

Extensions and Limits of the Specker-Blatter Theorem

Eldar Fischer ✉

Faculty of Computer Science, Technion – Israel Institute of Technology, Haifa, Israel

Johann A. Makowsky ✉

Faculty of Computer Science, Technion – Israel Institute of Technology, Haifa, Israel

Abstract

The original Specker-Blatter Theorem (1983) was formulated for classes of structures \mathcal{C} of one or several binary relations definable in Monadic Second Order Logic MSOL. It states that the number of such structures on the set $[n]$ is modularly C-finite (MC-finite). In previous work we extended this to structures definable in CMSOL, MSOL extended with modular counting quantifiers. The first author also showed that the Specker-Blatter Theorem does not hold for one quaternary relation (2003).

If the vocabulary allows a constant symbol c , there are n possible interpretations on $[n]$ for c . We say that a constant c is *hard-wired* if c is always interpreted by the same element $j \in [n]$. In this paper we show:

- (i) The Specker-Blatter Theorem also holds for CMSOL when hard-wired constants are allowed. The proof method of Specker and Blatter does not work in this case.
- (ii) The Specker-Blatter Theorem does not hold already for \mathcal{C} with one ternary relation definable in First Order Logic FOL. This was left open since 1983.

Using hard-wired constants allows us to show MC-finiteness of counting functions of various restricted partition functions which were not known to be MC-finite till now. Among them we have the restricted Bell numbers $B_{r,A}$, restricted Stirling numbers of the second kind $S_{r,A}$ or restricted Lah-numbers $L_{r,A}$. Here r is a non-negative integer and A is an ultimately periodic set of non-negative integers.

2012 ACM Subject Classification Theory of computation → Finite Model Theory; Mathematics of computing → Enumeration

Keywords and phrases Specker-Blatter Theorem, Monadic Second Order Logic, MC-finiteness

Digital Object Identifier 10.4230/LIPIcs.CSL.2024.26

Related Version *Full Version:* <https://arxiv.org/abs/2206.12135>

Funding *Eldar Fischer:* Israel Science Foundation ISF 879/22

1 Introduction

A sequence of natural numbers $s(n)$ is *C-finite* if it satisfies a linear recurrence relation with constant coefficients. $s(n)$ is MC-finite if it satisfies a linear recurrence relation with constant coefficients modulo m for each m separately. A C-finite sequence $s(n)$ must have limited growth: $s(n) \leq 2^{cn}$ for some constant c . No such bound exists for MC-finite sequences: for every monotone increasing sequence $s(n)$ the sequence $s'(n) = n!s(n)$ is MC-finite.

A typical example of a C-finite sequence is the sequence $f(n)$ of Fibonacci numbers. A typical example of an MC-finite sequence which is not C-finite is the sequence $B(n)$ of Bell numbers. The Bell number $B(n)$ counts the number of partitions of the set $[n]$ of the numbers $\{1, 2, \dots, n\}$. Let $Eq(n)$ be number of equivalence relations over $[n]$. Clearly, $B(n) = Eq(n)$. Let $Eq_2(n)$ be the number of equivalence relations on $[n]$ with exactly two equivalence classes of the same size. $Eq_2(n)$ is not MC-finite since the value of $Eq_2(n)$ is odd iff n is an even power of 2, see [3].



© Eldar Fischer and Johann A. Makowsky;

licensed under Creative Commons License CC-BY 4.0

32nd EACSL Annual Conference on Computer Science Logic (CSL 2024).

Editors: Aniello Murano and Alexandra Silva; Article No. 26; pp. 26:1–26:20

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In [23] G. Pfeiffer discusses counting other transitive relations besides $Eq(n)$, in particular, partial orders $PO(n)$, quasi-orders (aka preorders) $QO(n)$ and just transitive relations $Tr(n)$. Using a growth argument one can see that none of these functions is C-finite. It follows directly from the Specker-Blatter Theorem stated below, see Corollary 2, that $PO(n)$, $QO(n)$ and $Tr(n)$ are MC-finite. However, to the best of our knowledge, this has not been stated in the literature. This may be due to the fact that no explicit formulas for these functions are known. The Specker-Blatter Theorem establishes MC-finiteness even in the absence of explicit formulas. It derives MC-finiteness solely from the assumption that \mathcal{C} is definable in Monadic Second Order Logic (MSOL), or in MSOL augmented by modular counting quantifiers (CMSOL).

The present paper grew out of our study of modular recurrence relations for restricted partition functions, [11]. We provide a short review of the Specker-Blatter Theorem, and show how to extend its applicability by extending the allowed vocabulary to include constants with a fixed interpretation (“hard-wired”). The reduction allowing this extension can be made to work in the other direction. Using it we also close the gap between the Specker-Blatter Theorem and its known limits, left open in [12], by constructing an FOL statement over a single ternary relation for which the theorem does not hold.

Formal definitions with more examples and details about C-finite and MC-finite sequences are given in Section 6.

2 Background in logic

We generally follow the notation of [8], and assume basic knowledge of model theory. Standard texts for Finite Model Theory are [8, 21]. In the following, we always refer to a set $\bar{R} = \{R_1, \dots, R_{\ell_{\bar{R}}}\}$ of distinct binary relation symbols, a set $\bar{a} = \{a_1, \dots, a_{\ell_{\bar{a}}}\}$ of distinct constant symbols, and so on. By $a \in \bar{a}$ we mean that there exists $1 \leq i \leq \ell_{\bar{a}}$ for which $a = a_i$. We also use the shorthand $[n] = \{1, \dots, n\}$.

Let $\tau = \bar{R} \cup \bar{a}$ be a vocabulary, i.e., a set of non-logical constants. We denote by $\text{FOL}(\tau)$ the set of first order formulas with its non-logical constants in τ . If τ is clear from the context, we omit it. We denote by $\text{MSOL}(\tau)$ the set of Monadic Second Order Logic, obtained from FOL by allowing unary relation variables and quantification over them. The logic CMSOL is obtained from MSOL by allowing also quantification of the form $C_{m,a}x\phi(x)$, which are interpreted by

$$\mathcal{A} \models C_{m,a}x\phi(x) \text{ iff } |\{a \in A : \phi(a)\}| \equiv a \pmod{m}.$$

In the following we will be interested in the set of models of a logic sentence ϕ over a vocabulary τ whose universe is $[n]$ for any natural number n . We denote this set by

$$\mathcal{C}_{\phi} = \{\mathfrak{M} = ([n], A_1, \dots, A_m) : n \in \mathbb{N}, A_i \in [n]^{\rho_i}, \mathfrak{M} \models \phi\}.$$

3 The original Specker-Blatter Theorem

Let ϕ_E be the formula in First Order Logic (FOL) which says that $E(x, y)$ is an equivalence relation. $Eq(n)$ can be written as

$$Eq(n) = |\{E \subseteq [n]^2 : ([n], E) \models \phi_E\}|.$$

$PO(n)$, $QO(n)$ and $Tr(n)$ can be written in a similar way.

The original Specker-Blatter Theorem from 1981, [1, 2, 3, 26], gives a general criterion for certain integer sequences to be MC-finite. Let $\bar{R} = \{R_1, \dots, R_{\ell_{\bar{R}}}\}$ be a finite set of relation symbols of arities $\rho_1, \dots, \rho_{\ell_{\bar{R}}}$ respectively, and ϕ be a formula of Monadic Second Order Logic (MSOL) using relation symbols from R without free variables.

Let $Sp_{\phi}(n)$ be the number of ways we can interpret the relation symbols in R on $[n]$ such that the resulting structures where A_i is the interpretation of R_i satisfies ϕ . Formally

$$Sp_{\phi}(n) = |\{A_i \subseteq [n]^{\rho_i}, i \leq m : ([n], A_1, \dots, A_m) \models \phi\}|.$$

► **Theorem 1** (Specker-Blatter). *Let \bar{R} be a finite set of binary relations and ϕ be a formula of MSOL(\bar{R}) using relation symbols in \bar{R} . Then the sequence $Sp_{\phi}(n)$ is MC-finite.*

► **Corollary 2.** *The sequences counting the number of partial orders $PO(n)$, quasi-orders $QO(n)$, and transitive relations $Tr(n)$ on $[n]$, are MC-finite.*

The idea behind the proof of the Specker-Blatter theorem consists of two parts, both of which use the assertion that $\tau = \bar{R}$ contains only binary relation symbols. Unary symbols can also be incorporated, since these can be simulated with binary symbols in a way that preserves the number of satisfying models.

The first part is combinatorial and applies to any family \mathcal{C} of structures over the vocabulary τ satisfying a property that we outline below. For such a family, we let $Sp_{\mathcal{C}}(n)$ be the number of members of \mathcal{C} whose universe is $[n]$. In particular, $Sp_{\phi}(n)$ is just a shorthand for $Sp_{\mathcal{C}_{\phi}}(n)$.

A pointed \bar{R} -structure is an \bar{R} -structure $\mathcal{A} = ([n], A_1, \dots, A_{\mu}, a)$ with an additional distinguished point $a \in [n]$. Given a pointed \bar{R} -structure \mathcal{A}_1 with universe $[n_1]$ and an \bar{R} -structure \mathcal{A}_2 with universe $[n_2]$ we define $\mathcal{A} = Subst(\mathcal{A}_1, a, \mathcal{A}_2)$ as follows:

- (i) The universe A of \mathcal{A} is the disjoint union of A_1 and A_2 with the point a removed. It can be assumed to be the set $[n_1 + n_2 - 1]$.
- (ii) The binary relations are defined such that \mathcal{A}_2 is a module in \mathcal{A} , i.e., for $u \in A_1 \setminus \{a\}$ and $v \in A_2$ and $R \in \bar{R}$, the relation $R(u, v)$ holds in $\mathcal{A} = Subst(\mathcal{A}_1, a, \mathcal{A}_2)$ iff $R(u, a)$ holds in \mathcal{A}_1 . For $u, v \in A_1 \setminus \{a\}$ (respectively $u, v \in A_2$), $R(u, v)$ holds in \mathcal{A} iff it holds in \mathcal{A}_1 (respectively \mathcal{A}_2).

By using an arbitrary enumeration of all possible pointed \bar{R} -structures and all possible (non-pointed) \bar{R} -structures, we construct an $\mathbb{N} \times \mathbb{N}$ matrix $M_{\mathcal{C}}$ over $\{0, 1\}$, by setting for every i and j the value $M_{\mathcal{C}}(i, j)$ to be the indicator as to whether the substitution of the j 'th structure in the i 'th pointed structure results in a member of \mathcal{C} . The main combinatorial part is the following.

► **Theorem 3** (Specker-Blatter, combinatorial version). *Let \bar{R} be a finite set of binary relations and \mathcal{C} be a class of finite \bar{R} -structures whose substitution rank is finite under \mathbb{Z}_{p^q} for any prime number p and $q \in \mathbb{N}$. Then the sequence $Sp_{\mathcal{C}}(n)$ is MC-finite.*

The above applies to an uncountable number of families \mathcal{C} . Theorem 1 follows from it by the following lemma, which forms the second part of the original proof:

► **Lemma 4.** *Let \bar{R} be a finite set of binary relations and \mathcal{C} be a finite class of R -structures defined by an R -sentences ϕ in MSOL. Then the substitution rank of \mathcal{C} is finite.*

In [15] it is shown that the lemma still holds if MSOL is replaced by CMSOL. On the other hand, when considering relations of arity higher than 2, the substitution operation is no longer well-defined as it is written here. As it later turned out, this is not a merely technical limitation, but an essential one.

Also, it is not clear how to handle hard-wired constant in the definition of the substitution operation. In this paper, instead of incorporating the hard-wired constants directly into the original mechanism, we show a reduction from the question of the original count to a sum of counts over other sentences that do not involve the constants. This approach turns out to be useful also in the other direction, of proving a new *limit* on the Specker-Blatter theorem.

4 Previous limitations and extensions

Limitations and extensions of the Specker-Blatter Theorem have been previously discussed in [15, 14].

It is well known that Eulerian graphs and regular graphs of even degree are not definable in MSOL, but they are definable in CMSOL. In [15], the Specker-Blatter Theorem was shown to hold also for CMSOL. It follows in particular that $Eul(n)$, which counts the number of Eulerian graphs over $[n]$ (i.e. connected graphs all of whose degrees are even), is also MC-finite.

In [13] the first author showed that the Specker-Blatter Theorem does not hold for quaternary relations:

► **Theorem 5** (E. Fischer, 2002). *There is an FOL-sentence with only one quaternary relation symbol ϕ , such that $Sp_\phi(n)$ is not an MC-sequence.*

The question of whether Specker-Blatter Theorem holds in the presence of ternary relation symbols remained open.

5 Main new results

Due to space constraints, some proofs are deferred to the full version of this paper¹ [17].

The Bell numbers $B(n)$ and the Stirling numbers of the second kind $S_k(n)$ for fixed k can be shown to be MC-finite using the Specker-Blatter Theorem. A. Broder in 1984, [4], introduced the restricted Bell numbers $B_r(n)$ and the restricted Stirling numbers of the second kind $S_{k,r}(n)$. Let $r \in \mathbb{N}^+$. $S_{k,r}(n)$ is defined as the number of set partitions of $[r+n]$ into $k+r$ blocks with the additional condition that the first r elements are in distinct blocks. $B_r(n)$ is defined as

$$B_r(n) = \sum_k S_{k,r}(n).$$

The class of equivalence relations on $[r+n]$ where the first r elements are in different equivalence classes is definable in FOL with one binary relation and r hard-wired constants. The Specker-Blatter Theorem does not directly apply to this case. In [11] it is shown how to circumvent this obstacle in the case of one equivalence relation. It followed that both $S_{k,r}(n)$ and $B_r(n)$ are MC-finite.

In this paper we prove a more general theorem:

► **Theorem 6** (Elimination of hard-wired constants).

- (i) *Let τ consist of a finite set of (hard-wired) constant symbols \bar{a} , unary relations symbols \bar{U} , and binary relation symbols \bar{R} . For every class \mathcal{C} of τ -structures there exist classes $\mathcal{C}_1, \dots, \mathcal{C}_r$ of τ' -structures, where τ' -contains only a finite number $r(\bar{a}, \bar{U}, \bar{R})$ of binary relation symbols, such that*

¹ The full version can be downloaded at <https://arxiv.org/abs/2206.12135>.

$$Sp_{\mathcal{C}}(n) = \sum_{i=1}^r Sp_{\mathcal{C}_i}(n).$$

Equality here is not modular.

(ii) Furthermore, if \mathcal{C} is FOL-definable (MSOL-definable, CMSOL-definable), so are the \mathcal{C}_i .

► **Corollary 7.** *Let τ consist of a finite set of (hard-wired) constant symbols \bar{a} , unary relations symbols \bar{U} , and binary relation symbols \bar{R} , and let \mathcal{C} be class of finite τ -structures definable in CMSOL. Then the sequence $Sp_{\mathcal{C}}(n)$ is MC-finite.*

The proof of Theorem 6 is given in Section 7. In Section 8 we state Theorem 19, which is an extension of Theorem 6 that works for higher arities and other logics, and sketch its proof. The full proof details of Theorem 19 are deferred to [17]. The extension to higher arities is needed for proving Theorem 8 below.

We have seen in Theorem 5 that the Specker-Blatter Theorem does not hold for a single quaternary relation. The question of whether Specker-Blatter Theorem holds in the presence of a single ternary relation symbol remained open. Our second main result here answers this.

► **Theorem 8 (Ternary Counter-Example).** *There is a FOL-sentence ϕ with only one ternary relation symbol (and some lower arity relations), such that $Sp_{\phi}(n)$ is not an MC-sequence.*

The proof of Theorem 8 first produces a sentence ψ which also uses one symbol for a hard-wired constant. This will be shown in Section 9. To construct ϕ without the hard-wired constants, we deploy the aforementioned Theorem 19, which provides a sentence with one ternary relation and several lower arity relations. We can then also eliminate all relations except the ternary one, to arrive at Theorem 31 stated at Section 9, whose proof is deferred to the full version [17]. A sketch thereof is still provided.

We conclude this paper with Section 10, containing a summary and open problems.

6 More details about C-finite and MC-finite sequences of integers

A sequence of integers $s(n)$ is *C-finite*² if there are constants $p, q \in \mathbb{N}$ and $c_i \in \mathbb{Z}$, $0 \leq i \leq p-1$ such that for all $n \geq q$ the linear recurrence relation below holds for $s(n)$.

$$s(n+p) = \sum_{i=0}^{p-1} c_i s(n+i).$$

A sequence of integers $s(n)$ is modular C-finite, abbreviated as *MC-finite*, if for every $m \in \mathbb{N}$ there are constants $p_m, q_m \in \mathbb{N}^+$ such that for every $n \geq q_m$ there is a linear recurrence relation

$$s(n+p_m) \equiv \sum_{i=0}^{p_m-1} c_{i,m} s(n+i) \pmod{m}$$

with constant coefficients $c_{i,m} \in \mathbb{Z}$.

We denote by $s^m(n)$ the sequence $s(n) \pmod{m}$. Note that the coefficients $c_{i,m}$ and both p_m and q_m generally do depend on m .

² These are also called constant-recursive sequences or linear-recursive sequences in the literature.

► **Proposition 9.** *The sequence $s(n)$ is MC-finite iff $s^m(n)$ is ultimately periodic for every m .*

Proof. MC-finiteness implies periodicity. The converse is from [24]. ◀

Clearly, if a sequence $s(n)$ is C-finite then it is also MC-finite with $r_m = r$ and $c_{i,m} = c_i$ for all m . The converse is not true as there are uncountably many MC-finite sequences, but only countably many C-finite sequences with integer coefficients, see Proposition 11 below.

► **Example 10.**

- (i) The Fibonacci sequence is C-finite.
- (ii) If $s(n)$ is C-finite then it has at most simple exponential growth. There is $c \in \mathbb{N}^+$ such that $s(n) \leq 2^{cn}$ for all $n \in \mathbb{N}$, see e.g. [9, 19].
- (iii) The Bell numbers $B(n)$ are *not C-finite*, but are *MC-finite*.
- (iv) Let $f(n)$ be any integer sequence. The sequence $s_1(n) = 2 \cdot f(n)$ is ultimately periodic modulo 2, but not necessarily MC-finite.
- (v) Let $g(n)$ be any integer sequence which is not almost everywhere zero. The sequence $s_2(n) = n! \cdot g(n)$ is MC-finite but not C-finite due to its growth.
- (vi) The sequence $s_3(n) = \frac{1}{2} \binom{2n}{n}$ is not MC-finite: $s_3(n)$ is odd if and only if n is a power of 2 (Lucas, 1878). A proof may be found in [18, Exercise 5.61] or in [26].
- (vii) The Catalan numbers $C(n) = \frac{1}{n+1} \binom{2n}{n}$ are not MC-finite, since $C(n)$ is odd iff n is a Mersenne number, i.e., $n = 2^m - 1$ for some m , see [20, Chapter 13].
- (viii) Let p be a prime and $f(n)$ monotone increasing. The sequence $s(n) = p \cdot f(n) + z(n)$, where $z(n)$ is defined to equal 1 if n is a power of p and to equal 0 for any other n , is monotone increasing but not ultimately periodic modulo p , hence not MC-finite.

► **Proposition 11.**

- (i) *There are uncountably many monotone increasing sequences which are MC-finite, and uncountably many which are not MC-finite.*
- (ii) *Almost all integer sequences (under a suitable measure) are not MC-finite.*

Proof. (i) follows from Example 10 (v) and (viii). (ii) follows from almost all integer sequences being absolutely normal (see [9]); the full proof is deferred to [17]. ◀

7 Proving the reduction

7.1 Introduction

In the following we consider extending the language with “hard-wired” constants. Specifically, assume that we have a class \mathcal{C} that is defined by a sentence ϕ involving a set of constant symbols \bar{a} , unary symbols \bar{U} and binary symbols \bar{R} . The function $f_{\mathcal{C}}(n)$ is defined as the number of models over the universe $[n + \ell_{\bar{a}}]$ which satisfy ϕ , for which a_i is interpreted as $n + i$ for all $i \in [\ell_{\bar{a}}]$. Note the distinction from the non-hard-wired setting, where we would have had to also count the possible interpretations of the constants.

Our main result is an expression for the function $f_{\mathcal{C}}(n)$ (when constants are allowed) that is based on counting functions for classes that do not utilize constants. We first show this reduction for languages using only unary and binary relations. The reduction preserves many of the common logics, in particular an FOL expression would be reduced to functions involving FOL expressions, and so on. This extends the Specker-Blatter theorem to languages involving hard-wired constants, allowing modular ultimate periodicity proofs of new functions.

In this section we prove Theorem 6. For convenience we state it again as Theorem 12.

► **Theorem 12** (Reducing model counts to the case without constants). *For any class \mathcal{C} defined by an FOL (resp. MSOL, CMSOL) sentence involving a set of constant symbols \bar{a} , unary symbols \bar{U} and binary symbols \bar{R} , there exist classes $\mathcal{C}_1, \dots, \mathcal{C}_r$ (where r depends on the original language), definable by FOL (resp. MSOL, CMSOL) sentences involving \bar{U}' (which contains \bar{U}), \bar{R} and no constants, satisfying $f_{\mathcal{C}}(n) = \sum_{i=1}^r f_{\mathcal{C}_i}(n)$ for all $n \in \mathbb{N}$.*

The following is the immediate corollary it produces for the Specker-Blatter Theorem, which is a restatement of Corollary 2.

► **Corollary 13** (Extended Specker-Blatter Theorem). *For a class \mathcal{C} definable in CMSOL with (hard-wired) constants, unary and binary relation symbols only, the function $f_{\mathcal{C}}$ is MC-finite.*

Theorem 12 is proved by induction over the number of constants. The basis, $\ell_{\bar{a}} = 0$, is trivial (with $\bar{U}' = \bar{U}$, $r = 1$ and $\mathcal{C}_1 = \mathcal{C}$). The induction step is provided by the following.

► **Lemma 14** (Removing a single constant). *For any class \mathcal{C} defined by an FOL (resp. MSOL, CMSOL) sentence involving a set of constant symbols \bar{a} with $\ell_{\bar{a}} > 0$, unary symbols \bar{U} and binary symbols \bar{R} , there exist classes $\mathcal{C}_1, \dots, \mathcal{C}_r$ (where r depends on the original language), definable by FOL (resp. MSOL, CMSOL) sentences over the language $(\bar{a}', \bar{U}', \bar{R}')$, where $\bar{a}' = \bar{a} \setminus \{a_{\ell_{\bar{a}}}\}$, $\bar{U}' = \bar{U} \cup \bar{I} \cup \bar{O}$ where $\ell_{\bar{I}} = \ell_{\bar{O}} = \ell_{\bar{R}}$, and $\bar{R}' = \bar{R}$, satisfying $f_{\mathcal{C}}(n) = \sum_{i=1}^r f_{\mathcal{C}_i}(n)$ for all $n \in \mathbb{N}$.*

The main idea in the proof of this lemma is to encode the “interaction” of the constant $a_{\ell_{\bar{a}}}$ with the rest of the universe using the additional unary relations. For every $i \in [\ell_{\bar{R}}]$, we will use the new relation I_i to hold every $x \neq a_{\ell_{\bar{a}}}$ for which (x, a) was in R_i , and the relation O_i to hold every $x \neq a_{\ell_{\bar{a}}}$ for which (a, x) was in R_i .

We cannot directly keep track whether (a, a) was in R_i , or whether a was in U_i for $i \in [\ell_{\bar{U}}]$, so we count the number of models for each of these options separately. This sets $r = 2^{\ell_{\bar{U}} + \ell_{\bar{R}}}$. Instead of a running index, we index each such option with a set $\mathfrak{U} \subseteq [\ell_{\bar{U}}]$ denoting which of the relations in \bar{U} include the constant to be removed $a = a_{\ell_{\bar{a}}}$, and a set $\mathfrak{R} \subseteq [\ell_{\bar{R}}]$ denoting which of the relations in \bar{R} include (a, a) . Using these we can define the case where a model \mathfrak{N} over the language $(\bar{a}', \bar{U}', \bar{R})$ with universe $[n + \ell_{\bar{a}} - 1]$ corresponds (along with \mathfrak{U} and \mathfrak{R}) to an “original model” \mathfrak{M} with universe $[n + \ell_{\bar{a}}]$ over the original language.

► **Definition 15.** *Given a model \mathfrak{M} over the language $(\bar{a}, \bar{U}, \bar{R})$ with universe $[n + \ell_{\bar{a}}]$, a model \mathfrak{N} over the language $(\bar{a}', \bar{U}', \bar{R})$ with universe $[n + \ell_{\bar{a}} - 1]$, and sets $\mathfrak{U} \subseteq [\ell_{\bar{U}}]$ and $\mathfrak{R} \subseteq [\ell_{\bar{R}}]$, where (as always) in both models every constant a_i is interpreted to be $n + i$, we say that $(\mathfrak{N}, \mathfrak{U}, \mathfrak{R})$ correspond to \mathfrak{M} if the following hold.*

- For every $U \in \bar{U}$ and $x \in [n + \ell_{\bar{a}} - 1]$, we have $\mathfrak{N} \models U(x)$ if and only if $\mathfrak{M} \models U(x)$.
- For every $i \in [\ell_{\bar{U}}]$, we have $i \in \mathfrak{U}$ if and only if $\mathfrak{M} \models U_i(a)$.
- For every $R \in \bar{R}$ and $x, y \in [n + \ell_{\bar{a}} - 1]$, we have $\mathfrak{N} \models R(x, y)$ if and only if $\mathfrak{M} \models R(x, y)$.
- For every $i \in [\ell_{\bar{R}}]$ and $x \in [n + \ell_{\bar{a}} - 1]$, we have $\mathfrak{N} \models I_i(x)$ if and only if $\mathfrak{M} \models R_i(x, a)$.
- For every $i \in [\ell_{\bar{R}}]$ and $x \in [n + \ell_{\bar{a}} - 1]$, we have $\mathfrak{N} \models O_i(x)$ if and only if $\mathfrak{M} \models R_i(a, x)$.
- For every $i \in [\ell_{\bar{R}}]$, we have $i \in \mathfrak{R}$ if and only if $\mathfrak{M} \models R_i(a, a)$.

It is important to note, for the purpose of counting, the following observation.

► **Observation 16.** *Definition 15 provides a bijection between the set of possible models \mathfrak{M} over the universe $[n + \ell_{\bar{a}}]$ (where the constants are interpreted as in Definition 15), and the set of possible triples $(\mathfrak{N}, \mathfrak{U}, \mathfrak{R})$ where \mathfrak{N} is a model over $[n + \ell_{\bar{a}} - 1]$ (where the constants are interpreted as in Definition 15) and $\mathfrak{U} \subseteq [\ell_{\bar{U}}]$ and $\mathfrak{R} \subseteq [\ell_{\bar{R}}]$.*

Suppose we are given an expression $\phi(\bar{x})$ where $\bar{x} = \{x_1, \dots, x_{\ell_{\bar{x}}}\}$ is a set of variable symbols over the language $(\bar{a}, \bar{U}, \bar{R})$, as well as a set $\mathfrak{U} \subseteq [\ell_{\bar{U}}]$ and a set $\mathfrak{R} \subseteq [\ell_{\bar{R}}]$. We will construct, by induction over the structure of ϕ , several expressions, where one of which is an expression $\phi'_{\mathfrak{U}, \mathfrak{R}}(\bar{x})$ over the language $(\bar{a}', \bar{U}', \bar{R})$. It will be constructed so that for any \mathfrak{M} over the language $(\bar{a}, \bar{U}, \bar{R})$ with universe $[n + \ell_{\bar{a}}]$ and \mathfrak{N} over the language $(\bar{a}', \bar{U}', \bar{R})$ with universe $[n + \ell_{\bar{a}} - 1]$, where $(\mathfrak{N}, \mathfrak{U}, \mathfrak{R})$ correspond to \mathfrak{M} , and any fixing of $x_1, \dots, x_{\ell_{\bar{x}}} \in [n + \ell_{\bar{a}} - 1]$, we will have $\mathfrak{M} \models \phi(\bar{x})$ if and only if $\mathfrak{N} \models \phi'_{\mathfrak{U}, \mathfrak{R}}(\bar{x})$.

Lemma 14 then immediately follows from the case $\ell_{\bar{x}} = 0$ (i.e. where ϕ is a sentence). To be precise, for a class \mathcal{C} defined by a sentence ϕ over the language $(\bar{a}, \bar{U}, \bar{R})$, we obtain $fc(n) = \sum_{\mathfrak{U} \subseteq [\ell_{\bar{U}}], \mathfrak{R} \subseteq [\ell_{\bar{R}}]} fc_{\mathfrak{U}, \mathfrak{R}}(n)$, where $\mathcal{C}_{\mathfrak{U}, \mathfrak{R}}$ is the class respectively defined by $\phi'_{\mathfrak{U}, \mathfrak{R}}(\bar{x})$ over the language $(\bar{a}', \bar{U}', \bar{R})$.

To sustain the induction, the above will not be enough. This is because we need to account under the model \mathfrak{N} also for the case where some variables are “assigned the value $a = a_{\ell_{\bar{a}}}$ ”, a value which does not exist in its universe (it exists only in that of \mathfrak{M}). We henceforth consider also a set $\mathfrak{X} \subseteq [\ell_{\bar{x}}]$, and denote the set of variable symbols $x_{\mathfrak{X}} = \{x_i : i \in \mathfrak{X}\}$. In our induction we will construct the expressions $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$, where $\phi'_{\mathfrak{U}, \mathfrak{R}}(\bar{x})$ is just the special case $\phi'_{\emptyset, \mathfrak{U}, \mathfrak{R}}(\bar{x})$. With models \mathfrak{M} and \mathfrak{N} as above and a fixing of the variables in $\bar{x} \setminus x_{\mathfrak{X}}$, denote by $\bar{x}_{\mathfrak{X} \rightarrow a}$ the completion of this fixing to all of \bar{x} that is obtained by fixing x_i to be equal to a for all $i \in \mathfrak{X}$. We will then have $\mathfrak{M} \models \phi(\bar{x}_{\mathfrak{X} \rightarrow a})$ if and only if $\mathfrak{N} \models \phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$.

The rest of this section is concerned with the recursive definition of $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$. There is a subsection for the base cases, a subsection for Boolean connectives, and a subsection for each class of quantifiers (first order quantifiers, counting quantifiers, and monadic second order quantifiers). In every construction we argue (at times trivially) that we keep the correspondence invariant, namely that $\mathfrak{M} \models \phi(\bar{x}_{\mathfrak{X} \rightarrow a})$ if and only if $\mathfrak{N} \models \phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ whenever \mathfrak{M} and $(\mathfrak{N}, \mathfrak{U}, \mathfrak{R})$ satisfy the correspondence condition of Definition 15.

7.2 The base constructions

We use the Boolean “true” and “false” statements in the following, so for formality’s sake they are also considered as atomic statements here. Clearly, if $\phi(\bar{x})$ is simply the “true” statement \top (respectively the “false” statement \perp), then setting $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ to \top (respectively \perp) gives us the equivalent statement satisfying the correspondence invariant.

For $i \in [\ell_{\bar{U}}]$ and $j \in [\ell_{\bar{x}}]$, let us now consider the expression $\phi(\bar{x}) = U_i(x_j)$. To produce $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$, we partition to cases according to whether $j \in \mathfrak{X}$. In the case where $j \notin \mathfrak{X}$, we simply set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = U_i(x_j)$ as well, which clearly satisfies the invariant for $(\mathfrak{N}, \mathfrak{U}, \mathfrak{R})$ correlated with \mathfrak{M} (recall that the “if and only if” condition in this case should hold when the value of x_i is in $[n + \ell_{\bar{a}} - 1]$).

Similarly, for $i \in [\ell_{\bar{U}}]$ and $j \in [\ell_{\bar{a}} - 1]$, for the expression $\phi(\bar{x}) = U_i(a_j)$, we produce $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = U_i(a_j)$, noting that in our setting the value of a_j is guaranteed to be equal to $n + j \in [n + \ell_{\bar{a}} - 1]$.

Now consider the expression $\phi(\bar{x}) = U_i(x_j)$ for the case where $x_j \in \mathfrak{X}$. Recall that in this case $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ should not depend on x_j . Moreover, to preserve the invariant for corresponding sets and models, $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ should hold if and only if $U_j(a)$ holds (recall that we use the shorthand $a = a_{\ell_{\bar{a}}}$ throughout). We hence define $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ to be \top (“true”) if $i \in \mathfrak{U}$, and define it to be \perp (“false”) if $i \notin \mathfrak{U}$.

The remaining case for a unary relation is the expression $\phi(\bar{x}) = U_i(a)$. Again, we define $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ to be \top if $i \in \mathfrak{U}$, and define it to be \perp if $i \notin \mathfrak{U}$.

We now move on to handle the atomic expressions involving a binary relation R_i where $i \in [\ell_{\bar{R}}]$. The first case here is the one analogous to the first case we discussed involving a unary relation. Namely, it is the case where $\phi(\bar{x}) = R_i(x_j, x_k)$ where both $j \notin \mathfrak{X}$ and $k \notin \mathfrak{X}$. In this case we set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = R_i(x_j, x_k)$, and argue the same argument as above about satisfying the correspondence invariant.

The next four cases we discuss resemble the last two cases we discussed about a unary relation. Namely, these are the cases where $\phi(\bar{x}) = R_i(x_j, x_k)$ with $j, k \in \mathfrak{X}$, $\phi(\bar{x}) = R_i(x_j, a)$ or $\phi(\bar{x}) = R_i(a, x_j)$ with $j \in \mathfrak{X}$, and $\phi(\bar{x}) = R_i(a, a)$. In all these cases the resulting expression should reflect on whether $\mathfrak{M} \models R_i(a, a)$, which for the corresponding $(\mathfrak{N}, \mathfrak{U}, \mathfrak{R})$ is handled by the set \mathfrak{R} . We hence set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = \top$ if $i \in \mathfrak{R}$, and set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = \perp$ if $i \notin \mathfrak{R}$.

Next we handle the cases where $\phi(\bar{x}) = R_i(x_j, x_k)$ with $j \notin \mathfrak{X}$ and $k \in \mathfrak{X}$, and $\phi(\bar{x}) = R_i(x_j, a)$ with $j \notin \mathfrak{X}$. For both these cases, for the correspondence invariant to follow we need to look at whether $\mathfrak{M} \models R_i(x_j, a)$, where the value of x_j is in $[n + \ell_{\bar{a}} - 1]$. For the corresponding $(\mathfrak{N}, \mathfrak{U}, \mathfrak{R})$ this occurs if and only if $\mathfrak{N} \models I_i(x_j)$, where we recall that I_i is a relation from $\bar{U}' \setminus \bar{U}$. We therefore set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = I_i(x_j)$ in these cases. Similarly, for the cases $\phi(\bar{x}) = R_i(a_j, x_k)$ and $\phi(\bar{x}) = R_i(a_j, a)$, where $j \in [\ell_{\bar{a}} - 1]$ and $k \in \mathfrak{X}$, we set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = I_i(a_j)$.

Moving on to the remaining cases for a binary relation, we consider $\phi(\bar{x}) = R_i(x_k, x_j)$ with $j \notin \mathfrak{X}$ and $k \in \mathfrak{X}$, and $\phi(\bar{x}) = R_i(a, x_j)$ with $j \notin \mathfrak{X}$. These are analogous to the cases handled in the last paragraph, only here we use O_i instead of I_i . We set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = O_i(x_j)$ in these two cases. Finally, for the cases $\phi(\bar{x}) = R_i(x_k, a_j)$ and $\phi(\bar{x}) = R_i(a, a_j)$, where $j \in [\ell_{\bar{a}} - 1]$ and $k \in \mathfrak{X}$, we set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = O_i(a_j)$.

The final atomic formula to consider is the ‘‘builtin relation’’ of equality. We skip all cases involving only constants (e.g. $a_i = a_j$), since these are equivalent to \top or \perp . We also skip cases that are equivalent by the symmetry of the equality relation to those that we discuss.

First, if $\phi(\bar{x})$ is $x_i = x_j$ or $x_i = a_k$ for $i, j \notin \mathfrak{X}$ and $k \in [\ell_{\bar{a}} - 1]$, then since we are dealing with values that are guaranteed to be in $[n + \ell_{\bar{a}} - 1]$ (the universe of \mathfrak{N}), we set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ respectively to $x_i = x_j$ or $x_i = a_k$ as well (so it is ‘‘unaltered’’ from $\phi(\bar{x})$).

On the other hand, if $\phi(\bar{x})$ is $x_i = x_j$ or $x_i = a$ for $i, j \in \mathfrak{X}$, then for the correspondence principle to hold, we need $\mathfrak{N} \models \phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ to hold if $\mathfrak{M} \models (a = a)$. In other words, we have to set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = \top$ here.

The final cases are those where $\phi(\bar{x})$ is $x_i = x_j$ or $x_i = a$ for $i \notin \mathfrak{X}$ and $j \in \mathfrak{X}$. For the correspondence principle to hold, we need $\mathfrak{N} \models \phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ to hold if and only if $\mathfrak{M} \models (x_i = a)$. However, we make here the subtle yet important observation that this should occur for any value that x_i can take from the universe of \mathfrak{N} , which does not include a . Therefore, we can (and should) set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = \perp$ in these cases.

7.3 Boolean connectives

Handling Boolean connectives is the most straightforward part of this construction. For example, suppose that we have $\phi(\bar{x}) = \neg\psi(\bar{x})$ for some expression $\psi(\bar{x})$, for which we have already established (by the induction hypothesis) that $\mathfrak{M} \models \psi(\bar{x}_{\mathfrak{X} \rightarrow a})$ if and only if $\mathfrak{N} \models \psi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ whenever \mathfrak{M} and $(\mathfrak{N}, \mathfrak{U}, \mathfrak{R})$ correspond. Here we can clearly set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = \neg\psi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$, and obtain that $\mathfrak{M} \models \phi(\bar{x}_{\mathfrak{X} \rightarrow a})$ if and only if $\mathfrak{N} \models \phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ whenever \mathfrak{M} and $(\mathfrak{N}, \mathfrak{U}, \mathfrak{R})$ correspond.

The same idea and analysis follow for all other Boolean connectives. For example, for the expression $\phi(\bar{x}) = \psi_1(\bar{x}) \wedge \psi_2(\bar{x})$, we set $\phi'_{\mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = \psi'_{1, \mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) \wedge \psi'_{2, \mathfrak{X}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$.

7.4 First order quantifiers

To deal with quantifiers over variables, we make some assumptions on the structure of our expressions that can easily be justified by the appropriate variable substitutions. Namely, we require that every quantified variable is quantified only once in the expression, and is not used at all outside the scope of the quantification. In particular, this means that the set \mathfrak{X} that appears in the subscript of our expression cannot contain a reference to the quantified variable.

For notational convenience, when $\phi(\bar{x})$ is our formula, we denote by $x = x_{\ell_{\bar{x}}+1}$ the quantified variable. So the two cases that we consider in this subsection are the existential quantification $\phi(\bar{x}) = \exists_x \psi(\bar{x} \cup \{x\})$ and the universal quantification $\phi(\bar{x}) = \forall_x \psi(\bar{x} \cup \{x\})$, and for both of them we would like to construct a corresponding $\phi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ where $\mathfrak{X} \subseteq [\ell_{\bar{x}}]$.

In the existential case, we want $\mathfrak{N} \models \phi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ to occur whenever there is at least one value of x for which $\mathfrak{M} \models \psi(\bar{x} \cup \{x\})$. For the values of x within $[n + \ell_{\bar{a}} - 1]$, by the induction hypothesis, this is covered by the expression $\exists_x \psi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \cup \{x\} \setminus x_{\mathfrak{X}})$. However, there is one possible value of x not covered in this way, and that is the value $n + \ell_{\bar{a}}$, which we identify with the constant a . But by the induction hypothesis, $\mathfrak{M} \models \psi(\bar{x} \cup \{x\})$ for $x = a$ if and only if $\mathfrak{N} \models \psi'_{\mathfrak{X} \cup \{\ell_{\bar{x}}+1\}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$. The combined expression that satisfies the correspondence invariant is hence $\phi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = \exists_x \psi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \cup \{x\} \setminus x_{\mathfrak{X}}) \vee \psi'_{\mathfrak{X} \cup \{\ell_{\bar{x}}+1\}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$.

The universal case follows an analogous argument, only here $\mathfrak{M} \models \psi(\bar{x} \cup \{x\})$ needs to hold for all values of x , those in $[n + \ell_{\bar{a}} - 1]$ as well as the value of a . The combined expression is $\phi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}) = \forall_x \psi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \cup \{x\} \setminus x_{\mathfrak{X}}) \wedge \psi'_{\mathfrak{X} \cup \{\ell_{\bar{x}}+1\}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$.

7.5 Modular counting quantifiers

We briefly recall the definition of a modular counting quantifier. Given $\phi(\bar{x}) = C_x^{r,m} \psi(\bar{x} \cup \{x\})$, this expression is said to hold in \mathfrak{M} for a specific assignment to the variable of \bar{x} , if the size of the set $\{x : \mathfrak{M} \models \psi(\bar{x} \cup \{x\})\}$ is congruent to r modulo m . As with the previous subsection, we assume that the quantified variable is not used outside the quantification scope, and that no variable is quantified more than once. We again denote for notational convenience the quantified variable by $x = x_{\ell_{\bar{x}}+1}$, and note that $\mathfrak{X} \subseteq [\ell_{\bar{x}}]$ cannot include a reference to x .

When working with $(\mathfrak{N}, \mathfrak{M}, \mathfrak{R})$ that corresponds to \mathfrak{M} , to obtain the original modular count, we have to count the set (satisfying the induction hypothesis) $\{x : \mathfrak{N} \models \psi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \cup \{x\} \setminus x_{\mathfrak{X}})\}$, as well as check whether $\mathfrak{N} \models \psi'_{\mathfrak{X} \cup \{\ell_{\bar{x}}+1\}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$ (which if true adds 1 to the count). This gives $(C_x^{r-1,m} \psi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \cup \{x\} \setminus x_{\mathfrak{X}}) \wedge \psi'_{\mathfrak{X} \cup \{\ell_{\bar{x}}+1\}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})) \vee (C_x^{r,m} \psi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \cup \{x\} \setminus x_{\mathfrak{X}}) \wedge \neg \psi'_{\mathfrak{X} \cup \{\ell_{\bar{x}}+1\}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}}))$ as the combined expression for $\phi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$.

7.6 Monadic second order quantifiers

Here we deal with quantifiers over unary relations. The cases we cover are the existential quantification $\phi(\bar{x}) = \exists_U \psi(\bar{x})$ and the universal quantification $\phi(\bar{x}) = \forall_U \psi(\bar{x})$, where U is a new unary relation that does not appear in the language $(\bar{a}, \bar{U}, \bar{R})$ of $\phi(\bar{x})$, while being part of the language of $\psi(\bar{x})$. As before, we assume that the quantified relation symbol U appears only in the scope of this quantification, and is not quantified anywhere else, and again denote for convenience $U = U_{\ell_{\bar{U}}+1}$. In particular, when analyzing expressions of the type $\psi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$, we may allow \mathfrak{M}' to contain $[\ell_{\bar{U}} + 1]$ (the same is not allowed for the expression $\phi'_{\mathfrak{X}, \mathfrak{M}, \mathfrak{R}}(\bar{x} \setminus x_{\mathfrak{X}})$, whose language does not contain U).

Consider now the family of possible models \mathfrak{M}' that extend \mathfrak{M} with an interpretation of the relation U . Now consider $(\mathfrak{N}', \mathfrak{M}', \mathfrak{R}')$ which correspond to \mathfrak{M}' , in relation to $(\mathfrak{N}, \mathfrak{M}, \mathfrak{R})$ which correspond to \mathfrak{M} . Referring to Definition 15, every relation already appearing in \bar{U} will

have the same interpretation in \mathfrak{M} and \mathfrak{M}' . Also, $\mathfrak{R}' = \mathfrak{R}$, since the binary relations are the same in the languages of both models. Additionally, from the definition, the interpretation of $U = U_{\ell_{\bar{U}}+1}$ in \mathfrak{M}' is the restriction of its interpretation in \mathfrak{M} to $[n + \ell_{\bar{a}} - 1]$. As for the final ingredient \mathfrak{U}' , for every $i \in [\ell_{\bar{U}}]$, the condition on whether it is in \mathfrak{U} or in \mathfrak{U}' is the same. However, \mathfrak{U}' may also include $\ell_{\bar{U}} + 1$ according to whether $\mathfrak{M}' \models U(a)$. So considering all possible models \mathfrak{M}' , we have two possibilities. Either $\mathfrak{U}' = \mathfrak{U}$, or $\mathfrak{U}' = \mathfrak{U} \cup \{\ell_{\bar{U}} + 1\}$.

We can now construct our expression that corresponds to all models extending \mathfrak{M} . For the existential case we have $\phi'_{\bar{x}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\bar{x}}) = \exists_U \psi'_{\bar{x}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\bar{x}}) \vee \exists_U \psi'_{\bar{x}, \mathfrak{U} \cup \{\ell_{\bar{U}}+1\}, \mathfrak{R}}(\bar{x} \setminus x_{\bar{x}})$, and for the universal one we have $\phi'_{\bar{x}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\bar{x}}) = \forall_U \psi'_{\bar{x}, \mathfrak{U}, \mathfrak{R}}(\bar{x} \setminus x_{\bar{x}}) \wedge \forall_U \psi'_{\bar{x}, \mathfrak{U} \cup \{\ell_{\bar{U}}+1\}, \mathfrak{R}}(\bar{x} \setminus x_{\bar{x}})$.

8 Handling relations of other arities

8.1 Nullary relations and a many-one version of the reduction

Before we consider relations of higher arities, let us show how incorporating nullary (“arity zero”) relations can replace the reduction of Theorem 6 into a many-one reduction. That is, instead of a reduction into of the original $Sp_{\phi}(n)$ into a finite sum $\sum_{i=1}^r Sp_{\phi_i}(n)$, we will have a reduction into a single $Sp_{\phi'}(n)$ where ϕ' may also involve nullary relations. Later, when we consider higher arity relations, nullary relations add much needed consistency to the notation.

Formally, for a nullary relation Z , the corresponding atomic formula is $Z()$, and a model \mathfrak{M} over a language that includes Z interprets this formula as either true or false, that is, either $\mathfrak{M} \models Z()$ or $\mathfrak{M} \models \neg Z()$.

Note that nullary relations can be simulated using higher arity relations. To replace a nullary relation Z in the language with a unary relation U (while preserving the model count), the logical expression under discussion should be conjuncted with “ $\forall_x \forall_y (U(x) \leftrightarrow U(y))$ ”, and then every instance of “ $Z()$ ” in the formula should be replaced with “ $\exists_x U(x)$ ”.

The corresponding reduction theorem is the following.

► **Theorem 17** (Many-one reduction to the case without constants). *For any class \mathcal{C} defined by an FOL (resp. MSOL, CMSOL) sentence involving a set of constant symbols \bar{a} , nullary symbols \bar{Z} , unary symbols \bar{U} and binary symbols \bar{R} , there exists a class \mathcal{C}' definable by an FOL (resp. MSOL, CMSOL) sentence involving \bar{Z}' , \bar{U}' (which contain \bar{Z} and \bar{U} respectively), \bar{R} and no constants, satisfying $f_{\mathcal{C}}(n) = f_{\mathcal{C}'}(n)$ for all $n \in \mathbb{N}$.*

Also here, the theorem follows from a single constant removal lemma, which is used for an inductive argument over $\ell_{\bar{a}}$.

► **Lemma 18** (Removing a single constant in a many-one manner). *For any class \mathcal{C} defined by an FOL (resp. MSOL, CMSOL) sentence involving a set of constant symbols \bar{a} with $\ell_{\bar{a}} > 0$, nullary symbols \bar{Z} , unary symbols \bar{U} and binary symbols \bar{R} , there exists a class \mathcal{C}' , definable by an FOL (resp. MSOL, CMSOL) sentence over the language $(\bar{a}', \bar{Z}', \bar{U}', \bar{R}')$, where $\bar{a}' = \bar{a} \setminus \{a_{\ell_{\bar{a}}}\}$, $\bar{Z}' = \bar{Z} \cup \bar{S} \cup \bar{D}$ where $\ell_{\bar{S}} = \ell_{\bar{U}}$ and $\ell_{\bar{D}} = \ell_{\bar{R}}$, $\bar{U}' = \bar{U} \cup \bar{I} \cup \bar{O}$ where $\ell_{\bar{I}} = \ell_{\bar{O}} = \ell_{\bar{R}}$, and $\bar{R}' = \bar{R}$, satisfying $f_{\mathcal{C}}(n) = f_{\mathcal{C}'}(n)$ for all $n \in \mathbb{N}$.*

The proofs are deferred to the full version of this paper [17]. We provide here a “sketch by example” on how this is done.

Adding nullary relations essentially allows us to get rid of the role of \mathfrak{U} and \mathfrak{R} , so we will only inductively construct the expressions $\phi'_{\bar{x}}$ and let nullary relations “hold the information” as to whether some relations hold with a substituted to all their variables.

If for instance we consider a binary relation $R(x, y)$ in the original vocabulary of ϕ , then the new vocabulary will include $R(x, y)$ (holding the information about the contents of R that does not involve the constant a), the unary relations $I(x)$ (holding the information about $R(x, a)$ for all $x \neq a$) and $O(y)$ (holding the information about $R(y, a)$), and the nullary relation $Z()$ that holds the information whether $R(a, a)$ holds.

In the inductive construction, $\phi'_x(\bar{x} \setminus x_x)$ will tell us whether ϕ holds where all variables enumerated by \mathfrak{X} are assigned the constant a , while all other variables are guaranteed to be different from a .

As an example, for the expression $\phi(x) = \exists_y R(x, y)$, we will have $\phi'_\emptyset(x) = (\exists_y R(x, y)) \vee I(x)$, which goes over the two options for the existence of y for which $R(x, y)$ holds: the first option being $y \neq a$, and the second one being $y = a$. Similarly we will have $\phi'_{\{1\}}() = (\exists_y O(y)) \vee Z()$, which goes over the two options for the existence of y for which $R(a, y)$ holds.

For a universal quantifier the reduction would similarly go over the two cases, but use a conjunction this time. Thus for $\psi(x) = \forall_y R(x, y)$ we will have $\psi'_\emptyset(x) = (\forall_y R(x, y)) \wedge I(x)$, and $\psi'_{\{1\}}() = (\forall_y O(y)) \wedge Z()$.

8.2 Handling higher arity relations and extended logics

The mechanism behind the proof of Theorem 17 can be extended higher arity relations, as well as more expressive logics, such as Second Order Logic (SOL) and Guarded Second Order Logic (GSOL)³.

► **Theorem 19** (Many-one reduction allowing higher arity). *For any class \mathcal{C} defined by an FOL (resp. MSOL, CMSOL, GSOL, SOL) sentence involving a set of constant symbols \bar{a} , and relation symbols \bar{R} of arbitrary arities, there exists a class \mathcal{C}' definable by an FOL (resp. MSOL, CMSOL, GSOL, SOL) sentence involving \bar{R}' , which contains \bar{R} , has the same maximum arity as \bar{R} , and has no new relations of maximum arity, satisfying $f_{\mathcal{C}}(n) = f_{\mathcal{C}'}(n)$ for all $n \in \mathbb{N}$.*

Also here, this follows by induction using a corresponding extension of Lemma 18, which allows us to eliminate hard-wired constants one at a time. As a way of demonstrating the proof of this theorem (deferred to [17]), we explain how the single elimination lemma works when considering a ternary relation $R(x, y, z)$.

As before, for eliminating a hard-wired constant a we need to add lower arity relations to R . Namely, since after the reduction $R(x, y, z)$ itself would refer only to values different from a , we add a lower arity relation for every option of substituting the value a in any of these variables. This would add $\binom{3}{1} = 3$ binary relations which we will denote here by $R_1(y, z)$, $R_2(x, z)$ and $R_3(x, y)$, in addition to $\binom{3}{2} = 3$ unary relations which we will denote by $R_{12}(z)$, $R_{13}(y)$ and $R_{23}(x)$, and a single nullary relation which we will denote by $R_{123}()$.

Thus, given for example the expression $\phi(x, y) = \exists_z R(x, y, z)$, we would have for example $\phi'_\emptyset(x, y) = (\exists_z R(x, y, z)) \vee R_3(x, y)$, $\phi'_{\{1\}}(y) = (\exists_z R_1(y, z)) \vee R_{13}(y)$, $\phi'_{\{2\}}(x) = (\exists_z R_2(x, z)) \vee R_{23}(x)$ and $\phi'_{\{1,2\}}() = (\exists_z R_{12}(z)) \vee R_{123}()$.

Going another recursion level, for $\psi(x) = \exists_{y,z} R(x, y, z) = \exists_y \phi(x)$, we would correspondingly have $\psi'_\emptyset(x) = (\exists_{y,z} R(x, y, z)) \vee (\exists_z R_2(x, z)) \vee (\exists_y R_3(x, y)) \vee R_{23}(x)$, and $\psi'_{\{1\}}() = (\exists_{y,z} R_1(y, z)) \vee (\exists_z R_{12}(z)) \vee (\exists_y R_{13}(y)) \vee R_{123}()$.

³ All discussion of these logics is deferred to [17], since the limit to the Specker-Blatter theorem discussed here only uses FOL.

9 An FOL-definable class \mathcal{C} where $f_{\mathcal{C}}(n)$ is not MC-finite

In this section we negatively settle the question of whether the Specker-Blatter theorem holds for classes whose language contains only ternary and lower-arity relations.

9.1 Using one hard-wired constant

We first construct a class whose language includes a single ternary relation and a single hard-wired constant. Our counterexample builds on ideas used in [13].

► **Theorem 20** (Ternary relation counterexample with a constant). *There exists an FOL sentence $\phi_{\mathcal{M}}$ over the language (a, R) , where a is a single (hard-wired) constant and R is a single relation of arity 3, so that the corresponding class \mathcal{C} satisfies $f_{\mathcal{C}}(n-1) = 0$ for any n that is not a power of 2, and $f_{\mathcal{C}}(n-1) \equiv 1 \pmod{2}$ for $n = 2^m$ for every $m \in \mathbb{N}$. In particular, $f_{\mathcal{C}}$ is not ultimately periodic modulo 2.*

The statement uses $f_{\mathcal{C}}(n-1)$ instead of $f_{\mathcal{C}}(n)$, but recalling the definition of $f_{\mathcal{C}}$, this refers to the universe $[n-1] \cup \{a\}$ whose size is n . We explain later how to modify this class to produce a counterexample modulo other prime numbers p instead of 2.

By Theorem 19, we have the following immediate corollary that does away with the constant, at the price of adding some additional smaller arity relations. This corollary is effectively a restatement of Theorem 8.

► **Corollary 21** (Ternary counterexample without constants). *There exists an FOL sentence $\phi'_{\mathcal{M}}$ over the language (\bar{R}) , where \bar{R} includes one relation of arity 3 and other relations of lower arities, so that the corresponding class \mathcal{C} satisfies $f_{\mathcal{C}}(n) = 0$ for every n for which $n+1$ is not a power of 2, and $f_{\mathcal{C}}(n) \equiv 1 \pmod{2}$ for $n = 2^m - 1$ for every $m \in \mathbb{N}$. In particular, $f_{\mathcal{C}}$ is not ultimately periodic modulo 2.*

At the end of this section we sketch how to further reduce the language so that it includes only one ternary relation and no lower arity relations. The full details are deferred to [17].

9.2 The first construction

The starting point of the construction is a structure that is defined over a non-constant length sequence (and hence not yet expressible in FOL) of unordered graphs. This definition follows the streamlining by Specker [25] of the original construction from [13].

► **Definition 22** (Iterated matching sequence). *Given a set V of vertices, An iterated matching sequence is a sequence of graphs over V , identified by their edge sets $\bar{E} = E_1, \dots, E_{\ell_{\bar{E}}}$, satisfying the following for every $1 \leq i \leq \ell_{\bar{E}}$.*

- *The connected components of E_i are (vertex-disjoint) complete bipartite graphs.*
- *The two vertex classes of every complete bipartite graph in E_i as above are two connected components of $\bigcup_{j=1}^{i-1} E_j$ (for $i = 1$ this means that E_1 is a matching).*
- *Every connected component of $\bigcup_{j=1}^{i-1} E_j$ is a vertex class of some bipartite graph of E_i (so in particular E_1 is a perfect matching).*

An iterated matching sequence \bar{E} is full if every vertex pair $u, v \in V$ (where $u \neq v$) appears in some E_i .

The following properties of iterated matching sequences are easily provable by induction.

► **Observation 23.** For an iterated matching \bar{E} , every E_i corresponds to a perfect matching over the set of connected components of $\bigcup_{j=1}^{i-1} E_j$. Additionally, every connected component of $\bigcup_{j=1}^i E_j$ is a clique with exactly 2^i vertices.

The above implies that there can be a full iterated matching sequence over $[n]$ if and only if n is a power of 2, in which case $\ell_{\bar{E}} = \log_2(n)$. Denoting the number of possible full iterated matching sequences over $[n]$ by $f_{\mathcal{M}}(n)$, note the following lemma.

► **Lemma 24** (see [25]). For every n which is not a power of 2 we have $f_{\mathcal{M}}(n) = 0$, while $f_{\mathcal{M}}(n) \equiv 1 \pmod{2}$ for $n = 2^m$ for every $m \in \mathbb{N}$.

The rest of this section concerns the construction of a sentence $\phi_{\mathcal{M}}$ over a language with one constant and one ternary relation, so that the corresponding class \mathcal{C} satisfies $f_{\mathcal{C}}(n-1) = f_{\mathcal{M}}(n)$. In the original construction utilizing a quaternary relation Q , essentially we had $(u, v, x, y) \in Q$ if $(u, v) \in E_i$ and $(x, y) \in E_{i-1}$ for some $1 < i \leq \ell_{\bar{E}}$, or $(u, v) \in E_1$ and $x = y$. For the construction here, we only have a ternary relation R , and we encode the placement of (u, v) within \bar{E} by the set $\{w : (u, v, w) \in R\}$. We will have to utilize the hard-wired constant a to make sure that there is exactly one way to encode every full iterated matching sequence.

9.3 Setting up and referring to an order over the vertex pairs

We simulate the structure of a full iterated matching sequence over $[n]$ (where $n \in [n]$ is identified with the constant a) by assigning “ranks” to pairs of members of $[n]$, which we consider as vertices, where each pair (x, y) is assigned the set $r_{x,y} = \{z : (x, y, z) \in R\}$. First we need to make sure that “graphness” is satisfied, which means that $r_{x,y}$ is symmetric and is empty for loops.

$$\phi_{\text{graph}} = \forall_{x,y,z} (R(x, y, z) \rightarrow (x \neq y \wedge R(y, x, z)))$$

Next we make sure that every two vertex pairs have ranks that are comparable by containment. This means that for every (x_1, y_1) and (x_2, y_2) either $r_{x_1, y_1} \subseteq r_{x_2, y_2}$ or $r_{x_2, y_2} \subseteq r_{x_1, y_1}$.

$$\phi_{\text{comp}} = \forall_{x_1, y_1, x_2, y_2} \neg \exists_{z_1, z_2} (R(x_1, y_1, z_1) \wedge \neg R(x_2, y_2, z_1) \wedge R(x_2, y_2, z_2) \wedge \neg R(x_1, y_1, z_2))$$

Finally, we want every non-loop vertex pair to have a non-empty rank, and moreover for it to include the constant a . This is crucial, because a will eventually serve as an “anchor” making sure that there is only one way to assign ranks when encoding a full iterated matching sequence using the ternary relation R .

$$\phi_{\text{full}} = \forall_{x,y} ((x \neq y) \rightarrow R(x, y, a))$$

It is a good time to sum up the full statement that sets up our pair ranks.

$$\phi_{\text{rank}} = \phi_{\text{graph}} \wedge \phi_{\text{comp}} \wedge \phi_{\text{full}}$$

Whenever this statement is satisfied, we can use it to construct expressions that compare ranks. We will use the following expressions, which compare the ranks of (x_1, y_1) and (x_2, y_2) , when we formulate further conditions on R that will eventually force it to conform to a full iterated matching sequence. Note that conveniently, these comparison expression also work against loops (whose “rank”, the empty set, is considered to be the lowest).

$$\begin{aligned} \phi_{=} (x_1, y_1, x_2, y_2) &= \forall_z (R(x_1, y_1, z) \leftrightarrow R(x_2, y_2, z)) \\ \phi_{\leq} (x_1, y_1, x_2, y_2) &= \forall_z (R(x_1, y_1, z) \rightarrow R(x_2, y_2, z)) \\ \phi_{<} (x_1, y_1, x_2, y_2) &= \phi_{\leq} (x_1, y_1, x_2, y_2) \wedge \neg \phi_{=} (x_1, y_1, x_2, y_2) \end{aligned}$$

9.4 Making the ordered pairs correspond to an iterated matching

In this subsection we consider a ternary relation R that is known to satisfy ϕ_{rank} as defined in Subsection 9.3, and impose further conditions that will force it to correspond to an iterated matching sequence (which will also be full by virtue of every pair having a rank).

For every rank appearing in R , that is for every set A which is equal to $r_{x,y}$ for some $x, y \in [n]$, we refer to the set of vertex pairs having this rank as E_i , where i is the number of ranks that appear in R (including the empty set, which is the “rank” of loops) and are strictly contained in A . So in particular $E_0 = \{(x, x) : x \in [n]\}$, and E_1 for example would be the set of vertex pairs that have the smallest non-empty set as their ranks.

We first impose the restriction that for any i , the graph defined by $\bigcup_{j=1}^i E_j$ is a transitive graph, that is a disjoint union of cliques. By Observation 23 this is a necessary condition for \bar{E} to be an iterated matching sequence (note that allowing also the 0-ranked loops does not change the condition). This is the same as saying that for any three vertices x, y, z , it cannot be the case that the rank of (x, z) is larger than the maximum ranks of (x, y) and (y, z) .

$$\phi_{\text{trans}} = \forall_{x,y,z} (\phi_{\leq}(x, z, x, y) \vee \phi_{\leq}(x, z, y, z))$$

Whenever R satisfies the above, it is not hard to add the restriction that E_i consists of disjoint complete bipartite graphs such that each of them connects exactly two components of $\bigcup_{j=1}^{i-1} E_j$, with all such components being covered. First we state that if some rank A exists, that is, there exists some (x, y) for which $A = r_{x,y}$, then every vertex z is a part of an edge with such rank.

$$\phi_{\text{cover}} = \forall_{x,y} \forall_z \exists_w \phi_{=} (x, y, z, w)$$

Then, using the prior knowledge that all connected components of both $\bigcup_{j=1}^{i-1} E_j$ and $\bigcup_{j=1}^i E_j$ are cliques, to make sure that every connected component of E_i is exactly a bipartite graph encompassing two components of $\bigcup_{j=1}^{i-1} E_j$, it is enough to state that it contains no triangles, excluding of course “triangles” of the type (x, x, x) .

$$\phi_{\text{part}} = \forall_{x,y,z} ((x \neq y) \rightarrow \neg(\phi_{=} (x, y, y, z) \wedge \phi_{=} (x, y, x, z)))$$

All of the above is sufficient to guarantee that the relation R corresponds to a full iterated matching sequence. However, as things stand now there can be many relations that correspond to the same iterated matching. This occurs because we still have unwanted freedom in choosing the sets that correspond to the possible ranks. To remove this freedom, we now require that the rank of every pair (x, y) for $x \neq y$ consists of exactly one connected component of the union of the lower ranked pairs. This will be sufficient, because by ϕ_{full} the only option for the rank would be the connected component that contains the constant a .

Noting that by ϕ_{trans} these components are cliques, it is enough to require that every member of $r_{x,y}$ is connected via a lower rank edge to a , while every vertex that is connected to a member of $r_{x,y}$ via a lower rank edge is also a member of $r_{x,y}$. We obtain the following statement.

$$\phi_{\text{anchor}} = \forall_{x,y,z} (R(x, y, z) \rightarrow (\phi_{<}(z, a, x, y) \wedge \forall_w (\phi_{<}(z, w, x, y) \rightarrow R(x, y, w))))$$

The final statement that counts the number of full iterated matching sequences, and hence provides the example proving Theorem 20 is the following.

$$\phi_{\mathcal{M}} = \phi_{\text{rank}} \wedge \phi_{\text{trans}} \wedge \phi_{\text{cover}} \wedge \phi_{\text{part}} \wedge \phi_{\text{anchor}}$$

9.5 Adapting the example to other primes

We show here how to adapt the FOL sentence from Theorem 20 to provide a sequence that is not ultimately periodic modulo p for any prime number $p \geq 2$. The analogous corollary about removing the constant also follows.

► **Theorem 25** (Ternary relation counterexample for $p \geq 2$). *For any prime number p , there exists an FOL sentence $\phi_{\mathcal{M}_p}$ over the language (a, R) , where a is a (hard-wired) constant and R is a relation of arity 3, so that the corresponding class \mathcal{C}_p satisfies $f_{\mathcal{C}_p}(n-1) = 0$ for every n that is not a power of p , and $f_{\mathcal{C}_p}(n-1) \equiv 1 \pmod{p}$ for $n = p^m$ for every $m \in \mathbb{N}$. In particular, $f_{\mathcal{C}_p}$ is not ultimately periodic modulo p .*

The construction follows the same lines as the extension from $p = 2$ to $p \geq 2$ in previous works. For completeness we give some details on how it works with respect to the version of [25]. The basic idea is to use a “matching” of p -tuples instead of pairs.

► **Definition 26.** *A p -matching over the vertex set $[n]$ is a spanning graph, each of whose connected components is either a clique with p vertices or a single vertex. A perfect p -matching is a p -matching in which there are no single vertex components (in other words, it is a partition of $[n]$ into sets of size p).*

The following is not hard to prove.

► **Lemma 27.** *There are no perfect p -matchings over $[n]$ unless n is a multiple of p , in which case their number is congruent to 1 modulo p .*

Proof. The case where n is not a multiple of p is trivial. Otherwise, consider the number of possible partitions of the set $[p]$ to a sequence of subsets of sizes i_1, \dots, i_r , where $\sum_{k=1}^r i_k = p$. Note that unless $i_1 = p$ (and hence $r = 1$), the number of such partitions is divisible by $\binom{p}{i_1}$, which is divisible by p (since p is a prime).

Denoting by $f_{M_p}(n)$ the number of perfect p -matchings over $[n]$, We consider for any p -matching its restriction to $[p]$ (which corresponds to a partition of $[p]$ – the reason we need to consider the partitions as sequences rather than as unordered families of sets is that we need to consider which sets in the restriction of the p -matching over $[n] \setminus [p]$ they are “attached” to). This implies that $f_{M_p}(n) \equiv f_{M_p}(n-p) \pmod{p}$ for every $n > p$, allowing us to prove by induction that $f_{M_p}(n) \equiv 1 \pmod{p}$ if p divides n . ◀

The definition of an iterated p -matching sequence is what one would expect.

► **Definition 28** (Iterated p -matching sequence). *Given a set V of vertices, An iterated p -matching sequence is a sequence graphs over V , identified by their edge sets $\bar{E} = E_1, \dots, E_{\ell_{\bar{E}}}$, satisfying the following for every $1 \leq i \leq \ell_{\bar{E}}$.*

- *The connected components of E_i are (vertex-disjoint) complete p -partite graphs.*
- *The p vertex classes of every complete p -partite graph in E_i as above are p connected components of $\bigcup_{j=1}^{i-1} E_j$ (for $i = 1$ this means that E_1 is a p -matching).*
- *Every connected component of $\bigcup_{j=1}^{i-1} E_j$ is a vertex class of some p -partite graph of E_i (so in particular E_1 is a perfect p -matching).*

An iterated matching sequence \bar{E} is full if every vertex pair $u, v \in V$ (where $u \neq v$) appears in some E_i .

Again we have the following properties, analogous to those of iterated matching sequences.

► **Observation 29.** For an iterated p -matching \bar{E} , every E_i corresponds to a perfect p -matching over the set of connected components of $\bigcup_{j=1}^{i-1} E_j$. Additionally, every connected component of $\bigcup_{j=1}^i E_j$ is a clique with exactly p^i vertices.

The above implies that there can be a full iterated matching sequence over $[n]$ if and only if n is a power of p , in which case $\ell_{\bar{E}} = \log_p(n)$. Denoting the number of possible full iterated matching sequences over $[n]$ by $f_{\mathcal{M}_p}(n)$, note the following lemma.

► **Lemma 30.** For every n that is not a power of p we have $f_{\mathcal{M}_p}(n) = 0$, while for $n = p^m$ for every $m \in \mathbb{N}$ we have $f_{\mathcal{M}_p}(n) \equiv 1 \pmod{p}$.

Proof. The case where n is not a power of p was already discussed above. The case $n = p^m$ is proved by induction over m using Lemma 27. ◀

From here on the construction of $\phi_{\mathcal{M}_p}$ is identical to that of $\phi_{\mathcal{M}}$ in Subsection 9.3 and Subsection 9.4, with the only exceptions being the replacements for ϕ_{cover} and ϕ_{part} .

To construct ϕ_{cover_p} , we need to state that for every existing rank, each vertex is a part of a size p clique consisting of edges from this rank.

$$\phi_{\text{cover}_p} = \forall_{x,y} \forall_{z_1} \exists_{z_2, \dots, z_p} \bigwedge_{1 \leq i < j \leq p} \phi_{=}(x, y, z_i, z_j)$$

To construct ϕ_{part_p} , we need to state that no E_i may contain a clique with $p + 1$ vertices.

$$\phi_{\text{part}_p} = \forall_{z_1, \dots, z_{p+1}} ((z_1 \neq z_2) \rightarrow \neg(\bigwedge_{1 \leq i < j \leq p+1} \phi_{=}(z_1, z_2, z_i, z_j)))$$

The final expression is the following.

$$\phi_{\mathcal{M}_p} = \phi_{\text{rank}} \wedge \phi_{\text{trans}} \wedge \phi_{\text{cover}_p} \wedge \phi_{\text{part}_p} \wedge \phi_{\text{anchor}}$$

9.6 Reducing the example further to have a single relation

We provide here a sketch on how to produce, starting with Theorem 8, a sentence with a single relation that provides a class that is not MC-finite.

► **Theorem 31 (A sentence with a single relation).** For every prime number $p \geq 1$ there exists an FOL-sentence ϕ_p over a language consisting of a single relation of arity 3, so that for the class \mathcal{C} corresponding to ϕ_p , its counting function $f_{\mathcal{C}}(n)$ is not ultimately periodic modulo p .

Starting with an expression ϕ that results from invoking Theorem 25 over $\phi_{\mathcal{M}_p}$, we explain how to reduce it further to an expression that involves a single ternary relation. This transformation is ad-hoc and uses certain specific features and symmetries of models satisfying $\phi_{\mathcal{M}_p}$, and their reflection in the corresponding models satisfying ϕ . The full details require delving into the specifics of the proof of Theorem 19, and are deferred to [17].

As $\phi_{\mathcal{M}_p}$ involves a single ternary relation R and a single constant a , the resulting ϕ involves a corresponding ternary relation, as well as three binary relations, three unary relations, and a single nullary relation. Since the lower order relations result from substituting the constant at some of the places of the relation R (while restricting the other places to hold values different from a), looking at the working of $\phi_{\mathcal{M}_p}$ allows us to immediately rule out most options for the lower arity relations.

For example, the nullary relation would correspond to whether $R(a, a, a)$ holds, so it must evaluate to \perp (“false”) for all models of ϕ (since $\phi_{\mathcal{M}_p}$ implies $\neg R(x, x, y)$ for all x and y , equal or unequal to a). Thus we may just remove it and replace its occurrences in ϕ with the \perp symbol.

Similar considerations allow us to eliminate the unary relation corresponding to $R(a, a, x)$ for $x \neq a$ (always false), and the unary relations corresponding to $R(a, x, a)$ and $R(x, a, a)$ (always true by ϕ_{full} since $x \neq a$).

Next, the binary relation corresponding to $R(x, y, a)$ for $x, y \neq a$, while not constant, can be “fully deduced” and replaced with $x \neq y$ by ϕ_{full} and ϕ_{graph} .

This leaves us with the two relations corresponding to $R(x, a, y)$ and $R(a, x, y)$. By first noting that for satisfying models they are equal to each other (by ϕ_{graph}), we can reduce them to a single relation. In the final step, we “fold” this relation into the ternary relation, after noting that R has to satisfy $\neg R(x, x, y)$ for all x, y . The last operation requires first replacing the occurrences of $R(x, y, z)$ in the sentence with “ $(x \neq y) \wedge R(x, y, z)$ ”, which now “frees” this part of R to be used instead of the binary relation. This does not change the number of satisfying models, since we make sure that this “region” of R is completely used through a bijection for the role of the binary relation that it replaces.

10 Conclusions and open problems

In this work we have extended the Specker-Blatter Theorem to classes of τ -structures definable in CMSOL for vocabularies τ which contain a finite number of hard-wired constants, unary and binary relation symbols, Corollary 7. We have also shown that it does not hold already when τ consists of only one ternary relation symbol, Theorem 31. We note that in [15, 16] we have shown that for \mathcal{C} definable in CMSOL such that all structures have degree bounded by a constant d , $S_{\mathcal{C}}(n)$ is always MC-finite. The degree of a structure \mathcal{A} is defined via the Gaifman graph of \mathcal{A} . With this the MC-finiteness of $S_{\mathcal{C}}(n)$ for CMSOL-definable classes of τ -structures as a function of τ is completely understood. Applications of our results in this paper to restricted Bell numbers and various restricted partition functions are given in [11].

A sequence of integers $s(n)$ is MC-finite if for every $m \in \mathbb{N}^+$ there are constants $r(m), p(m) \in \mathbb{N}^+$ and coefficients $\alpha_1(m), \dots, \alpha_{p(m)} \in \mathbb{N}^+$, such that for all $n \geq r(m)$ we have

$$s(n + p(m) + 1) \equiv \sum_{i=0}^{p(m)} \alpha_i s(i) \pmod{m}.$$

The Specker-Blatter Theorem gives little information on the constants $r(m), p(m)$ or the coefficients $\alpha_1(m), \dots, \alpha_{p(m)}$. These in particular depend on the substitution rank of the class \mathcal{C} . In fact Theorem 3 gives a very bad estimate of the substitution rank in the case of binary relation symbols. The constants are computable, but it is not known whether they are always computable in feasible time or whether their size is bounded by an elementary function. In the presence of constants the substitution rank is not defined. Our main Theorem 12 allows to eliminate the constants, and therefore gives a formula for which the substitution rank is defined. However, due to the increased complexity of the resulting formula, the estimate of the substitution rank will be even worse.

► **Problem 32.** *Given a sentence ϕ in CMSOL(τ) where τ consists only of constants, unary and binary relation symbols,*

- (i) *what is the time complexity of computing the constants $r(m), p(m)$ and the coefficients $\alpha_1(m), \dots, \alpha_{p(m)}$?*
- (ii) *what can we say about the size of these constants?*

The proof of Theorem 3 depends on the Feferman-Vaught Theorem which also holds for CMSOL(τ) for any finite relational τ , [10, 22]. In our context, the Feferman-Vaught Theorem allows to check whether a formula of CMSOL(τ) holds in $\text{Subst}(\mathcal{A}_1, a, \mathcal{A}_2)$ by checking a

sequence of CMSOL(τ)-formulas in \mathcal{A}_1 and \mathcal{A}_2 independently. This sequence is called a reduction sequence, cf. [14]. In [6] it is shown that even for FOL(τ) the size of the reduction sequences for the Feferman-Vaught Theorem cannot, in general, be bounded by an elementary function.

► **Problem 33.** *Does there exist an elementary function $f(k)$, so that for any sentence ϕ in CMSOL(τ) where τ consists only of constants, unary and binary relation symbols, the size of the constants $r(m)$ and $p(m)$ is bounded by $f(\max\{|\phi|, m\})$?*

The Specker-Blatter Theorem also applies to hereditary, monotone and minor-closed graph classes, provided they are definable using a finite set of forbidden (induced) subgraphs or minors. In the first two cases such a class is FOL-definable. In the case of a minor-closed class, B. Courcelle showed that it is MSOL-definable, see [5]. By the celebrated theorem of N. Robertson and P. Seymour, [7], every minor-closed class of graphs is definable by a finite set of forbidden minors. However, there are monotone (hereditary) classes of graphs where a finite set of forbidden (induced) subgraphs does not suffice.

► **Problem 34.** *Are there hereditary or monotone classes of graphs \mathcal{C} such that $Sp_{\mathcal{C}}(n)$ is not MC-finite?*

An analogue question arises when we replace graphs by finite relational τ -structures. In this case one speaks of classes of τ -structures closed under substructures. Every class of finite τ -structures \mathcal{C} closed under substructures can be characterized by a set of forbidden substructures. If this set is finite, \mathcal{C} is again FOL-definable, and the Specker-Blatter Theorem applies.

► **Problem 35.**

- (i) *Let τ be a relational vocabulary. Are there substructure closed classes \mathcal{C} of τ -structures such that $Sp_{\mathcal{C}}(n)$ is not MC-finite?*
- (ii) *Same question when all the relations are at most binary?*

References

- 1 C. Blatter and E. Specker. Le nombre de structures finies d'une théorie à caractère fini. *Sciences Mathématiques, Fonds Nationale de la recherche Scientifique, Bruxelles*, pages 41–44, 1981.
- 2 C. Blatter and E. Specker. Modular periodicity of combinatorial sequences. *Abstracts of the AMS*, 4:313, 1983.
- 3 C. Blatter and E. Specker. Recurrence relations for the number of labeled structures on a finite set. In E. Börger, G. Hasenjaeger, and D. Rödding, editors, *In Logic and Machines: Decision Problems and Complexity*, volume 171 of *Lecture Notes in Computer Science*, pages 43–61. Springer, 1984.
- 4 Andrei Z Broder. The r-stirling numbers. *Discrete Mathematics*, 49(3):241–259, 1984.
- 5 B. Courcelle and J. Engelfriet. *Graph Structure and Monadic Second-order Logic, a Language Theoretic Approach*. Cambridge University Press, 2012.
- 6 Anuj Dawar, Martin Grohe, Stephan Kreutzer, and Nicole Schweikardt. Model theory makes formulas large. In *International Colloquium on Automata, Languages, and Programming*, pages 913–924. Springer, 2007.
- 7 R. Diestel. *Graph Theory*. Graduate Texts in Mathematics. Springer, 3 edition, 2005.
- 8 H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer Verlag, 1995.
- 9 Graham Everest, Alfred J van der Poorten, Igor Shparlinski, Thomas Ward, et al. *Recurrence sequences*, volume 104. American Mathematical Society Providence, RI, 2003.

- 10 S. Feferman and R. Vaught. The first order properties of algebraic systems. *Fundamenta Mathematicae*, 47:57–103, 1959.
- 11 Yuval Filmus, Eldar Fischer, Johann A. Makowsky, and Vsevolod Rakita. MC-finiteness of restricted set partitions, 2023. [arXiv:2302.08265](#).
- 12 E. Fischer. The Specker-Blatter theorem does not hold for quaternary relations. *Journal of Combinatorial Theory, Series A*, 103:121–136, 2003.
- 13 E. Fischer. The Specker-Blatter theorem does not hold for quaternary relations. *Journal of Combinatorial Theory, Series A*, 103:121–136, 2003.
- 14 E. Fischer, T. Kotek, and J.A. Makowsky. Application of logic to combinatorial sequences and their recurrence relations. In M. Grohe and J.A. Makowsky, editors, *Model Theoretic Methods in Finite Combinatorics*, volume 558 of *Contemporary Mathematics*, pages 1–42. American Mathematical Society, 2011.
- 15 E. Fischer and J. A. Makowsky. The Specker-Blatter theorem revisited. In *COCOON*, volume 2697 of *Lecture Notes in Computer Science*, pages 90–101. Springer, 2003.
- 16 Eldar Fischer, Tomer Kotek, and Johann A Makowsky. Application of logic to combinatorial sequences and their recurrence relations. *Model Theoretic Methods in Finite Combinatorics*, 558:1–42, 2011.
- 17 Eldar Fischer and Johann A. Makowsky. Extensions and limits of the specker-blatter theorem, 2022. [arXiv:2206.12135](#).
- 18 Ronald L Graham, Donald E Knuth, and Oren Patashnik. *Concrete mathematics: a foundation for computer science*. Addison-Wesley, 1989.
- 19 Manuel Kauers and Paule. *The Concrete Tetrahedron: Symbolic Sums, Recurrence Equations, Generating Functions, Asymptotic Estimates*. Springer, 2011.
- 20 Thomas Koshy. *Catalan numbers with applications*. Oxford University Press, 2008.
- 21 L. Libkin. *Elements of Finite Model Theory*. Springer, 2004.
- 22 J.A. Makowsky. Algorithmic uses of the Feferman-Vaught theorem. *Annals of Pure and Applied Logic*, 126.1-3:159–213, 2004.
- 23 Götz Pfeiffer. Counting transitive relations. *Journal of Integer Sequences*, 7(2):3, 2004.
- 24 James A Reeds and Neil JA Sloane. Shift register synthesis (modulo m). *SIAM Journal on Computing*, 14(3):505–513, 1985.
- 25 E. Specker. Modular counting and substitution of structures. *Combinatorics, Probability and Computing*, 14:203–210, 2005.
- 26 Ernst Specker. Application of logic and combinatorics to enumeration problems. In *Ernst Specker Selecta*, pages 324–350. Springer, 1990.