSP}(\mathbb{Z}; \mathrm{succ}, \mathrm{Diff}_S)$ is NP-hard. Hence, if R is not a d-progression for any $d \geq 1$, then $\mathrm{CSP}(\mathcal{B})$ is NP-hard. So let us assume that R is a d-progression for some $d \geq 1$.

Since $\mathcal{B}$ is not preserved by $\max_d$, it contains a relation T_1 containing tuples $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ such that $(\max_d(a_1, b_1), \ldots, \max_d(a_n, b_n)) \notin T_1$. Since $\mathcal{B}$ is positive, the binary projections of T_1 have finite distance degree by Lemma 6.41. By the same argument as above, we can suppose that these binary projections are arithmetic progressions unless $\mathrm{CSP}(\mathcal{B})$ is NP-hard. Lemma 45 in [15] establishes that $\mathrm{CSP}(\mathcal{B})$ is NP-hard unless all these arithmetic progressions are d-progressions. Using succ, we easily see that T_1 pp-defines a relation T_1' that is not preserved by $\max_d$ and such that all the differences $a_i - a_j$, for $(a_1, \ldots, a_n) \in T_1'$, are divisible by d. Therefore we can pick two tuples $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ in T_1' whose entries are divisible by d, and such that $(\max_d(a_1, b_1), \ldots, \max_d(a_n, b_n)) \notin T_1'$. Similarly, we obtain a relation T_2' which has a pp-definition in $\mathcal{B}$ and which contains tuples $(c_1, \ldots, c_m), (d_1, \ldots, d_m)$ whose entries are all divisible by d and such that $(\min_d(c_1, d_1), \ldots, \min_d(c_m, d_m))$ is not in T_2'. Let $\mathcal{B}'$ be $(\mathbb{Z}; \mathrm{succ}^d, R, T_1', T_2')$. It follows from our construction that $\mathcal{B}'/d$ is not preserved by max nor by min. Moreover $\mathcal{B}'$ contains the non-basic d-progression R, so that $\mathcal{B}'/d$ contains a non-basic 1-progression. By Proposition 6.48, we obtain that $\mathrm{CSP}(\mathcal{B}'/d)$ is NP-hard.

Now we reduce $\mathrm{CSP}(\mathcal{B}'/d)$ to $\mathrm{CSP}(\mathcal{B})$ to prove that the latter is also NP-hard. By Lemma 43 in [15], the relation $\mathrm{Dist}_{[0,d|d]}$ has a pp-definition in $\mathcal{B}$. Let q be the qe-degree

of $\mathcal{B}$ and note that an instance of $\mathcal{B}$ on n variables has a solution iff it has a solution on the interval $[0, qn]$. From an instance Φ of $\mathrm{CSP}(\mathcal{B}'/d)$ we build an instance Ψ of $\mathrm{CSP}(\mathcal{B})$. The variables of Ψ consist of the variables of Φ and additionally $qn - 1$ new variables $x_1, \ldots, x_{qn-1}$ for each extant variable x of Φ, and finally an additional new variable z. The constraints of Ψ are as follows:

- for each constraint of Φ using the relations $succ^d, R, T_1'$, and T_2', we add a constraint to Ψ using the pp-definitions of these relations in $\mathcal{B}$,
- for each variable x of Φ, Ψ contains the constraint $\mathrm{Dist}_{[0,qdn|d]}(x, z)$, that we define by the conjunction $\mathrm{Dist}_{[0,d|d]}(x, x_1) \wedge \mathrm{Dist}_{[0,d|d]}(x_1, x_2) \wedge \ldots \wedge \mathrm{Dist}_{[0,d|d]}(x_{qn-1}, z)$.

It is straightforward to see that $\mathcal{B}/d \models \Phi$ iff $\mathcal{B} \models \Psi$ and the result follows. □

6.6.5 Concluding the classification

We can finally combine the structural classification of first-order reducts of $(\mathbb{Z}; <)$ (Theorem 6.6) with the complexity classification of the previous sections.

Theorem 6.5. Let $\mathcal{B}$ be a reduct of $(\mathbb{Z}; <)$ with finite signature. Then there exists a structure $\mathcal{C}$ such that $\mathrm{CSP}(\mathcal{C})$ equals $\mathrm{CSP}(\mathcal{B})$ and one of the following cases applies.

1. $\mathcal{C}$ has a finite domain, and the CSP for $\mathcal{B}$ is in P or NP-complete.
2. $\mathcal{C}$ is a reduct of $(\mathbb{Q}; <)$ and $\mathrm{CSP}(\mathcal{C})$ is in P or NP-complete.
3. $\mathcal{C}$ is a reduct of $(\mathbb{Z}; <)$ and preserved by max or by min. In this case, $\mathrm{CSP}(\mathcal{C})$ is in P.
4. $\mathcal{C}$ is a reduct of $(\mathbb{Z}; \mathrm{succ})$ such that $\mathcal{C}$ is preserved by a modular max or min, or $\mathbb{Q}.\mathcal{C}$ is preserved by a binary injective function preserving succ. In this case, $\mathrm{CSP}(\mathcal{C})$ is in P.
5. $\mathrm{CSP}(\mathcal{B})$ is NP-complete.

Proof. Let $\mathcal{B}$ be a finite signature reduct of $(\mathbb{Z}; <)$. By Proposition 6.1, $\mathrm{CSP}(\mathcal{B})$ is in NP. If $\mathcal{B}$ is homomorphically equivalent to a finite structure, we are in case (1) of the statement and there is nothing to be shown. Otherwise, Theorem 6.6 implies that there exists a reduct $\mathcal{C}$ of $(\mathbb{Z}; <)$ such that $\mathrm{CSP}(\mathcal{B})$ equals $\mathrm{CSP}(\mathcal{C})$, and one of the following cases applies.

(a) $\mathcal{C}$ is a reduct of $(\mathbb{Q}; <)$. We are in case (2) of the statement; the complexity of $\mathrm{CSP}(\mathcal{C})$ has been classified in Theorem 50 in [23].

(b) For all $k \geq 1$, the relation $\mathrm{Dist}_{\{k\}}$ is pp-definable; in this case, $\mathrm{CSP}(\mathcal{B})$ and $\mathrm{CSP}(\mathcal{C})$ are NP-hard by Proposition 6.7. Hence, we are in case (5) of the statement.

(c) The relation succ is pp-definable in $\mathcal{C}$. If $<$ is pp-definable in $\mathcal{C}$, then Lemma 6.40 and Theorem 6.38 imply that we are in case (3) or (5) of the statement. Otherwise $\mathcal{C}$ is a reduct of $(\mathbb{Z}; \mathrm{succ})$, by Lemma 6.35. If $\mathcal{C}$ is non-positive then the statement follows from Proposition 6.45, if it is positive then the statement follows from Proposition 6.49.

□

6.7 Conclusion

In this chapter, we proved that for finitely many relations $R_1, \ldots, R_k$ that are first-order definable over $(\mathbb{Z}; <)$, the constraint satisfaction problem for $\mathcal{B} = (\mathbb{Z}; R_1, \ldots, R_k)$ satisfies a complexity dichotomy. In the case that $\mathcal{B}$ contains the successor relation (we showed that the complexity classification can be reduced to this situation), the dichotomy has an elegant algebraic formulation using the ω-saturated expansion $\mathbb{Q}.\mathcal{B}$ of $\mathcal{B}$:

- $\mathbb{Q}.\mathcal{B}$ has a modular maximum/minimum polymorphism or a binary injective polymorphism and the CSP of $\mathcal{B}$ is in P, or
- $\mathbb{Q}.\mathcal{B}$ omits these polymorphisms and the CSP of $\mathcal{B}$ is NP-complete.

These results are important foundations for the investigation of the complexity of CSPs for constraint languages that are definable in Presburger arithmetic, i.e., definable over $(\mathbb{Z}; +, <)$. In the next chapter, we continue this investigation with constraint languages that are definable over $(\mathbb{Z}; +, 1)$ and that contain $+$.

We note that the important Max-Atom problem [8] can be formulated as the CSP for a first-order reduct of $(\mathbb{Z}; <)$ whose signature is *infinite*. In order to define the computational problem $\mathrm{CSP}(\mathcal{B})$ for a structure with an infinite relational signature we have to discuss how the relation symbols of $\mathcal{B}$ are represented in the input. As explained in Chapter 2, the classical CSP is only defined for finite relational signatures; moving to infinite signatures requires specifying an encoding for the constraints of the input. For example, if we represent a relation symbol R in a first-order reduct $\mathcal{B}$ of $(\mathbb{Z}; <)$ by the first-order formula that defines $R^{\mathcal{B}}$, we can no longer expect polynomial-time algorithms for $\mathrm{CSP}(\mathcal{B})$ since already the problem to decide whether a single constraint in the input is satisfiable becomes PSPACE-complete. However, for first-order reducts of $(\mathbb{Z}; <)$ with infinite signature there is a natural candidate for an input encoding of the relation symbols of $\mathcal{B}$ that still allows for polynomial-time algorithms for $\mathrm{CSP}(\mathcal{B})$, which we describe in the following. Each constraint $R(x_1, \ldots, x_k)$ is represented by a quantifier-free definition of R over $(\mathbb{Z}; <)$ in reduced disjunctive normal form, where a literal $x_i \leq x_j + k$ is encoded by giving k in binary representation. When the input is represented in this way, $\mathrm{CSP}(\mathcal{B})$ is still in NP, by the same argument as in Proposition 6.1.

The Max-Atom problem is the CSP for the first-order reduct of $(\mathbb{Z}; <)$ that contains all the relations of the form

$$\{(x, y, z) \in \mathbb{Z}^3 \mid x \leq y + p \vee x \leq z + p\}.$$

Many decision problems reduce in polynomial time to Max-Atom: this is for example the case of mean-payoff games [50] (which are in fact polynomial-time equivalent to Max-Atom [80]), parity games, and the model-checking problem for the modal μ-calculus [69]. The precise complexity of the Max-Atom problem is still unknown: it is known to be in $\mathsf{NP} \cap \mathsf{coNP}$, but not known to be in P. Note that if the constants p in the Max-Atom constraints are encoded in unary, then there is a simple reduction of the Max-Atom problem to a discrete temporal CSP (which is max-closed and with finite signature; also see Example 11, (6)): the max-atom constraint $x \leq \max(y, z) + p$ is equivalent to

$$\exists x_1, \ldots, x_p \big(x_1 = \mathrm{succ}(x) \wedge \cdots \wedge x_p = \mathrm{succ}(x_{p-1}) \wedge x_p \leq \max(y, z)\big).$$

The hardness proofs in this chapter can be used even for infinite-signature reducts $\mathcal{B}$ of $(\mathbb{Z}; <)$: for any structure $\mathcal{B}'$ obtained by keeping finitely many relations from $\mathcal{B}$, there is a trivial polynomial-time reduction from $\mathrm{CSP}(\mathcal{B}')$ to $\mathrm{CSP}(\mathcal{B})$. Hence, Theorem 6.5 implies that if $\mathcal{B}$ contains the successor relation, and if $\max_d, \min_d$, and si are not polymorphisms of $\mathcal{B}$ or $\mathbb{Q}.\mathcal{B}$, then $\mathrm{CSP}(\mathcal{B})$ is NP-hard. In particular, if $\mathcal{B}$ contains additionally the relation $<$, then $\mathrm{CSP}(\mathcal{B})$ is NP-hard unless $\mathcal{B}$ is preserved by max or min.

Chapter 7

Disjunctive Linear Diophantine Constraints

In this chapter, we continue investigating CSPs in the Jonsson-Lööw programme. The templates of interest here are structures that are first-order definable in $(\mathbb{Z}; +, 1)$ and that have $+$ in their signature. The CSP of $(\mathbb{Z}; +, 1)$ itself is the problem of solving linear Diophantine equations. This problem is in P, and the class that we propose to study here is the class of all extensions of linear Diophantine satisfiability. Note that although $(\mathbb{Z}; \mathrm{succ})$ is first-order definable in $(\mathbb{Z}; +, 1)$, and so is every countable unary structure with finite signature, the class of structures studied in this chapter has an empty intersection with the classes studied in Chapter 4 and Chapter 6, because the structures considered here have to contain $+$ in their signature.

In a related work [19], the authors studied the problem of classifying the complexity of the CSP of *semi-algebraic* constraint languages, that is, constraint languages whose relations are first-order definable in $(\mathbb{Q}; +, \times, <, 1)$. They obtained a complete classification in the case that the constraint language contains $+$, $\leq$, and 1. Moving from semi-algebraic to semi-linear constraint languages, whose relations are first-order definable over the structure $(\mathbb{Q}; +, <, 1)$, Jonsson and Thapper classified the complexity of constraint languages containing $+$ in their signature [65]. The complexity of first-order reducts of $(\mathbb{Q}; +, 1)$ containing $+$ has also been classified [18].

In these works, the class of relations quantifier-free definable in Horn conjunctive normal form plays a key role. In this context, the atomic relations are inequalities and equalities, and each clause may have no more than one equality or inequality. That is, additional disjuncts in clauses must be disequalities. For first-order expansions of $(\mathbb{Q}; +)$, the tractable constraint languages are precisely those that are quantifier-free Horn definable on $(\mathbb{Q}; +)$ [18]. However, the integers behave very differently from the rationals or reals and even simple types of Horn definitions engender intractable constraint languages, as documented in [64]. In this chapter, we show that the tractability border for first-order reducts of $(\mathbb{Z}; +, 1)$ containing $+$ coincides with that for first-order expansions of $(\mathbb{Q}; +)$.

Theorem 7.1. *Let $\mathcal{A}$ be an expansion of $(\mathbb{Z}; +)$ by relations with a first-order definition in $(\mathbb{Z}; +, 1)$. Then $\mathrm{CSP}(\mathcal{A})$ is in P or NP-complete. Moreover, if $\mathcal{A}$ is a core, then $\mathrm{CSP}(\mathcal{A})$ is in P if the relations of $\mathcal{A}$ are Horn-definable, and $\mathrm{CSP}(\mathcal{A})$ is NP-complete otherwise.*

7.1 Preliminaries

A *linear equation* is a formula of the form $\sum_{i=1}^{n} \kappa_i x_i = c$ with $\kappa_1, \ldots, \kappa_n, c \in \mathbb{Z}$, whose free variables are $\{x_1, \ldots, x_n\}$. A *modular linear equation* is a formula of the form $\sum_{i=1}^{n} \kappa_i x_i = c \bmod d$ with $\kappa_1, \ldots, \kappa_n, c, d \in \mathbb{Z}$. Let $\mathcal{L}_{(\mathbb{Z};+,1)}$ be the infinite relational language containing the linear equations and modular linear equations. More formally, for every linear equation or modular linear equation $\mathcal{L}_{(\mathbb{Z};+,1)}$ contains a corresponding relational symbol. For convenience, we consider first-order logic to have native symbols for $\top$ (true) and $\bot$ (false). It is well-known that $(\mathbb{Z}; +, 1)$ admits quantifier elimination in the language $\mathcal{L}_{(\mathbb{Z};+,1)}$ (see [83], or [77, Corollary 3.1.21] for a more modern treatment). Call an $\mathcal{L}_{(\mathbb{Z};+,1)}$-formula *standardized* if it does not contain a negated modular linear equation. Every $\mathcal{L}_{(\mathbb{Z};+,1)}$-formula is equivalent over $\mathbb{Z}$ to a standardized $\mathcal{L}_{(\mathbb{Z};+,1)}$-formula, since a negated modular linear equation is equivalent to a disjunction of modular linear equations (i.e., $k \neq b \bmod c \Leftrightarrow \bigvee_{0 \leq a < c, a \neq b} k = a \bmod c$). We say that an equation *appears* in a formula if it is a positive or negative literal in this formula.

Any subgroup Γ of $\mathbb{Z}^k$ can be given by a finite set of *generators*, i.e., k-tuples $\overline{g}^1, \ldots, \overline{g}^m$, such that for every $\overline{g} \in \Gamma$, there are $\lambda_1, \ldots, \lambda_m \in \mathbb{Z}$ such that $\overline{g} = \sum_i \kappa_i \overline{g}^i$, where we write $\lambda \cdot \overline{g}$ for $(\lambda g_1, \ldots, \lambda g_k)$. A *coset* of a subgroup Γ of $\mathbb{Z}^k$ is any set of the form $\overline{a} + \Gamma := \{\overline{a} + \overline{g} \mid \overline{g} \in \Gamma\}$, where $\overline{a} \in \mathbb{Z}^k$. By moving to a standardized formula, we are in a position to deduce the following.

Proposition 7.2. *Suppose R is a unary relation first-order definable in $(\mathbb{Z}; +, 1)$. Then R has the form $(R' \cup A) \setminus B$, where R' is a finite union of cosets of nontrivial subgroups of $\mathbb{Z}$, and A and B are finite disjoint sets of integers.*

Proof. Consider a disjunction ϕ of equations (possibly negated and modular equations). If this disjunction contains a negated equation $ax \neq c$, then ϕ defines a relation that contains $\mathbb{Z} \setminus \{c/a\}$ and is therefore as in the statement. Otherwise, ϕ contains only positive linear equation and modular equations, and the relation that ϕ defines is clearly of the form $R' \cup A$ for some finite set A and some union R' of nontrivial subgroups of $\mathbb{Z}$.

Consider a quantifier-free formula ϕ in conjunctive normal form defining R. Each conjunct defines a relation of the right form, per the previous paragraph. It is easily checked that a conjunction of relations of this form is again a relation of the form $(R' \cup A) \setminus B$, so that we have proved that every quantifier-free formula with one free variables defines a relation of the right form. The proposition then follows from quantifier-elimination. $\square$

If $\mathcal{A}$ is first-order definable in $(\mathbb{Z}; <, +, 1)$, then $\mathrm{CSP}(\mathcal{A})$ is in NP. This follows from quantifier elimination and is noted, inter alia, in [64].

Definition 7.1. Let ϕ be an $\mathcal{L}_{(\mathbb{Z};+,1)}$-formula. We say that ϕ is *Horn* if it is a conjunction of clauses of the form

$$\bigvee_{i=1}^{n} \neg \phi_i \vee \phi_0$$

where $\phi_1, \ldots, \phi_n$ are linear equations and ϕ_0 is a linear or a modular linear equation.

Example 16. Singletons, cofinite unary relations, and cosets of subgroups of $\mathbb{Z}^n$ are all examples of relations definable over $\mathcal{L}_{(\mathbb{Z};+,1)}$ that are Horn-definable.

7.2 Cores

If $\mathcal{A}$ is a first-order reduct of $(\mathbb{Z};+,1)$ containing $+$, note that its endomorphisms are precisely of the form $x \mapsto \lambda x$ for some $\lambda \in \mathbb{Z}$. Therefore, we view in the following $\mathrm{End}(\mathcal{A})$ as a subset of $\mathbb{Z}$, where the monoid structure on $\mathrm{End}(\mathcal{A})$ implies that as a subset of $\mathbb{Z}$, it is closed under multiplication and contains 1. Remember that a core is a structure whose endomorphisms are self-embeddings. Under the identification of $\mathrm{End}(\mathcal{A})$ with a subset of $\mathbb{Z}$, we obtain that $\mathcal{A}$ is a core if, and only if, whenever $\lambda, \lambda\mu \in \mathrm{End}(\mathcal{A})$ then $\mu \in \mathrm{End}(\mathcal{A})$. We say that $\mathcal{B}$ is a core of $\mathcal{A}$ if $\mathcal{A}$ and $\mathcal{B}$ are homomorphically equivalent and $\mathcal{B}$ is a core.

Lemma 7.3. *Let $\mathcal{A}$ be first-order definable in $(\mathbb{Z};+,1)$, and suppose that $\mathcal{A}$ contains $+$. There exists a structure which is a core of $\mathcal{A}$, and which is either a 1-element structure or first-order definable in $(\mathbb{Z};+,1)$ and containing $+$.*

Proof. If $0 \in \mathrm{End}(\mathcal{A})$ then the lemma is clearly true ($\mathcal{A}$ being homomorphically equivalent to the substructure of $\mathcal{A}$ induced by $\{0\}$), so let us assume that $0 \notin \mathrm{End}(\mathcal{A})$. Similarly we can assume that $\mathrm{End}(\mathcal{A}) \not\subseteq \{-1,1\}$, otherwise $\mathcal{A}$ is already a core. For a quantifier-free formula ψ and an integer λ, define ψ/λ by induction on ψ as follows:

- if ψ is $\sum \kappa_i x_i = c$ and λ divides c, then ψ/λ is $\sum \kappa_i x_i = c/\lambda$,
- if ψ is $\sum \kappa_i x_i = c$ and λ does not divide c, then ψ/λ is $\bot$,
- if ψ is $\sum \kappa_i x_i = c \bmod d$ and $\ell := \gcd(\lambda, d)$ divides c, ψ/λ is $\sum \kappa_i x_i = ec/\ell \bmod d/\ell$ where e is the inverse of λ/ℓ modulo d/ℓ,
- if ψ is $\sum \kappa_i x_i = c \bmod d$ and $\ell := \gcd(\lambda, d)$ does not divide c, ψ/λ is $\bot$,
- extend to boolean combinations in the obvious fashion.

Note that for every tuple $\overline{a}$, we have that $\overline{a}$ satisfies ψ/λ iff $\lambda \cdot \overline{a}$ satisfies ψ. Indeed, if ψ is a linear equation then this is clear. Similarly, it is clear if ψ is a modular equation and $\ell := \gcd(\lambda, d)$ does not divide c. Suppose that ψ is a modular equation and $\ell := \gcd(\lambda, d)$ divides c. If $\sum \lambda\kappa_i x_i = c \bmod d$ then $\lambda/\ell \cdot (\sum \kappa_i x_i) = qd/\ell + c/\ell$ so that $e\lambda/\ell \cdot (\sum \kappa_i x_i) = (eq) \cdot d/\ell + ec/\ell$, where e is the inverse of λ/ℓ modulo d/ℓ and $q \in \mathbb{Z}$. We therefore obtain $\sum \kappa_i x_i = ec/\ell \bmod d/\ell$. Conversely if $\sum \kappa_i x_i = ec/\ell \bmod d/\ell$ then $\sum \lambda\kappa_i x_i = (\lambda e)c/\ell + (\frac{\lambda}{\ell} q)d = c \bmod d$.

Let ψ be any quantifier-free $\mathcal{L}_{(\mathbb{Z};+,1)}$-formula and suppose that $|\lambda| > 1$. The only cases where some magnitudes of the integers on the right-hand sides of terms in the formula ψ do not decrease by forming ψ/λ is when ψ only contains literals either of the form $\sum \kappa_i x_i = 0$ or of the form $\sum \kappa_i x_i = c \bmod d$ with λ and d coprime. Therefore, the sequence $\psi_0, \psi_1, \psi_2, \ldots$ where ψ_0 is ψ and where ψ_{i+1} is ψ_i/λ for some $\lambda \in \mathrm{End}(\mathcal{A})$ with $|\lambda| > 1$ reaches in a finite number of steps a fixpoint where all the literals are either of the form $\sum \kappa_i x_i = 0$ or are modular equations whose modulus d is such that λ and d are coprime. Let $n \geq 1$ be such that for every ψ defining a relation of $\mathcal{A}$, the formula ψ_n is a fixpoint. Let $\mathcal{B}$ be the structure whose domain is $\mathbb{Z}$ and whose relations are $+$ and the relations defined by ψ_n for each ψ defining a relation of $\mathcal{A}$.

We claim that $\mathcal{B}$ is homomorphically equivalent to $\mathcal{A}$ and is a core. The first claim is clear, since $\mathcal{B}$ is isomorphic to the structure obtained from $\mathcal{A}$ by successive applications

of endomorphisms $x \mapsto \lambda \cdot x$ (in particular $\mathcal{B}$ embeds into $\mathcal{A}$). Let now $x \mapsto \lambda \cdot x$ be an endomorphism of $\mathcal{B}$, and suppose that $\overline{a}$ is a tuple in a relation R of $\mathcal{B}$. Then we have that $\lambda \cdot \overline{a}$ in R since $x \mapsto \lambda \cdot x$ is an endomorphism. Conversely, note that λ is coprime to d or else we would not have reached a fixed point in the previous stage. Thus, $\lambda^{\phi(d)} = 1 \bmod d$, where $\phi(d)$ here is the totient of d. It follows then that $\lambda^{\phi(d)}\overline{a} = \overline{a} \bmod d$. Suppose $\lambda\overline{a} \in R$, then by applying $\phi(d) - 1$ times an endomorphism, we derive $\lambda^{\phi(d)}\overline{a} \in R$. It follows that $\overline{a} \in R$, for both the cases that atoms are of the form $\sum \kappa_i x_i = 0$ or are modular equations whose modulus d is such that λ and d are coprime. Hence, $x \mapsto \lambda \cdot x$ is an embedding of $\mathcal{A}$. □

Define an order on standardized formulas using the lexicographic ordering on:

1. number of non-Horn clauses,
2. number of literals,
3. sum of the absolute value of all numbers appearing in an equation.

A formula is *minimal* if no smaller formula is equivalent. The following properties follow from the construction of a core in the previous proof.

Proposition 7.4. *Let $\mathcal{A}$ be first-order definable in $(\mathbb{Z}; +, 1)$, and suppose that $\mathcal{A}$ contains $+$ and is a core. Let $\lambda \in \mathrm{End}(\mathcal{A})$. Let R be a relation of $\mathcal{A}$ and let ϕ be a standardized minimal formula defining R.*

- *If $\sum \kappa_i x_i = c$ is a linear equation appearing in ϕ, then either $c = 0$ or $|\lambda| = 1$.*
- *If $\sum \kappa_i x_i = c \bmod d$ is a modular linear equation in ϕ, then λ and d are coprime.*

Moreover, if $\mathrm{End}(\mathcal{A}) = 1 + d\mathbb{Z}$ for some $d \geq 2$, then every relation of $\mathcal{A}$ can be expressed with a minimal formula in which all modular linear equations are divisors of d.

Proof. The two items are clear from the proof of Lemma 7.3, as the syntactic manipulations in the proof decrease the size of the formula. For the last statement, let d' be a modulus appearing in a minimal definition of a relation of $\mathcal{A}$. By the second item, we have that d' and $1 + kd$ are coprime, for all $k \in \mathbb{Z}$. Let ℓ be such that $\ell d = -1 \bmod \frac{d'}{\gcd(d,d')}$. If d' and $1 + \ell d$ are coprime, there exist $u, v \in \mathbb{Z}$ such that $ud' + v(1 + \ell d) = 1$. Taking this equation modulo $\frac{d'}{\gcd(d,d')}$, we obtain that $0 = 1 \bmod \frac{d'}{\gcd(d,d')}$, so that $\gcd(d, d') = d'$ and d' divides d. □

7.3 Hardness

Our sources of hardness come from *one-dimensional pp-interpretations*, that we define now. A structure $\mathcal{B}$ is said to be *one-dimensional pp-interpretable* in $\mathcal{A}$ if there exists a partial map $h\colon A \to B$ such that the inverse image of every relation of $\mathcal{B}$ (including the equality relation and the unary relation B) under h has a pp-definition in A. Formally, we require that for every k-ary relation R of $\mathcal{B}$, there exists a pp-formula $\phi_R(x_1, \ldots, x_k)$ in the language of $\mathcal{A}$ such that

$$\mathcal{A} \models \phi_R(a_1, \ldots, a_k) \Leftrightarrow \mathcal{B} \models R(h(a_1), \cdots, h(a_k))$$

holds for all $a_1, \ldots, a_k \in A$. The constraint on the equality relation of $\mathcal{B}$ and the unary relation B force that the kernel of h and its domain have a pp-definition in $\mathcal{A}$. One sees that one-dimensional pp-interpretations are a particular case of pp-constructions, so that from Lemma 2.3 we obtain that if $\mathcal{B}$ is pp-interpretable in $\mathcal{A}$, then $\mathrm{CSP}(\mathcal{B})$ reduces in polynomial time to $\mathrm{CSP}(\mathcal{A})$.

7.3.1 The fully modular case

One of the sources of hardness for our problems are expansions of the *general subgroup problem* from [52]. The general subgroup problem of a finite abelian group Γ is the CSP of $(\Gamma; +)$ expanded with a k-ary relation for every coset $\overline{a} + \Delta$, where Δ is a subgroup of Γ^k. It is known that this problem is solvable in polynomial time (under some reasonable encoding of the input); in modern parlance, this follows from the fact that the operation $(x, y, z) \mapsto x - y + z$ is a Maltsev polymorphism of the template. Feder and Vardi [52, Theorem 34] proved that the problem becomes NP-hard if the template is further expanded by any other relation.

The general subgroup problem of $\mathbb{Z}/d\mathbb{Z}$ can be viewed as a CSP of a first-order reduct of $(\mathbb{Z}; +, 1)$ whose relations are defined by quantifier-free formulas only containing modular linear equations. This motivates the following definition.

Definition 7.2. A relation $R \subseteq \mathbb{Z}^k$ is called *fully modular* if it is definable by a conjunction of disjunctions of modular linear equations, in which case we can even assume that all the modular linear equations involved in such a definition of R have the same modulus $d \geq 1$.

Proposition 7.5. *Let $\mathcal{A}$ be finite-signature first-order definable in $(\mathbb{Z}; +, 1)$ and containing $+$. Suppose that $\mathcal{A}$ has a fully modular relation that is not Horn-definable. Then* $\mathrm{CSP}(\mathcal{A})$ *is NP-complete.*

Proof. Let R be a relation of $\mathcal{A}$ that is not Horn-definable and fully modular, and let $d \geq 1$ be such that R can be defined with only linear equalities modulo d. Let $\mathcal{A}/d\mathcal{A}$ be the structure $\mathcal{A}/d\mathcal{A}$ with domain $\mathbb{Z}/d\mathbb{Z}$ and with relations $+$ as well as for every relation S of arity k of $\mathcal{A}$, a relation S' defined by

$$S' = \{(a_1, \ldots, a_k) \mid \exists q \in \mathbb{Z} : (qd + a_1, \ldots, qd + a_k) \in S'\}.$$

Note that $\mathcal{A}/d\mathcal{A}$ is pp-interpretable in $\mathcal{A}$: the map h is the canonical projection $x \mapsto x \bmod d$, whose kernel is pp-definable by the formula $\phi_=(x, y) := \exists z(x - y = dz)$. As a consequence, $\mathrm{CSP}(\mathcal{A}/d\mathcal{A})$ reduces in logarithmic space to $\mathrm{CSP}(\mathcal{A})$. Moreover, if $\mathcal{A}$ is a core then $\mathcal{A}/d\mathcal{A}$ is also a core. It follows from Proposition 2.20 that $\mathrm{CSP}(\mathcal{A}/d\mathcal{A}, 1)$ reduces to $\mathrm{CSP}(\mathcal{A}/d\mathcal{A})$ and so to $\mathrm{CSP}(\mathcal{A})$. Note that every coset of a subgroup of $\mathbb{Z}/d\mathbb{Z}^k$ is pp-definable in $(\mathcal{A}/d\mathcal{A}, 1)$ and that if R is not Horn-definable then R' is not a coset of a subgroup. It follows from Theorem 34 in the bible [52] that $\mathrm{CSP}(\mathcal{A})$ is NP-complete. $\square$

7.3.2 The unary case

In order to prove Theorem 7.1, we now focus on the case of *parametrised unary relations*. For a unary relation $R \subseteq \mathbb{Z}$ that is definable in $(\mathbb{Z}; +, 1)$, we write R° for the unique union of cosets of nontrivial subgroups of $\mathbb{Z}$ such that R can be decomposed as $R = (R^\circ \cup A) \setminus B$ where A, B are finite and disjoint, $A \cap R^\circ = \emptyset$ and $B \subset R^\circ$.

Definition 7.3. Let $\mathcal{A}$ be a first-order reduct of $(\mathbb{Z};+,1)$ and $\Lambda \subseteq \mathbb{Z}$ be a set containing 1 and such that $0 \notin \Lambda$. Let $\{S_\lambda\}_{\lambda\in\Lambda}$ be a set of unary relations $S_\lambda \subseteq \mathbb{Z}$. We say that the unary relations are *compatible* if there exist disjoint finite sets $A, B \subseteq \mathbb{Z}$ such that $S_\lambda = (S_\lambda^\circ \cup \lambda \cdot A) \setminus \lambda \cdot B$ for all $\lambda \in \Lambda$ and such that for all $d \geq 1$ and $c \in \{0, \cdots, d-1\}$, we have $c + d\mathbb{Z} \subseteq S_1^\circ \Leftrightarrow \lambda c + d\mathbb{Z} \subseteq S_\lambda^\circ$. We say that $\{S_\lambda\}_{\lambda\in\Lambda}$ is *uniformly pp-definable in* $\mathcal{A}$ if there exists a pp-formula $\theta(x,y)$ such that $a \in S_\lambda$ if, and only if, $\mathcal{A} \models \theta(\lambda, a)$.

Note that the definition of being uniformly pp-definable implies that Λ has a pp-definition in $\mathcal{A}$, for $\exists y(\theta(x,y))$ is a pp-definition. Given a binary relation $S \subseteq \mathbb{Z}^2$ that is pp-definable in $\mathcal{A}$, we write S_λ for the set $\{a \in \mathbb{Z} \mid (\lambda, a) \in S\}$, and the family $\{S_\lambda\}_{\lambda\in\Lambda}$ with $\Lambda = \mathbb{Z} \setminus \{0\}$ is uniformly pp-definable in $\mathcal{A}$ (but does not necessarily satisfy the compatibility condition). For example, let $S = \{(a,b) \in \mathbb{Z}^2 \mid a = b \vee a = 2b\}$. Then we have $S_\lambda = \{\lambda\}$ for $\lambda = 1 \bmod 2$ and $S_\lambda = \{\lambda, \frac{\lambda}{2}\}$ for $\lambda = 0 \bmod 2$. Therefore, the compatibility condition is not satisfied by $\{S_\lambda\}_{\lambda\in\mathbb{Z}\setminus\{0\}}$.

Lemma 7.6. *Let $\mathcal{A}$ be a first-order reduct of $(\mathbb{Z};+,1)$ containing $+$. If $\{S_\lambda\}_{\lambda\in\Lambda}$ is a compatible set of unary relations that is uniformly pp-definable in $\mathcal{A}$ and such that $1 < |S_\lambda| < \infty$ for all $\lambda \in \Lambda$, then* $\mathrm{CSP}(\mathcal{A})$ *is* NP*-hard.*

Proof. Since every S_λ is finite, one sees that $S_\lambda = \lambda \cdot A$ for the finite set A coming from the compatibility condition. Let $m_1 := \min(A)$ and $m_2 := \min(A \setminus \{m_1\})$. The formula

$$\exists\lambda(x + y + z = (m_2 - m_1)\lambda \wedge x + m_1\lambda \in S_\lambda \wedge y + m_1\lambda \in S_\lambda \wedge z + m_1\lambda \in S_\lambda \wedge \lambda \in \Lambda)$$

defines the ternary relation consisting of $(a,b,c) \in \mathbb{Z}^3$ such that $a,b,c \in \{0, m_2 - m_1\}$ and exactly one of a,b,c is equal to $m_2 - m_1$. This gives an interpretation of 1-IN-3-SAT in $\mathcal{A}$, using the map $h\colon \{0, m_2 - m_1\} \to \{0,1\}$ such that $h(0) = 0$ and $h(m_2 - m_1) = 1$. Therefore $\mathrm{CSP}(\mathcal{A})$ is NP-hard. □

Proposition 7.7. *Let $\mathcal{A}$ be a core that is first-order definable in $(\mathbb{Z};+,1)$ and contains $+$. Let $\{S_\lambda\}_{\lambda\in\Lambda}$ be a compatible family of unary relations that is uniformly pp-definable in $\mathcal{A}$. If S_λ is not Horn-definable for all $\lambda \in \Lambda$, then* $\mathrm{CSP}(\mathcal{A})$ *is* NP*-hard.*

Proof. Let $A, B \subset \mathbb{Z}$ be finite and such that $S_\lambda = (S_\lambda^\circ \cup \lambda \cdot A) \setminus (\lambda \cdot B)$. If S_λ is finite for all $\lambda \in \Lambda$, then $\mathrm{CSP}(\mathcal{A})$ is NP-hard by Lemma 7.6. Therefore, we can assume that $S_\lambda^\circ \neq \emptyset$ for all $\lambda \in \Lambda$, and let $d \geq 1$ be such that S_λ° is a union of cosets of $d\mathbb{Z}$ for all $\lambda \in \Lambda$. Write $S_1^\circ = \bigcup_{i=1}^n c_i + d\mathbb{Z}$, with $c_i \in \{0, \ldots, d-1\}$.

If $n \in \{2, \ldots, d-1\}$, we claim that we can define a fully modular relation that is not Horn-definable. Indeed, let $\theta(x,y)$ be a formula that defines $\{S_\lambda\}_{\lambda\in\Lambda}$. Note that $\chi(x,y) := \theta(x,y) \wedge \theta(x, y+dx) \wedge \cdots \wedge \theta(x, y + \max(A \cup B)dx)$ holds precisely on the pairs (λ, a) such that $a \in S_\lambda^\circ$: since x is forced to be in Λ by θ, a satisfying assignment gives a nonzero value $\lambda$to x. Thus, if all of $y, y+d\lambda, \cdots, y+\max(A\cup B)d\lambda$ are in S_λ, then they all must be in the modular part S_λ°. The relation T that χ defines is fully modular and is such that $T_\lambda = S_\lambda^\circ$ and in particular T is not Horn-definable. It follows from Proposition 7.5 that $\mathrm{CSP}(\mathcal{A})$ is NP-hard.

Otherwise, the set S_λ° consists of a single coset of $d\mathbb{Z}$ for all $\lambda \in \Lambda$, and this coset is $\lambda c_1 + d\mathbb{Z}$ by the compatibility condition on $\{S_\lambda\}_{\lambda\in\Lambda}$. Since S_λ° is assumed to not be Horn-definable, then $A \neq \emptyset$. Let $a \in A$. We claim that we can define another family of

unary relations, where the unary relations are finite and not singletons. Indeed, consider the formula

$$\psi(x, y) := \exists z\, (\theta(x, y) \wedge \theta(x, z) \wedge y + z = (c_1 + a)x)$$

and let $T \subseteq \mathbb{Z}^2$ be the relation that it defines. First, note that $\psi(\lambda, c_1)$ and $\psi(\lambda, a)$ hold for all $\lambda \in \Lambda$, so that $|T_\lambda| > 1$. We claim that T_λ is finite. Since $A \cap S_1^\circ = \emptyset$, one has $a \neq c_1 \bmod d$. Consequently, $c_1 + a \neq 2c_1 \bmod d$ and $(c_1 + a)\lambda \neq 2c_1\lambda \bmod d$. The equation $y + z = (c_1 + a)\lambda$ therefore forces that one of y and z is in $\lambda \cdot A$. Since A is finite, there are only finitely many pairs satisfying this condition, thus showing that $1 < |T_\lambda| < \infty$. It follows from Lemma 7.6 that $\mathrm{CSP}(\mathcal{A})$ is NP-hard. □

A nice corollary of the hardness result for unary relations is the necessary condition on $\mathrm{End}(\mathcal{A})$ for $\mathrm{CSP}(\mathcal{A})$ to not be NP-hard that we give in the sequel (Corollary 7.10).

Lemma 7.8. *Let $\mathcal{A}$ be a finite-signature structure first-order definable in $(\mathbb{Z}; +, 1)$ that contains $+$. The set $\mathrm{End}(\mathcal{A})$ is pp-definable in $\mathcal{A}$.*

Proof. Let E be the set of all the formulas $R(a_1 \cdot x, \ldots, a_r \cdot x)$ for R in the language of $\mathcal{A}$ and $(a_1, \ldots, a_r) \in R$. We then have that $\mathcal{A} \models E(\lambda)$ iff $\lambda \in \mathrm{End}(\mathcal{A})$. We now show that there exists a finite subset $F \subseteq E$ that defines the same set of integers.

For each relation R of $\mathcal{A}$, fix a standardized definition ϕ_R in conjunctive normal form of R in $(\mathbb{Z}; +, 1)$. Let M be the largest absolute value of a constant appearing in ϕ_R. Consider the finite family $\mathcal{F}$ of equations $\sum \mu_i x_i = m$, where $\sum \mu_i x_i = m'$ is some equation in ϕ_R and $|m| \leq M$, together with all the equations $\sum \mu_i x_i = c \bmod d$ where $\sum \mu_i x_i = c' \bmod d$ is a modular equation in ϕ_R and $c \in \{0, \ldots, d-1\}$. For each subset of $\mathcal{F}$ that is satisfiable by a tuple in R, pick a tuple $\overline{b} \in R$ satisfying the formulas in this subset and add this tuple to a set $\mathcal{S}$. Repeat this operation for every relation of $\mathcal{A}$, and let $\mathcal{S}$ be the finite set of tuples (of possibly different arities) that we obtain. Finally, let F be the subset of E where only the formulas associated with tuples from $\mathcal{S}$ are kept.

We claim that F defines $\mathrm{End}(\mathcal{A})$. Since $F \subseteq E$, it suffices to show that every λ satisfying F is an endomorphism of $\mathrm{End}(\mathcal{A})$. Let $\lambda \in \mathbb{Z}$ satisfy F, and let $\overline{a} \in R$ be a tuple in some relation of $\mathcal{A}$. Let $\overline{b} \in \mathcal{S}$ be such that $\overline{b}$ satisfies exactly the same equations in $\mathcal{F}$ as $\overline{a}$. By construction, $\lambda\overline{b} \in R$ so that in each clause of ϕ_R, some equation is satisfied by $\lambda\overline{b}$. We show that $\lambda\overline{a}$ satisfies the same equations, so that $\lambda\overline{a} \in R$. If $\lambda = 0$, then $\lambda\overline{b} = \lambda\overline{a}$ so that $\lambda\overline{a} \in R$. Suppose now that $\lambda \neq 0$. Let $\sum \mu_i x_i = c$ be a linear equation that is satisfied by $\lambda\overline{b}$. Then necessarily λ divides c, so that $\overline{b}$ satisfies $\sum \mu_i x_i = \frac{c}{\lambda}$ and $|\frac{c}{\lambda}| \leq |c| \leq M$, so that $\sum \mu_i x_i = \frac{c}{\lambda}$ is an equation in $\mathcal{F}$. Consequently, $\overline{a}$ also satisfies this equation and $\lambda\overline{a}$ satisfies $\sum \mu_i x_i = c$. The proof for modular linear equations is similar. This proves that λ is an endomorphism of $\mathcal{A}$ and concludes the proof. □

Lemma 7.9. *Suppose that $(\mathbb{Z}; +, R)$ is a core, with $R \subseteq \mathbb{Z}$ being definable over $(\mathbb{Z}; +, 1)$. Write $R = (R' \cup A) \setminus B$ with R' being a union of cosets of a nontrivial subgroup of $\mathbb{Z}$ and A, B being disjoint minimal finite subsets of $\mathbb{Z}$.*

- *Suppose that $A \neq \emptyset$. Then $\{1\}$ or $\{1, -1\}$ is pp-definable in $(\mathbb{Z}; +, R)$.*
- *If $A = \emptyset$, we have $B = \emptyset$ or $R = \mathbb{Z} \setminus \{0\}$.*

Proof. Let n be such that R' is a union of n cosets of $d\mathbb{Z}$, i.e.,

$$R' = \bigcup_{i=1}^{n} c_i + d\mathbb{Z}.$$

Let us prove the first item. By Lemma 7.8, it suffices to prove that the only possible endomorphisms of the structure $(\mathbb{Z}; +, R)$ are $x \mapsto \lambda \cdot x$ with $\lambda \in \{1, -1\}$. Suppose that $x \mapsto \lambda \cdot x$ is an endomorphism. Then $\lambda \neq 0$ since the structure is a core, so suppose that $|\lambda| > 1$. Let a be the maximal element of A, and note that in particular $a+d \notin R$ (it cannot be in A because of the maximality assumption, and cannot be equal to any c_i modulo d). Then $a \in R$, so $\lambda^q a \in R$ for all $q \in \mathbb{N}$. In particular, if q is such that $\lambda^q > \max(A \cup B)$ we obtain $\lambda^q a \in R'$. This means that $\lambda^q a = c_i \bmod d$ for some $i \in \{1, \dots, n\}$. Finally, $\lambda^q(a + d) = \lambda^q a + \lambda^q d = c_i \bmod d$, so that $\lambda^q(a + d) \in R$. This implies that $x \mapsto \lambda^q \cdot x$ is not an embedding, contradicting the core assumption on $(\mathbb{Z}; +, R)$.

Let us now prove the second item. Let b be some element of B. We must have $b = c_i \bmod d$ for some $i \in \{1, \dots, n\}$, by the minimality assumption on B (otherwise $B \setminus \{b\}$ could be used to define the same relation). Note that the map $x \mapsto (d + 1)x$ is an endomorphism of $(\mathbb{Z}; +, R)$, so it has to be an embedding. It follows that $(d + 1)^m \cdot b \notin R$ for any m. Suppose that b is not 0. Choose m so that $(d + 1)^m \cdot |b| > \max_{e \in B} |e|$ so that $(d + 1)^m \cdot b \notin B$. But $(d + 1)^m b = c_i \bmod d$, a contradiction. It follows that $B \subseteq \{0\}$, which concludes the proof. □

Corollary 7.10. *Let $\mathcal{A}$ be finite-signature first-order definable in $(\mathbb{Z}; +, 1)$ and contain $+$. Suppose that $\mathcal{A}$ is a core. If $\mathrm{End}(\mathcal{A})$ is not Horn-definable, then $\mathrm{CSP}(\mathcal{A})$ is NP-hard. Moreover, if $\mathrm{End}(\mathcal{A})$ is Horn-definable, then it is either $\{1\}$, $\mathbb{Z} \setminus \{0\}$, or $1 + d\mathbb{Z}$ for some $d \geq 2$.*

Proof. One simply has to note that $(\mathbb{Z}; +, \mathrm{End}(\mathcal{A}))$ is a core. Indeed, let λ be an endomorphism of $(\mathbb{Z}; +, \mathrm{End}(\mathcal{A}))$. Since $1 \in \mathrm{End}(\mathcal{A}))$, we obtain that $\lambda \in \mathrm{End}(\mathcal{A})$, so that $x \mapsto \lambda x$ is a self-embedding of $\mathcal{A}$ by the fact that $\mathcal{A}$ is a core. Since $\mathrm{End}(\mathcal{A})$ is definable over $\mathcal{A}$ without quantifiers, it follows that $x \mapsto \lambda x$ is also a self-embedding of $(\mathbb{Z}; +, \mathrm{End}(\mathcal{A}))$. Te first part of the result then follows from Lemma 7.8 and Proposition 7.7. The second part follows from Lemma 7.9. □

7.3.3 Arbitrary arities

We finally present here the proof of hardness in the general case. The strategy is to cut from a non-Horn relation R a uniformly definable family $\{S_\lambda\}_{\lambda \in \Lambda}$ of lines for which each S_λ is not Horn-definable. In a second step, we ensure that we get a family satisfying the compatibility condition, and we conclude using Proposition 7.7. Call a formula ϕ in conjunctive normal form *reduced* if removing any literal or clause from ϕ yields a formula that is not equivalent to ϕ. Note that minimal formulas are necessarily reduced.

Lemma 7.11. *Let $\mathcal{A}$ be a reduct of $(\mathbb{Z}; +, 1)$ containing $+$ and such that $\mathcal{A}$ is a core. Suppose that $\mathcal{A}$ contains a relation that is not Horn-definable. Then $\mathrm{CSP}(\mathcal{A})$ is NP-hard, or $\mathcal{A}$ pp-defines a relation that is not Horn-definable and that has a minimal definition containing a non-Horn clause ψ such that:*

- *no negated linear equation is in ψ,*

- *at least one linear equation is in ψ.*

Proof. Let R be an n-ary relation of $\mathcal{A}$ that is not Horn-definable, let ϕ be a minimal formula defining R in conjunctive normal form, and let ψ be a clause of ϕ that is not Horn. From Corollary 7.10, we can suppose that $\mathrm{End}(\mathcal{A})$ is $\{1\}$, $\mathbb{Z} \setminus \{0\}$, or $1 + d\mathbb{Z}$ for $d \geq 2$. This implies that either $\{1\}$ is pp-definable or, by Proposition 7.4, all the linear equations appearing in ϕ are homogeneous. As a consequence, we can assume that ψ does not contain any negative literal, per the assumption that ϕ is reduced: indeed, consider the relation R' defined by

$$\phi \wedge \sum \kappa_i x_i = c \tag{†}$$

where $\sum \kappa_i x_i \neq c$ is a negated equation in ψ; either $c = 0$, in which case (†) is in the language of $\mathcal{A}$, or $\{1\}$ is pp-definable in $\mathcal{A}$ and (†) is also in the language of $\mathcal{A}$. This new relation R' is not Horn-definable and is pp-definable in $\mathcal{A}$, so it suffices to prove hardness for this relation.

Suppose now that ψ only contains modular linear equations. By Proposition 7.4, $\mathrm{End}(\mathcal{A})$ is not $\mathbb{Z} \setminus \{0\}$ (any modulus in a modular linear equation would have to be coprime with every nonzero integer!). Therefore, $\mathrm{End}(\mathcal{A})$ is $\{1\}$ or $1 + d\mathbb{Z}$ for $d \geq 2$. In the latter case, we can assume by Proposition 7.4 that all the modular linear equations in ψ are modulo a divisor of d. In the former case, let d be a common multiple of all the moduli appearing in a modular linear equation in ψ. Consider the structure $\mathcal{A}/d\mathcal{A}$ defined in Proposition 7.5. The relation T obtained from R in this structure is not a coset of a subgroup $\mathbb{Z}/d'\mathbb{Z}$ of $\mathbb{Z}/d\mathbb{Z}$ (otherwise one could replace ψ in ϕ by a Horn clause defining the corresponding coset modulo $d'\mathbb{Z}$ and obtain a smaller formula defining R, a contradiction to the minimality of ϕ). Moreover, $(\mathcal{A}/d\mathcal{A}, 1)$ is pp-interpretable in $\mathcal{A}$: in the two cases that $\mathrm{End}(\mathcal{A}) = \{1\}$ and $\mathrm{End}(\mathcal{A}) = 1 + d\mathbb{Z}$, the preimage of $\{1\}$ under the canonical projection is pp-definable in $\mathcal{A}$. We conclude as in Proposition 7.5 that $\mathrm{CSP}(\mathcal{A})$ is NP-hard.

Therefore, in the remaining case, ψ contains at least one linear equation and no negated linear equation, as required. □

Theorem 7.12. *Let $\mathcal{A}$ be a reduct of $(\mathbb{Z}; +, 1)$ containing $+$ and such that $\mathcal{A}$ is a core. Suppose that $\mathcal{A}$ contains a relation that is not Horn-definable. Then $\mathrm{CSP}(\mathcal{A})$ is NP-hard.*

Proof. From Lemma 7.11, we can suppose that $\mathcal{A}$ pp-defines a relation R that is not Horn-definable, that has a reduced standardized definition ϕ containing a non-Horn clause ψ with at least one linear equation (L) and no negated linear equation. Since ψ is not Horn, it contains at least another equation (L'), possibly modular. Let $(a_1, \ldots, a_n)$ satisfy ϕ and only (L) in ψ. Such a tuple exists by the assumption that ϕ is reduced. Similarly, let $(b_1, \ldots, b_n)$ satisfy ϕ and only (L') in ψ. Let S be the binary relation such that $(\lambda, t) \in S$ if, and only if, $\lambda \in \mathrm{End}(\mathcal{A})$ and $t(\overline{a} - \overline{b}) + \lambda\overline{b}$ is in R. Note that $\overline{a}$ and $\overline{b}$ being fixed, S is pp-definable over $\mathcal{A}$. Therefore, we obtain a family $\{S_\lambda\}_{\lambda \in \Lambda}$ that is uniformly definable in $\mathcal{A}$, where $\Lambda = \mathrm{End}(\mathcal{A})$. It is clear that $0, \lambda \in S_\lambda$. Moreover, note that

$$S_\lambda \cap \lambda \cdot \mathbb{Z} = \lambda \cdot S_1 \tag{‡}$$

holds for all $\lambda \in \mathrm{End}(\mathcal{A})$. Indeed:

$$\begin{aligned} t \in S_1 &\Leftrightarrow t(\overline{a}-\overline{b})+\overline{b} \in R \\ &\Leftrightarrow \lambda t(\overline{a}-\overline{b})+\lambda\overline{b} \in R && \text{because } \mathcal{A} \text{ is a core} \\ &\Leftrightarrow \lambda t \in S_\lambda. \end{aligned}$$

We prove that for all $\lambda \in \mathrm{End}(\mathcal{A})$, S_λ is not Horn-definable. Since $0, \lambda \in S_\lambda$, it suffices to prove that S_λ omits infinitely many multiples of λ, and by (‡) it suffices to prove that $\ell \notin S_1$ for infinitely many ℓ. Let ℓ be such that $\ell = 1 \bmod d'$, for every modulus d' appearing in ψ. It is clear that $\ell(\overline{a}-\overline{b})+\overline{b}$ does not satisfy any modular linear equation in ψ. Indeed, let $\sum_i \sigma_i x_i = c \bmod d'$ be such a modular linear equation. Then we have

$$\sum_i \sigma_i(\ell(a_i-b_i)+b_i) = c \bmod d' \Leftrightarrow \sum_i \sigma_i a_i = c \bmod d',$$

which is a contradiction to the choice of $\overline{a}$ since $\overline{a}$ only satisfies (L) in ψ and (L) is assumed to be non-modular. Consider now a linear equation $\sum \sigma_i x_i = c$ in ψ. This equation is satisfied by $\ell(\overline{a}-\overline{b})+\overline{b}$ if, and only if, the equation

$$\ell \cdot \sum \kappa_i(a_i-b_i) = c - \sum \kappa_i b_i \qquad (\star)$$

holds. Suppose first that $\sum \kappa_i(a_i-b_i) = 0$. Then $(\star)$ is satisfied if, and only if, we have $\sum \kappa_i b_i = c = \sum \kappa_i a_i$. This implies that both $\overline{a}$ and $\overline{b}$ satisfy the equation; this is a contradiction to our choice of the vectors $\overline{a}$ and $\overline{b}$, so that $\ell(\overline{a}-\overline{b})+\overline{b}$ does not satisfy $(\star)$. Suppose now that $\sum \kappa_i(a_i-b_i) \neq 0$. If $\ell > |c - \sum \kappa_i b_i|$, it is then clear that $(\star)$ is not satisfied. Therefore, for infinitely many ℓ, the tuple $\ell(\overline{a}-\overline{b})+\overline{b}$ does not satisfy any literal in ψ and $\ell \notin S_1$.

Let $\theta(x,y)$ be a reduced standardized definition of S of minimal size. By Proposition 7.2, one can decompose S_λ as $(S_\lambda^\circ \cup A_\lambda) \setminus B_\lambda$, where S_λ° is a union of cosets of a subgroup of $\mathbb{Z}$ (the cosets and the subgroup depend on λ). By inspection of the formula θ, one finds that A_λ and B_λ consist of points of the form $-\frac{a\lambda}{b}$, where $ax+by=0$ is an equation in θ. Note that since a and b are taken coprime by minimality of θ, if b divides $a\lambda$ then b divides λ. Let $m := \mathrm{lcm}\{|b| : ax+by=0 \text{ is an equation in } \theta\}$, and note that $m\mathbb{Z} \cap K$ is not empty by the previous remark. Let $T \neq \emptyset$ be the binary relation defined by $\theta(m \cdot x, y)$ and as above write $T_\lambda \subseteq \mathbb{Z}$ for the projections associated with T (i.e., $T_\lambda = S_{m\lambda}$ for all $\lambda \in \mathrm{End}(\mathcal{A})$). It is clear that T_λ is not Horn-definable and is of the form $(T_\lambda^\circ \cup (\lambda \cdot P)) \setminus (\lambda \cdot Q)$, where P and Q are finite sets that are independent of λ and $T_\lambda^\circ = S_{m\lambda}^\circ$. For the family of relations $\{T_\lambda\}_{\lambda \in \mathrm{End}(\mathcal{A})}$ to satisfy the compatibility condition, it remains to prove that $c + d\mathbb{Z} \subseteq T_1^\circ$ if, and only if, $\lambda c + d\mathbb{Z} \subseteq T_\lambda^\circ$, for all $d \geq 1$ and $c \in \{0, \ldots, d-1\}$. Suppose that $c + d\mathbb{Z} \subseteq T_1^\circ$ for some $d \geq 1$ and the cosets of T_λ° are cosets of $d'\mathbb{Z}$. By Proposition 7.4, λ and d' are coprime. Therefore, there exists $\mu \in \mathbb{Z}$ such that $\lambda\mu = 1 \bmod d'$. We have $c + d\mu\mathbb{Z} \subseteq T_1^\circ$, simply because d divides $d\mu$. It follows that $\lambda c + \lambda d\mu\mathbb{Z} \subseteq T_\lambda^\circ$. Now, let $x \in \lambda c + d\mathbb{Z}$, say $x = \lambda c + qd$. Then we have $x = \lambda c + q\lambda\mu d \bmod d'$. Note that $\lambda c + q\lambda\mu d \in \lambda c + \lambda d\mu\mathbb{Z} \subseteq T_\lambda^\circ$. Since the cosets in T_λ° are cosets of $d'\mathbb{Z}$, we obtain that $x \in T_\lambda^\circ$ and consequently that $\lambda c + d\mathbb{Z} \subseteq T_\lambda^\circ$. Conversely, if $\lambda c + d\mathbb{Z} \subseteq T_\lambda^\circ$ then $\lambda c + d\lambda\mathbb{Z} \subseteq T_\lambda^\circ$ (because d divides $d\lambda$), and since $\mathcal{A}$ is a core we have that $x \mapsto \lambda \cdot x$ is a self-embedding of $\mathcal{A}$, so that $c + d\mathbb{Z} \subseteq T_1^\circ$.

To conclude, the family of compatible relations $\{T_\lambda\}_{\lambda \in \mathrm{End}(\mathcal{A})}$ is uniformly pp-definable in $\mathcal{A}$ and consists of non-Horn relations. By Proposition 7.7, we obtain that $\mathrm{CSP}(\mathcal{A})$ is NP-hard. □

We illustrate our proofs in some examples below.

Example 17. Consider the binary relation

$$S = \{(\lambda, t) \mid (\lambda = 1 \bmod 4 \wedge t = 1 \bmod 4) \vee (\lambda = 3 \bmod 4 \wedge t = 3 \bmod 4) \vee t = 0\}.$$

One sees that the set of endomorphisms of $\mathcal{A} := (\mathbb{Z}; +, S)$ is equal to $\mathrm{End}(\mathcal{A}) := 1 + 2\mathbb{Z}$. Moreover, S is not Horn-definable. For every $\lambda \in \mathrm{End}(\mathcal{A})$, one has $S_\lambda = \{0\} \cup (\lambda + 4\mathbb{Z})$. When λ is fixed, one can define a finite set by $\exists y(x \in S_\lambda \wedge y \in S_\lambda \wedge x + y = \lambda)$, which defines $\{0, \lambda\}$. One then obtains a reduction from 1-IN-3-SAT by $\exists x, y, z \in \{0, \lambda\} : x + y + z \in \{0, \lambda\}$. Finally, by existentially quantifying over $\lambda \in \mathrm{End}(\mathcal{A})$ we obtain a reduction from 1-IN-3-SAT to $\mathrm{CSP}(\mathbb{Z}; +, S)$.

Example 18. Let $R := \{0\} \cup (1 + 3\mathbb{Z}) \cup (2 + 3\mathbb{Z})$ and $K = 1 + 3\mathbb{Z}$. Note that Proposition 7.7 does not apply to $\mathrm{CSP}(\mathbb{Z}; +, R)$ since $(\mathbb{Z}; +, R)$ is not a core, nor does Corollary 7.10 applies to $\mathrm{CSP}(\mathbb{Z}; +, R, K)$ since $\mathrm{End}(\mathbb{Z}; +, R, K)$ is K, which is Horn-definable. One obtains hardness of $\mathrm{CSP}(\mathbb{Z}; +, R, K)$ by Theorem 7.12 as follows. Pick $a = 0$ (satisfying the linear equation $x = 0$ in the definition of R) and $b = 1$ (satisfying the modular linear equation $x = 1 \bmod 3$ in the definition of R), and define the relation $S = \{(\lambda, t) \mid \lambda \in K \wedge \lambda - t \in S\}$. Note that for all $\lambda \in K$, we have $S_\lambda = \{\lambda\} \cup 3\mathbb{Z} \cup (2 + 3\mathbb{Z})$. The formula $\exists w(w \in K \wedge S(\lambda, t) \wedge S(\lambda, t + 3w))$ defines the relation $T = \{(\lambda, t) \mid \lambda = 1 \bmod 3 \wedge (t = 0 \bmod 3 \vee t = 2 \bmod 3)\}$, which is fully modular and not Horn-definable. Proposition 7.5 implies that $\mathrm{CSP}(\mathbb{Z}; +, R, S)$ is NP-hard.

7.4 Tractability

In this section we show the following.

Proposition 7.13. *Let $\mathcal{A}$ be a structure with finite relational signature, domain $\mathbb{Z}$, and whose relations have quantifier-free Horn definitions over $(\mathbb{Z}; +, 1)$. Then there is an algorithm that solves CSP($\mathcal{A}$) in polynomial time.*

This result follows from the following more general result.

Theorem 7.14. *Let ϕ be a quantifier-free Horn formula over $(\mathbb{Z}; +)$, allowing parameters from $\mathbb{Z}$ represented in binary. Then there exists a polynomial-time algorithm to decide whether ϕ is satisfiable over $(\mathbb{Z}; +)$.*

The proof of Theorem 7.14 can be found at the end of this section. We first show how to derive Proposition 7.13.

Proof of Proposition 7.13. The input of $\mathrm{CSP}(\mathcal{A})$ consists of a primitive positive sentence whose atomic formulas are of the form $R(x_1, \ldots, x_k)$ where R is quantifier-free Horn definable over $\mathcal{L}_{(\mathbb{Z};+,1)}$. Since $\sum_{i=1}^n a_i x_i = b \bmod c$ is equivalent to $\sum_{i=1}^n a_i x_i = b + ck$, where k is a new integer variable, we can as well assume that the input to our problem consists of a set of Horn clauses over $(\mathbb{Z}; +, 1)$. This is tacitly the process of quantifier introduction, the converse of quantifier elimination. Then apply Theorem 7.14. □

Our algorithm for the proof of Theorem 7.14 uses two other well-known algorithms:

1. a polynomial-time algorithm for satisfiability of linear diophantine equations, i.e., the subproblem of the computational problem from Theorem 7.14 where the input only contains atomic formulas (see, e.g., [85]).

2. a polynomial-time algorithm to compute the rank of a matrix over $\mathbb{Q}$; this allows us in particular to decide whether a given linear system of equalities implies another equality over the rationals (this is standard, using Gaussian elimination; again, see [85] for a discussion of the complexity).

These two algorithms can be combined to obtain the following.

Lemma 7.15. *There is a polynomial-time algorithm that decides whether a given system Φ of linear diophantine equations implies another given diophantine equation ψ over $\mathbb{Z}$.*

Proof. First, use the first algorithm above to test whether Φ has a solution over $\mathbb{Z}$. If no, return yes (false implies everything). If yes, we claim that Φ implies ψ over $\mathbb{Q}$ (which can be tested by the second algorithm above) if and only if Φ implies ψ over the integers. Clearly, if every rational solution of Φ satisfies ψ, then so does every integer solution. Suppose now that there exists a rational solution α to Φ which does not satisfy ψ. Also take an integer solution β to Φ. Then on the line L between α and β there are infinitely many integer points. If one of them does not satisfy ψ, then Φ does not imply ψ over the integers. If all of them satisfy ψ, then all points of L must satisfy ψ, in particular α, a contradiction. □

Given the two mentioned algorithms, our procedure for the proof of Theorem 7.14 is basically an implementation of positive unit clause resolution. It takes the same form as the algorithm presented in [18] for satisfiability over the rationals.

```
// Input: a set of Horn-clauses C over (Z;+) with parameters.
// Output: satisfiable if C is satisfiable in (Z;+), unsatisfiable otherwise
Let U be clauses from C that only contain a single positive literal.
If U is unsatisfiable then return unsatisfiable.
Do
    For all negative literals ¬φ in clauses from C
        If U implies φ, then delete the negative literal ¬φ from all clauses in C.
    If C contains an empty clause, then return unsatisfiable.
    If C contains a clause with a single positive literal ψ, then add {ψ} to U.
Loop until no literal has been deleted
Return satisfiable.
```

Figure 7.1: An algorithm for satisfiability of Horn formulas with parameters over $(\mathbb{Z};+)$.

Proof. We follow the proof of Proposition 3.1 from [18]. We first discuss the correctness of the algorithm.

When $\mathcal{U}$ logically implies ϕ (which can be tested with the algorithm from Lemma 7.15) then the negative literal $\neg\phi$ is never satisfied and can be deleted from all clauses without

affecting the set of solutions. Since this is the only way in which literals can be deleted from clauses, it is clear that if one clause becomes empty the instance is unsatisfiable.

If the algorithm terminates with *satisfiable*, then no negation of an inequality is implied by $\mathcal{U}$. If r is the rank of the linear equation system defined by $\mathcal{U}$, we can use Gaussian elimination to eliminate r of the variables from all literals in the remaining clauses. For each of the remaining inequalities, consider the sum of absolute values of all coefficients. Let S be one plus the maximum of the this sum over all the remaining inequalities. Then setting the i-th variable to S^i satisfies all clauses. To see this, take any inequality, and assume that i is the highest variable index in this inequality. Order the inequality in such a way that the variable with highest index is on one side and all other variables on the other side of the $\neq$ sign. The absolute value on the side with the i-th variable is at least S^i. The absolute value on the other side is less than $S^i - S$, since all variables have absolute value less than S^{i-1} and the sum of all coefficients is less than $S - 1$ in absolute value. Hence, both sides of the inequality have different absolute value, and the inequality is satisfied. Since all remaining clauses have at least one inequality, all constraints are satisfied.

Now let us address the complexity of the algorithm. With appropriate data structures, the time needed for removing negated literals $\neg\phi$ from all clauses when ϕ is implied by $\mathcal{U}$ is linearly bounded in the input size since each literal can be removed at most once. $\square$

7.5 Conclusion

We are finally in position to prove the main result.

Proof of Theorem 7.1. Let $\mathcal{A}$ be first-order definable in $(\mathbb{Z}; +, 1)$ and suppose that $\mathcal{A}$ contains $+$. By Lemma 7.3 there exists a core $\mathcal{B}$ of $\mathcal{A}$. If $\mathcal{B}$ has only one element then $\mathrm{CSP}(\mathcal{B})$ and $\mathrm{CSP}(\mathcal{A})$ are trivially in P. Otherwise, $\mathcal{B}$ is itself first-order definable in $(\mathbb{Z}; +, 1)$ and contains $+$, and the statement follows from Theorem 7.12 and Proposition 7.13. $\square$

Bibliography

[1] Albert Atserias, Andrei A. Bulatov, and Anuj Dawar. "Affine systems of equations and counting infinitary logic". In: *Theoretical Computer Science* 410.18 (2009), pp. 1666–1683. DOI: 10.1016/j.tcs.2008.12.049.

[2] Libor Barto, Michael Kompatscher, Miroslav Olšák, Michael Pinsker, and Trung Van Pham. "The equivalence of two dichotomy conjectures for infinite domain constraint satisfaction problems". In: *Proceedings of the Annual Symposium on Logic in Computer Science (LICS)*. 2017. DOI: 10.1109/LICS.2017.8005128.

[3] Libor Barto and Marcin Kozik. "Constraint Satisfaction Problems of Bounded Width". In: *Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS)*. 2009, pp. 595–603. DOI: 10.1109/FOCS.2009.32.

[4] Libor Barto and Marcin Kozik. "Absorbing subalgebras, cyclic terms and the constraint satisfaction problem". In: *Logical Methods in Computer Science* 8.1 (2012), pp. 1–26. DOI: 10.2168/LMCS-8(1:7)2012.

[5] Libor Barto, Marcin Kozik, and Todd Niven. "The CSP Dichotomy Holds for Digraphs with No Sources and No Sinks (A Positive Answer to a Conjecture of Bang-Jensen and Hell)". In: *SIAM Journal on Computing* 38.5 (2009). DOI: 10.1137/070708093.

[6] Libor Barto, Jakub Opršal, and Michael Pinsker. "The wonderland of reflections". In: *Israel Journal of Mathematics* 223.1 (2018). DOI: 10.1007/s11856-017-1621-9.

[7] Libor Barto and Michael Pinsker. "The algebraic dichotomy conjecture for infinite domain constraint satisfaction problems". In: *Proceedings of the Annual Symposium on Logic in Computer Science (LICS)*. 2016, pp. 615–622. DOI: 10.1145/2933575.2934544.

[8] Marc Bezem, Robert Nieuwenhuis, and Enric Rodríguez-Carbonell. "The Max-Atom Problem and Its Relevance". In: *Proceedings of the International Conference on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR)*. 2008, pp. 47–61. DOI: 10.1007/978-3-540-89439-1_4.

[9] Manuel Bodirsky. *Constraint Satisfaction with Infinite Domains*. Dissertation, Humboldt-Universität zu Berlin. 2004.

[10] Manuel Bodirsky. "Cores of countably categorical structures". In: *Logical Methods in Computer Science* 3.1 (2007), pp. 1–16. DOI: 10.2168/LMCS-3(1:2)2007.

[11] Manuel Bodirsky. "Complexity Classification in Infinite-Domain Constraint Satisfaction". Mémoire d'habilitation à diriger des recherches, Université Diderot – Paris 7. Available at arXiv:1201.0856v8. 2012.

[12] Manuel Bodirsky. "Ramsey Classes: Examples and Constructions". In: *Surveys in Combinatorics. London Mathematical Society Lecture Note Series 424*. 2015, pp. 1–48. DOI: 10.1017/CBO9781316106853.002.

[13] Manuel Bodirsky and Hubie Chen. "Qualitative Temporal and Spatial Reasoning Revisited". In: *Computer Science Logic*. 2007, pp. 194–207. DOI: 10.1007/978-3-540-74915-8_17.

[14] Manuel Bodirsky and Víctor Dalmau. "Datalog and Constraint Satisfaction with Infinite Templates". In: *Journal on Computer and System Sciences* 79.1 (2013), pp. 79–100. DOI: 10.1016/j.jcss.2012.05.012.

[15] Manuel Bodirsky, Víctor Dalmau, Barnaby Martin, Antoine Mottet, and Michael Pinsker. "Distance Constraint Satisfaction Problems". In: *Information and Computation* 247 (2016), pp. 87–105. DOI: 10.1016/j.ic.2015.11.010.

[16] Manuel Bodirsky and Martin Grohe. "Non-dichotomies in Constraint Satisfaction Complexity". In: *Proceedings of the International Colloquium on Automata, Languages and Programming (ICALP)*. Lecture Notes in Computer Science. 2008, pp. 184–196. DOI: 10.1007/978-3-540-70583-3_16.

[17] Manuel Bodirsky, Martin Hils, and Barnaby Martin. "On the scope of the universal-algebraic approach to constraint satisfaction". In: *Logical Methods in Computer Science* 8.3 (2012). DOI: 10.2168/LMCS-8(3:13)2012.

[18] Manuel Bodirsky, Peter Jonsson, and Timo von Oertzen. "Horn versus full first-order: complexity dichotomies for algebraic constraint satisfaction". In: *Journal of Logic and Computation* 22.3 (2011), pp. 643–660. DOI: 10.1093/logcom/exr011.

[19] Manuel Bodirsky, Peter Jonsson, and Timo von Oertzen. "Essential Convexity and Complexity of Semi-algebraic Constraints". In: *Logical Methods in Computer Science* 8.4 (2012). DOI: 10.2168/LMCS-8(4:5)2012.

[20] Manuel Bodirsky, Peter Jonsson, and Trung Van Pham. "The Complexity of Phylogeny Constraint Satisfaction Problems". In: *ACM Transactions on Computational Logic* 18.3 (2017), 23:1–23:42. DOI: 10.1145/3105907.

[21] Manuel Bodirsky and Jan Kára. "The Complexity of Equality Constraint Languages". In: *Theory of Computing Systems* 3.2 (2008), pp. 136–158. DOI: 10.1007/s00224-007-9083-9.

[22] Manuel Bodirsky and Jan Kára. "The Complexity of Temporal Constraint Satisfaction Problems". In: *Proceedings of the Annual Symposium on Theory of Computing (STOC)*. ACM, May 2008, pp. 29–38.

[23] Manuel Bodirsky and Jan Kára. "The Complexity of Temporal Constraint Satisfaction Problems". In: *Journal of the ACM* 57.2 (2009), 9:1–9:41. DOI: 10.1145/1667053.1667058.

[24] Manuel Bodirsky, Florent Madelaine, and Antoine Mottet. "A universal-algebraic proof of the dichotomy for Monotone Monadic SNP". In: *Proceedings of the Annual Symposium on Logic in Computer Science (LICS)*. 2018, pp. 105–114. DOI: 10.1145/3209108.3209156.

[25] Manuel Bodirsky and Marcello Mamino. "Constraint Satisfaction Problems over Numeric Domains". In: *The Constraint Satisfaction Problem: Complexity and Approximability*. Vol. 7. Dagstuhl Follow-Ups. 2017, pp. 79–111. DOI: 10.4230/DFU.Vol7.15301.79.

[26] Manuel Bodirsky, Barnaby Martin, Marcello Mamino, and Antoine Mottet. "The complexity of disjunctive linear Diophantine constraints". In: *Proceedings of the International Symposium on Mathematical Foundations of Computer Science (MFCS)*. 2018.

[27] Manuel Bodirsky, Barnaby Martin, and Antoine Mottet. "Constraint Satisfaction Problems over the Integers with Successor". In: *Proceedings of the International Colloquium on Automata, Languages and Programming (ICALP)*. 2015, pp. 256–267. DOI: 10.1007/978-3-662-47672-7_21.

[28] Manuel Bodirsky, Barnaby Martin, and Antoine Mottet. "Discrete Temporal Constraint Satisfaction Problems". In: *Journal of the ACM* 65.2 (2018), 9:1–9:41. DOI: 10.1145/3154832.

[29] Manuel Bodirsky, Barnaby Martin, Michael Pinsker, and András Pongrácz. "Constraint Satisfaction Problems for Reducts of Homogeneous Graphs". In: *Proceedings of the International Colloquium on Automata, Languages and Programming (ICALP)*. 2016, 119:1–119:14. DOI: 10.4230/LIPIcs.ICALP.2016.119.

[30] Manuel Bodirsky and Antoine Mottet. "Reducts of finitely bounded homogeneous structures, and lifting tractability from finite-domain constraint satisfaction". In: *Proceedings of the Annual Symposium on Logic in Computer Science (LICS)*. 2016, pp. 623–632. DOI: 10.1145/2933575.2934515.

[31] Manuel Bodirsky and Antoine Mottet. "A Dichotomy for First-Order Reducts of Unary Structures". In: *Logical Methods in Computer Science* 14.2 (2018). DOI: 10.23638/LMCS-14(2:13)2018.

[32] Manuel Bodirsky and Jaroslav Nešetřil. "Constraint Satisfaction with Countable Homogeneous Templates". In: *Journal of Logic and Computation* 16.3 (2006), pp. 359–373. DOI: 10.1093/logcom/exi083.

[33] Manuel Bodirsky and Michael Pinsker. "Minimal functions on the random graph". In: *Israel Journal of Mathematics* 200.1 (2014), pp. 251–296. DOI: 10.1007/s11856-014-1042-y.

[34] Manuel Bodirsky and Michael Pinsker. "Schaefer's theorem for graphs". In: *Journal of the ACM* 62.3 (2015). A conference version appeared in the Proceedings of STOC 2011, pages 655–664, 52 pages (article number 19).

[35] Manuel Bodirsky and Michael Pinsker. "Topological Birkhoff". In: *Transactions of the American Mathematical Society* 367 (2015), pp. 2527–2549.

[36] Manuel Bodirsky and Michael Pinsker. "Canonical Functions: a Proof via Topological Dynamics". Preprint available under http://arxiv.org/abs/1610.09660. 2016.

[37] Manuel Bodirsky, Michael Pinsker, and András Pongrácz. "Projective clone homomorphisms". Accepted for publication in the Journal of Symbolic Logic, Preprint arXiv:1409.4601. 2014.

[38] Manuel Bodirsky, Michael Pinsker, and András Pongrácz. "The 42 reducts of the random ordered graph". In: *Proceedings of the LMS* 111.3 (2015). Preprint available from arXiv:1309.2165, pp. 591–632.

[39] Manuel Bodirsky and Stefan Woelfl. "RCC8 is tractable on instances of bounded treewidth". In: *Proceedings of International Joint Conferences on Artificial Intelligence (IJCAI)*. 2011, pp. 756–761.

[40] Manuel Bodirsky and Michał Wrona. "Equivalence Constraint Satisfaction Problems". In: *Proceedings of Computer Science Logic*. Vol. 16. LIPICS. Dagstuhl Publishing, Sept. 2012, pp. 122–136.

[41] Mathias Broxvall and Peter Jonsson. "Point algebras for temporal reasoning: Algorithms and complexity". In: *Artificial Intelligence* 149.2 (2003), pp. 179–220.

[42] Andrei A. Bulatov. "H-Coloring dichotomy revisited". In: *Theoretical Computer Science* 349.1 (2005), pp. 31–39.

[43] Andrei A. Bulatov. "A dichotomy theorem for constraint satisfaction problems on a 3-element set". In: *Journal of the ACM* 53.1 (2006), pp. 66–120.

[44] Andrei A. Bulatov. "A Dichotomy Theorem for Nonuniform CSPs". In: *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*. 2017, pp. 319–330.

[45] Andrei A. Bulatov and Peter G. Jeavons. *Algebraic Structures in Combinatorial Problems*. Technical report MATH-AL-4-2001, Technische Universität Dresden. 2001.

[46] Andrei A. Bulatov, Peter G. Jeavons, and Andrei A. Krokhin. "Classifying the complexity of constraints using finite algebras". In: *SIAM Journal on Computing* 34 (2005), pp. 720–742.

[47] Stanley N. Burris and Hanamantagouda P. Sankappanavar. *A Course in Universal Algebra*. Springer Verlag, Berlin, 1981.

[48] Gregory Cherlin, Saharon Shelah, and Niangdong Shi. "Universal graphs with forbidden subgraphs and algebraic closure". In: *Advances in Applied Mathematics* 22 (1999), pp. 454–491.

[49] Tsu-Wu J. Chou and George E. Collins. "Algorithms for the solution of systems of linear Diophantine equations". In: *SIAM Journal on Computing* 11 (1982), pp. 687–708.

[50] Andrzej Ehrenfeucht and Jan Mycielski. "Positional strategies for mean payoff games". In: *International Journal of Game Theory* 8.2 (1979), pp. 109–113. ISSN: 1432-1270.

[51] Tomás Feder and Moshe Y. Vardi. "Monotone monadic SNP and constraint satisfaction". In: *Proceedings of the Symposium on Theory of Computing (STOC)*. 1993, pp. 612–622.

[52] Tomás Feder and Moshe Y. Vardi. "The computational structure of monotone monadic SNP and constraint satisfaction: a study through Datalog and group theory". In: *SIAM Journal on Computing* 28.1 (1998), pp. 57–104.

[53] Roland Fraıssé. "Sur l'extension aux relations de quelques propriétés des ordres". In: *Annales Scientifiques de l'École Normale Supérieure* 71 (1954), pp. 363–388.

[54] Mark H. Siggers. "A Strong Mal'cev Condition for Varieties Omitting the Unary Type". In: *Algebra Universalis* 64.1 (2010), pp. 15–20.

[55] Pavol Hell and Jaroslav Nešetřil. "On the complexity of H-coloring". In: *Journal of Combinatorial Theory, Series B* 48 (1990), pp. 92–110.

[56] Wilfrid Hodges. *A shorter model theory.* Cambridge: Cambridge University Press, 1997.

[57] Jan Hubička and Jaroslav Nešetřil. "Universal Structures with Forbidden Homomorphisms". In: *Logic Without Borders - Essays on Set Theory, Model Theory, Philosophical Logic and Philosophy of Mathematics.* arXiv:0907.4079. De Gruyter, 2015, pp. 241–264.

[58] Jan Hubička and Jaroslav Nešetřil. "All those Ramsey classes (Ramsey classes with closures and forbidden homomorphisms)". arXiv:1606.07979v2. 2016.

[59] Jan Hubička and Jaroslav Nešetřil. "Homomorphism and Embedding Universal Structures for Restricted Classes". In: *Multiple-Valued Logic and Soft Computing* 27.2-3 (2016). arXiv:0909.4939, pp. 229–253.

[60] Neil Immerman. *Descriptive Complexity.* Graduate Texts in Computer Science. Springer, 1999.

[61] Peter Jeavons, David Cohen, and Martin C. Cooper. "Constraints, Consistency and Closure". In: *Artificial Intelligence* 101.1-2 (1998), pp. 251–265.

[62] Peter Jeavons, David Cohen, and Marc Gyssens. "Closure Properties of Constraints". In: *Journal of the ACM* 44.4 (1997), pp. 527–548.

[63] Peter Jonsson and Thomas Drakengren. "A Complete Classification of Tractability in RCC-5". In: *Journal of Artificial Intelligence Research* 6 (1997), pp. 211–221.

[64] Peter Jonsson and Tomas Lööw. "Computation complexity of linear constraints over the integers". In: *Artificial Intelligence* 195 (2013), pp. 44–62.

[65] Peter Jonsson and Johan Thapper. "Constraint satisfaction and semilinear expansions of addition over the rationals and the reals". In: *J. Comput. Syst. Sci.* 82.5 (2016), pp. 912–928. DOI: 10.1016/j.jcss.2016.03.002.

[66] Bartek Klin, Eryk Kopczynski, Joanna Ochremiak, and Szymon Toruńczyk. "Locally Finite Constraint Satisfaction Problems". In: *30th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2015, Kyoto, Japan.* 2015, pp. 475–486.

[67] Bartek Klin, Slawek Lasota, Joanna Ochremiak, and Szymon Toruńczyk. "Homomorphism Problems for First-Order Definable Structures". In: *FSTTCS.* 2016.

[68] Michael Kompatscher and Trung Van Pham. "A Complexity Dichotomy for Poset Constraint Satisfaction". In: *34th Symposium on Theoretical Aspects of Computer Science (STACS 2017).* Vol. 66. Leibniz International Proceedings in Informatics (LIPIcs). 2017, 47:1–47:12. ISBN: 978-3-95977-028-6.

[69] Dexter Kozen. "Results on the propositional μ-calculus". In: *Theoretical Computer Science* 27 (1983), pp. 333–354.

[70] Marcin Kozik, Andrei Krokhin, Matt Valeriote, and Ross Willard. "Characterizations of several Maltsev conditions". In: *Algebra universalis* 73.3 (2015), pp. 205–224. DOI: 10.1007/s00012-015-0327-2.

[71] Gábor Kun. "Constraints, MMSNP, and Expander Relational Structures". In: *Combinatorica* 33.3 (2013), pp. 335–347.

[72] Richard E. Ladner. "On the Structure of Polynomial Time Reducibility". In: *Journal of the ACM* 22.1 (1975), pp. 155–171.

[73] Carsten Lutz and Frank Wolter. "On the Relationship between Consistent Query Answering and Constraint Satisfaction Problems". In: *Proceedings of the 18th International Conference on Database Theory (ICDT15).* 2015.

[74] Alan K. Mackworth. "Consistency in networks of relations". In: *Artificial Intelligence* 8 (1977), pp. 99–118.

[75] Florent Madelaine. "Constraint satisfaction problems and related logic". PhD thesis. University of Leicester, 2003.

[76] Florent Madelaine and Iain A. Stewart. "Constraint Satisfaction, Logic and Forbidden Patterns". In: *SIAM Journal on Computing* 37.1 (2007), pp. 132–163.

[77] David Marker. *Model Theory: An Introduction.* New York: Springer, 2002.

[78] Miklós Maróti and Ralph McKenzie. "Existence theorems for weakly symmetric operations". In: *Algebra Universalis* 59.3 (2008).

[79] Yuri V. Matiyasevich. *Hilbert's Tenth Problem.* Cambridge, Massachusetts: MIT Press, 1993.

[80] Rolf H. Möhring, Martin Skutella, and Frederik Stork. "Scheduling with AND/OR Precedence Constraints". In: *SIAM Journal on Computing* 33.2 (2004), pp. 393–415.

[81] Jaroslav Nešetřil. "Ramsey Classes and Homogeneous Structures". In: *Combinatorics, Probability & Computing* 14.1-2 (2005), pp. 171–189.

[82] Vladimir Pestov. "On free actions, minimal flows and a problem by Ellis". In: *Transations of the American Mathematical Society* 350.10 (1998), pp. 4149–4165.

[83] Mojżesz Presburger. "Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt". In: *Comptes Rendus du I congres de Mathématiciens des Pays Slaves* (1929), pp. 92–101.

[84] Thomas J. Schaefer. "The complexity of satisfiability problems". In: *Proceedings of the Symposium on Theory of Computing (STOC).* 1978, pp. 216–226.

[85] Alexander Schrijver. *Theory of Linear and Integer Programming.* Wiley - Interscience Series in Discrete Mathematics and Optimization, 1998.

[86] Katrin Tent and Martin Ziegler. *A course in model theory.* Lecture Notes in Logic. Cambridge University Press, 2012.

[87] Dmitriy Zhuk. "A Proof of CSP Dichotomy Conjecture". In: *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017.* 2017, pp. 331–342.

List of Figures

1.1 Overview of the results in the dissertation 3

2.1 The algebraic and topological approach for infinite-domain CSPs I 20
2.2 The algebraic and topological approach for infinite-domain CSPs II 21

4.1 A mashup of two binary functions. 38

5.1 The No-Mono-Triangle problem 53
5.2 Illustration of the formula $\chi_M^{(2)}(x)$ for Example 11 75
5.3 Illustration for the proof of Proposition 5.36 76
5.4 Illustration of Lemma 5.43. 81
5.5 Illustration of the proof of Proposition 5.46. 84
5.6 Proof of Theorem 5.47 86

6.1 Illustration for Lemma 6.9 102
6.2 Illustration for Lemma 6.9 103
6.3 Illustration of the proof of Lemma 6.15 107
6.4 Illustration of the proof of Lemma 6.15 107

7.1 An algorithm for satisfiability of Horn formulas with parameters over $(\mathbb{Z};+)$.146

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This paper has not previously been presented in identical or similar form to any other German or foreign examination board.

Antoine Mottet
Dresden, May 31, 2018