Vectorization hierarchies of some graph quantifiers

Lauri Hella *
Department of Mathematics
P.O. Box 4 (Yliopistonkatu 5)
00014 University of Helsinki
Finland
email: lauri.hella@helsinki.fi

Juha Nurmonen †
Department of Mathematics
and Computer Science
University of Leicester
University Road
Leicester LE1 7RH
United Kingdom
email: j.nurmonen@mcs.le.ac.uk

Abstract

We give a sufficient condition for the inexpressibility of the k-th extended vectorization of a generalized quantifier Q in $\mathrm{FO}(\mathbf{Q}_k)$, the extension of first-order logic by all k-ary quantifiers. The condition is based on a model construction which, given two $\mathrm{FO}(\mathbf{Q}_1)$-equivalent models with certain additional structure, yields a pair of $\mathrm{FO}(\mathbf{Q}_k)$-equivalent models. We also consider some applications of this condition to quantifiers that correspond to graph properties, such as connectivity and planarity.

1 Introduction

A lot of research in finite model theory has lately focused on its connections with computational complexity theory. The starting point of research in this area was Fagin's theorem [Fag74] which states that a property of finite structures is in NP if and only if it is expressible in existential second-order logic Σ^1_1. After this result practically all other central complexity classes have been characterized by expressibility in some logic, although such characterizations are usually restricted to the case of ordered finite structures. The problem whether such a characterization is possible on unordered finite structures remains open unfortunately often. In particular, this is the case for perhaps the most important complexity class PTIME. On ordered structures PTIME is characterizated by expressibility in least fixed point logic ([Imm86, Var82]).

*Partially supported by Grant 40734 from the Academy of Finland. The research was initiated while the first author was supported by a grant from the University of Helsinki.

†Supported by EPSRC grant GR/K96564. This research was initiated when the second author was a student at the Department of Mathematics, University of Helsinki and supported by the Academy of Finland.

Mathematics Subject Classification: 03C13, 03C80

Keywords: Ehrenfeucht-Fraïssé games, finite model theory, generalized quantifiers, vectorization

It is well known that, for many purposes, the expressive power of first-order logic FO on finite structures is too weak. Therefore finite model theory is mainly concerned with the study of various extensions of FO. Generalized quantifiers provide a simple and general way of enhancing the expressive power of FO. Indeed, according to the definition of Lindström [Lin66], any class $\mathcal{C}$ of structures which is closed under isomorphisms can be taken as the interpretation of a generalized quantifier $\mathsf{Q}_\mathcal{C}$. The corresponding logic $\mathrm{FO}(\mathsf{Q}_\mathcal{C})$ is then the least extension of FO with the substitution property in which the class $\mathcal{C}$ is definable (see [Ebb85] for a definition of the substitution property). Furthermore, if L is any logic such that every formula of L refers to at most finitely many symbols in the vocabulary, then there is a set $\mathbf{Q}$ of quantifiers such that L and $\mathrm{FO}(\mathbf{Q})$ have the same expressive power.

As $\mathrm{FO}(\mathbf{Q})$ is a minimal well behaved logic in which $\mathcal{C}$ is definable for each $\mathsf{Q}_\mathcal{C} \in \mathbf{Q}$, it is no wonder that, for finite sets $\mathbf{Q}$, the expressive power of $\mathrm{FO}(\mathbf{Q})$ is still severely restricted. Especially, it is known that there is no finite set $\mathbf{Q}$ of quantifiers such that $\mathrm{FO}(\mathbf{Q})$ (or even the extension of fixed point logic by $\mathbf{Q}$) captures PTIME on unordered structures ([Hel96, DH95]). The main reason for the weakness of $\mathrm{FO}(\mathbf{Q})$ has turned out to be in the fact that $\mathrm{FO}(\mathbf{Q})$ lacks certain natural closure properties, which many other logics (including first-order logic and fixed point logics) studied in finite model theory have. Most notably, as a rule $\mathrm{FO}(\mathbf{Q})$ is not closed under *vectorization*. Also, $\mathrm{FO}(\mathbf{Q})$ is often not *congruence closed* or closed under *relativization*. (See Section 2.3 for the definitions of these notions.)

Fortunately there is a simple modification of $\mathrm{FO}(\mathbf{Q})$ which is closed under vectorization: just replace $\mathbf{Q}$ by the set of all vectorizations Q^k, $k \geq 1$, of the quantifiers Q in $\mathbf{Q}$. We use the notation $\mathrm{FO}^{(\omega)}(\mathbf{Q})$ for this modified logic. Similarly, it is straightforward to define a variant of $\mathrm{FO}(\mathbf{Q})$ which is congruence closed, closed under relativization, or which has any combination of these three closure properties. In particular, we will consider the *extended vectorization* $\mathrm{FO}_*^{(\omega)}(\mathbf{Q})$ of $\mathrm{FO}(\mathbf{Q})$ which is the least extension of $\mathrm{FO}(\mathbf{Q})$ that is closed under all the three properties. This logic is generated by the set of all extended vectorizations Q_*^k, $k \geq 1$, of quantifiers Q in $\mathbf{Q}$ (see Section 2.3 for a definition).

The significance of vectorization has been recognized from the beginning in studies concerning logical characterizations of complexity classes. Immerman [Imm86, Imm87] defined three different versions of transitive closure operator, and proved that they capture LOGSPACE, NLOGSPACE and PTIME, respectively. Later Stewart [Ste91, Ste92] proved that the Hamiltonian path operator, as well as some other NP-complete operators, captures NP. In all these characterizations the corresponding operators are used in all *arities*, i.e., all vectorizations of the operators are needed. In some recent papers dealing with generalized quantifiers on finite structures the assumption of being closed under vectorizations has even been taken as a part of the definition of logics with generalized quantifiers (see, e.g., [MP95]).

Vectorization also remedies the defect of $\mathrm{FO}(\mathbf{Q})$ that prevented capturing PTIME by finitely many generalized quantifiers: in [Daw95] Dawar proved that if there is any reasonable logic that captures PTIME on unordered structures, then there is a single quantifier Q such that $\mathrm{FO}^{(\omega)}(\mathsf{Q})$ captures PTIME.

Given a quantifier Q, it is natural to ask whether its vectorization hierarchy is unbounded, i.e., whether there are arbitrarily large natural numbers k such that Q^k is not expressible in $\mathrm{FO}(\{\mathsf{Q}^l \mid l < k\})$. Similarly, one can ask whether the extended vectorization hierarchy Q^k_*, $k < \omega$, of Q is unbounded. In some interesting cases the answer is known to be "yes". In particular, this holds for the quantifiers Q_{TC} and $\mathsf{Q}_{\mathrm{HAM}}$ which correspond to the transitive closure and hamiltonian path operators, respectively ([GH96]). There are also some interesting examples of quantifiers whose vectorization hierarchy collapses; for example this is the case for the even cardinality quantifier (cf. [Mos]). However, it would be desirable to find general criteria for the unboundedness of the vectorization hierarchy of a quantifier.

In this paper we provide a new method for proving the unboundedness of the extended vectorization hierarchy of a quantifier. As a first step towards this goal we prove in Section 3 that if Q is not definable in the extension of first-order logic by all unary quantifiers, and if this can be shown by using 1-bijective Ehrenfeucht-Fraïssé games and counterexample models of a special type, then Q^k_* is not definable in $\mathrm{FO}(\mathbf{Q}_k)$, the extension of first-order logic by all k-ary quantifiers. The proof of this result is based on a model construction which, given a pair $\mathbb{A}, \mathbb{B}$ of models that are $\mathrm{FO}(\mathbf{Q}_1)$-equivalent up to quantifier rank kr in the presence of certain additional structure, yields a pair $\mathbb{A}_k, \mathbb{B}_k$ of models that are $\mathrm{FO}(\mathbf{Q}_k)$-equivalent up to quantifier rank r.

Although the requirements for this additional structure are apparently quite restrictive, we show that the method is useful and reasonably general by giving a whole class of applications. More precisely, in Section 4.1 we prove that if $\mathcal{C}$ is any non-trivial class of graphs such that both $\mathcal{C}$ and its complement $\overline{\mathcal{C}}$ are closed under stretching edges to paths, then there are counterexample graphs showing that the corresponding graph quantifier $\mathsf{Q}_{\mathcal{C}}$ is not definable in $\mathrm{FO}(\mathbf{Q}_1)$ even in the presence of the required additional structure. Hence, for such classes $\mathcal{C}$, the k-th extended vectorization of $\mathsf{Q}_{\mathcal{C}}$ is not definable in $\mathrm{FO}(\mathbf{Q}_k)$, and consequently, the extended vectorization hierarchy of $\mathsf{Q}_{\mathcal{C}}$ is unbounded. In Section 4.2 we apply this result to some concrete graph quantifiers, such as $\mathsf{Q}_{\mathrm{CONN}}$ and $\mathsf{Q}_{\mathrm{PLAN}}$, where CONN is the class of all connected graphs and PLAN is the class of all planar graphs.

2 Preliminaries

A vocabulary $\sigma = \{R_1, \ldots, R_s\}$ is a finite set of relation symbols each with a fixed arity $\mathrm{ar}(R_i)$. Our main interest in this paper concerns finite structures. Thus in any σ-structure $\mathbb{A} = (A, R_1(\mathbb{A}), \ldots, R_s(\mathbb{A}))$, the *universe* A of $\mathbb{A}$ is assumed to be finite.

For example, (undirected) graphs $\mathbb{G} = (G, E(\mathbb{G}))$ are structures over a vocabulary that consists of a single binary relation symbol E, where $E(\mathbb{G})$ is irreflexive and symmetric, i.e., for all vertices $u \in G$ we have $(u, u) \notin E(\mathbb{G})$ and $(u, v) \in E(\mathbb{G})$ implies $(v, u) \in E(\mathbb{G})$. If symmetry is not required, then $\mathbb{G}$ is a directed graph.

The definitions of syntax and semantics of first-order logic FO are standard. Equality is treated as a special relation symbol, which is not a member of the vocabulary. The fragment of FO consisting of formulas of FO with quantifier rank at most r, is denoted by FO_r (recall that the quantifier rank of a formula is the depth of nesting of quantifiers).

In *infinitary logic* $\mathcal{L}_{\infty\omega}$ also infinite disjunctions and conjunctions are allowed. The expressive power of this logic is too strong on finite structures, since any class of finite structures is definable in $\mathcal{L}_{\infty\omega}$. The interest of this logic comes from the fragment $\mathcal{L}^k_{\infty\omega}$, where only k distinct variables, free or bound, are allowed. The *finite variable logic* $\mathcal{L}^\omega_{\infty\omega}$ is the union of the logics $\mathcal{L}^k_{\infty\omega}$ over all natural numbers k. It is well known that $\mathcal{L}^k_{\infty\omega}$

has stronger expressive power than FO. For example, the class CONN of connected finite graphs is expressible in $\mathcal{L}^{\omega}_{\infty\omega}$ but not in FO (see e.g. [EF95]).

We use the standard notation for comparing the expressive power of logics: if L and L' are logics, then $L \leq L'$ means that any class $\mathcal{C}$ of (finite) structures that is definable in L is also definable in L'. Furthermore, we write $L \equiv L'$ if L and L' have equal expressive power, i.e., if both $L \leq L'$ and $L' \leq L$.

2.1 Generalized quantifiers

Let L be a logic and $\overline{n} = (n_1, \dots, n_k)$ a sequence of natural numbers. A structure is of type $\overline{n}$ if it is of the form $\mathbb{A} = (A, R_1(\mathbb{A}), \dots, R_k(\mathbb{A}))$, where $R_i(\mathbb{A}) \subseteq A^{n_i}$ for each $i = 1, \dots, k$. Let $\mathcal{C}$ be a class of structures of type $\overline{n}$ which is closed under isomorphisms. Then $\mathcal{C}$ can be associated with a generalized quantifier $\mathsf{Q}_{\mathcal{C}}$. The set of formulas of the logic $L(\mathsf{Q}_{\mathcal{C}})$ is defined as for the logic L with the following additional rule:

> If $\psi_1, \dots, \psi_k$ are formulas and $\overline{x}_i$ are n_i-tuples of distinct variables for $i = 1, \dots, k$, then $\mathsf{Q}_{\mathcal{C}}\overline{x}_1, \dots, \overline{x}_k(\psi_1, \dots, \psi_k)$ is a formula.

Thus $\mathsf{Q}_{\mathcal{C}}$ is an operator that binds k formulas together, and n_i variables in the i-th formula. In particular, a free occurrence of a variable x in one of the formulas ψ_i remains free in $\mathsf{Q}_{\mathcal{C}}\overline{x}_1, \dots, \overline{x}_k(\psi_1, \dots, \psi_k)$ unless x is contained in the tuple $\overline{x}_i$ that binds ψ_i. The semantics of the quantifier $\mathsf{Q}_{\mathcal{C}}$ is determined by the class $\mathcal{C}$ as follows:

$$\mathbb{A} \models \mathsf{Q}_{\mathcal{C}}\overline{x}_1, \dots, \overline{x}_k(\psi_1(\overline{x}_1, \overline{b}_1), \dots, \psi_k(\overline{x}_k, \overline{b}_k)) \text{ if and only if}$$
$$(A, \psi_1(\mathbb{A}, \overline{b}_1), \dots, \psi_k(\mathbb{A}, \overline{b}_k)) \in \mathcal{C}, \text{ where}$$
$$\psi_i(\mathbb{A}, \overline{b}_i) = \{\overline{a} \in A^{n_i} \mid \mathbb{A} \models \psi_i(\overline{a}, \overline{b}_i)\} \text{ for each } i = 1, \dots, k.$$

Here $\overline{b}_i$ is a tuple of parameters that gives the interpretation for those free variables of ψ_i which are not contained in $\overline{x}_i$. For a set $\mathbf{Q}$ of generalized quantifiers, the logic $L(\mathbf{Q})$ is defined similarly with the corresponding rule for each quantifier in $\mathbf{Q}$.

Note that the existential quantifier can be identified with the quantifier $\mathsf{Q}_{\mathcal{C}}$, where $\mathcal{C}$ is the class of all structures (A, P) such that $\emptyset \neq P \subseteq A$.

2.1 Example. (*a*) The *Rescher quantifier* R is defined to be the class of all structures (A, P, S), where $P, S \subseteq A$ and $|P| \leq |S|$. In the definition of the *Härtig quantifier* H, the last condition is replaced by $|P| = |S|$. These two quantifiers are of type $(1, 1)$. Since in FO(R) and FO(H) it is possible to compare cardinalities, these logics are stronger in expressive power than first-order logic.

(*b*) For any property of finite graphs there is a quantifier of type (2) which expresses that property. For instance, the PTIME computable property of being a planar graph is captured by the quantifier $\mathsf{Q}_{\mathrm{PLAN}}$, where PLAN is the class of all finite planar graphs. Similarly we can define quantifiers $\mathsf{Q}_{\mathrm{CONN}}$ and $\mathsf{Q}_{\mathrm{3COL}}$ which capture the notions of connectivity and 3-colorability of finite graphs, respectively.

(*c*) Let TC to be the class of all structures (A, R_1, R_2), where $R_1, R_2 \subseteq A^2$ and R_2 contains a single pair $(a_1, a_2) \in A^2$ which is in the transitive closure of the relation R_1. With the corresponding quantifier Q_{TC} of type $(2, 2)$ it is easy to define the transitive closure of any binary relation.

The *arity* of a quantifier $\mathsf{Q}_{\mathcal{C}}$ of type $(n_1, \ldots, n_k)$ is defined by

$$\mathrm{ar}(\mathsf{Q}_{\mathcal{C}}) = \max\{n_1, \ldots, n_k\}.$$

A quantifier $\mathsf{Q}_{\mathcal{C}}$ is n-ary if $\mathrm{ar}(\mathsf{Q}_{\mathcal{C}}) \leq n$. We denote the set of all n-ary quantifiers by $\mathbf{Q}_n$. The quantifier rank $qr(\varphi)$ of an $L(\mathbf{Q})$ formula φ is defined as for the logic L with the following additional rule:

$$\mathrm{qr}(\mathsf{Q}_{\mathcal{C}}\overline{x}_1, \ldots, \overline{x}_k(\varphi_1, \ldots, \varphi_k)) = \max\{\mathrm{qr}(\varphi_i) \mid 1 \leq i \leq k\} + 1 \text{ for each } \mathsf{Q}_{\mathcal{C}} \in \mathbf{Q}.$$

If $\mathbb{A}$ and $\mathbb{B}$ are models of the same vocabulary, which satisfy exactly the same $\mathrm{FO}(\mathbf{Q}_n)$ -sentences of quantifier rank at most r, then we write $\mathbb{A} \equiv \mathbb{B}$ $(\mathrm{FO}_r(\mathbf{Q}_n))$. Correspondingly, if $\mathbb{A}$ and $\mathbb{B}$ satisfy the same sentences of $\mathcal{L}^k_{\infty\omega}(\mathbf{Q}_n)$, we write $\mathbb{A} \equiv \mathbb{B}$ $(\mathcal{L}^k_{\infty\omega}(\mathbf{Q}_n))$.

In [Hel89, Hel96] Hella considered the following bijective Ehrenfeucht-Fraïssé -games. Let $\mathbb{A}$ and $\mathbb{B}$ be structures and assume positive integers r and n are given. The rules of the *n-bijective r-round game* $\mathrm{BEF}^r_n(\mathbb{A}, \mathbb{B})$ are as follows. There are two players, the spoiler and the duplicator. In each round $1 \leq i \leq r$ the duplicator selects a bijection $f_i : A \to B$ and the spoiler answers by choosing a subset $C_i \subseteq A$ with $|C_i| \leq n$. The duplicator wins this game if in each round $1 \leq i \leq r$, $p_i = \bigcup_{1 \leq j \leq i} f_j \restriction C_j$ is a partial isomorphism $\mathbb{A} \to \mathbb{B}$. Observe that for each $j < r$, the bijection f_{j+1} chosen by the duplicator must extend the partial isomorphism p_j; otherwise the spoiler wins in one move by choosing an element a from the domain of p_j such that $p_j(a) \neq f_{j+1}(a)$. The interest of these games comes from the following result.

2.2 THEOREM. ([HEL89, HEL96]) *Let $\mathbb{A}$ and $\mathbb{B}$ be structures of the same vocabulary. The duplicator has a winning strategy in $BEF^r_n(\mathbb{A}, \mathbb{B})$ if and only if $\mathbb{A} \equiv \mathbb{B}$ $(FO_r(\mathbf{Q}_n))$.*

The corresponding pebble game characterization for $\mathcal{L}^k_{\infty\omega}(\mathbf{Q}_n)$ goes as follows ([Hel96]). Consider two σ-structures $\mathbb{A}$ and $\mathbb{B}$. In the *n-bijective k-pebble game* $\mathrm{BP}^k_n(\mathbb{A}, \mathbb{B})$, k corresponds the number of variables and n the maximal arity of the quantifiers. After each round j the moves of the players determine a partial isomorphism $p_j : \mathbb{A} \to \mathbb{B}$. In the first round, the duplicator selects a bijection $f_1 : A \to B$ and the spoiler answers by choosing a subset C_1 with $|C_1| \leq n$. If $p_1 = f_1 \restriction C_1$ is not a partial isomorphism $\mathbb{A} \to \mathbb{B}$, the spoiler wins. Otherwise the game continues. In general, in round $j \geq 1$, assume that $p_j : \mathbb{A} \to \mathbb{B}$ is a partial isomorphism. Again the duplicator selects a bijection $f_{j+1} : A \to B$ which extends p_j. Now the spoiler chooses subsets $D_j \subseteq \mathrm{dom}(p_j)$ and $C_{j+1} \subseteq A$ such that $|C_{j+1}| \leq n$ and $|D_j \cup C_{j+1}| \leq k$. The function p_{j+1} is defined to be $p_j \restriction D_j \cup f_{j+1} \restriction C_{j+1}$. If p_{j+1} is not a partial isomorphism $\mathbb{A} \to \mathbb{B}$, the spoiler wins; otherwise the game goes on. In other words, if p_j is not a partial isomorphism after some round $j < \omega$, the spoiler wins — otherwise the duplicator wins. The following theorem describes the role of these games.

2.3 THEOREM. ([HEL96]) *Let $\mathbb{A}$ and $\mathbb{B}$ be structures of the same vocabulary. The duplicator has a winning strategy in $BP^k_n(\mathbb{A}, \mathbb{B})$ if and only if $\mathbb{A} \equiv \mathbb{B}$ $(\mathcal{L}^k_{\infty\omega}(\mathbf{Q}_n))$.*

2.4 REMARK. Immerman and Lander [IL90] defined a pebble game which characterizes equivalence with respect to the logic $\mathcal{L}^k_{\infty\omega}(\mathbf{C})$, where $\mathbf{C}$ is the set of all counting quantifiers 'there are at least m elements' for each natural number m. It can be seen that the game of Immerman and Lander is equivalent to the bijective game BP^k_1. Hence for every k, $\mathcal{L}^k_{\infty\omega}(\mathbf{Q}_1)$ has the same expressive power as the logic $\mathcal{L}^k_{\infty\omega}(\mathbf{C})$ (see also [KV95]).

2.2 Hanf's technique and unary quantifiers

Hanf [Han65] introduced a technique based on the number of local isomorphism types to guarantee elementary equivalence of two structures (finite or infinite) with respect to first-order logic. Fagin, Stockmeyer and Vardi [FSV95] formulated this technique in a form which is better suited for finite model theory. In [Nur96] it was shown that with this technique one gets structures that are equivalent with respect to $\mathrm{FO}_r(\mathbf{Q}_1)$. In the following we explain shortly this technique.

Let $\mathbb{A}$ be a σ-structure and $a, b \in A$. Elements a and b are *adjacent*, if there is some $R_i \in \sigma$ and a tuple $\bar{t} \in R_i(\mathbb{A})$ such that a and b are components of $\bar{t}$. The *degree* $deg(a)$ of an element a is the number of elements adjacent to a but not equal to a. Whenever $X \subseteq A$, $\mathbb{A}\restriction X$ is the structure with universe X, where the interpretation of R_i is the set of tuples $\bar{t}$ in $R_i(\mathbb{A})$ such that every component of $\bar{t}$ is in X, for $1 \leq i \leq s$.

Consider a σ-structure $\mathbb{A}$. The *neighborhood* $N(e, a)$ *of radius* e *of* $a \in A$ is defined recursively by

$$N(1, a) = \{a\};$$

$$N(e+1, a) = \{v \mid v \text{ is adjacent to some } b \in N(e, a)\} \cup N(e, a).$$

The *e-type* of an element a is the isomorphism type of $(\mathbb{A}, a)\restriction N(e, a)$. Two structures $\mathbb{A}$ and $\mathbb{B}$ are *e-equivalent* if for every e-type τ, they have exactly the same number of elements with e-type τ.

The concept of e-equivalence gives an easy way to prove inexpressibility results for $\mathrm{FO}(\mathbf{Q}_1)$. The following theorem is used later in this paper.

2.5 THEOREM. ([NUR96]) *For every positive integer r there is a positive integer e such that whenever $\mathbb{A}$ and $\mathbb{B}$ are e-equivalent structures, then $\mathbb{A} \equiv \mathbb{B}$ ($FO_r(\mathbf{Q}_1)$).*

2.3 Extended vectorization of quantifiers

Let σ be a relational vocabulary, where a relation symbol R in σ has arity r. We denote by σ^k the signature obtained by replacing each $R \in \sigma$ by a new rk-ary relation symbol R_k. If $\mathcal{C}$ is a class of σ-structures, the *k-ary version* of $\mathcal{C}$ is the class $\mathcal{C}^k$ of σ^k-structures defined as follows. A σ^k-structure $\mathbb{A}$ belongs to $\mathcal{C}^k$ if and only if the structure $\overline{\mathbb{A}} = (A^k, (\overline{R}_k(\mathbb{A}))_{R\in\sigma})$ belongs to $\mathcal{C}$, where

$$\overline{R}_k(\mathbb{A}) = \{(\bar{a}_1, \ldots, \bar{a}_r) \in (A^k)^r \mid \bar{a}_1{}^\frown\cdots{}^\frown\bar{a}_r \in R_k(\mathbb{A})\}.$$

In the case $k = 1$, obviously $\mathcal{C}^k = \mathcal{C}$.

Let a structure $\mathbb{A}$ and a congruence relation S on $\mathbb{A}$ be given. The quotient of $\mathbb{A}$ with respect to S is denoted by $\mathbb{A}/S$. The *quotient class* of $\mathcal{C}$ is

$$q(\mathcal{C}) = \{(\mathbb{A}, S) \mid S \text{ is a congruence relation on } \mathbb{A},\ \mathbb{A}/S \in \mathcal{C}\}.$$

We call $\mathsf{Q}_{q(\mathcal{C})}$ the *quotient of* $\mathsf{Q}_{\mathcal{C}}$. The *relativization* of $\mathcal{C}$ is

$$r(\mathcal{C}) = \{(\mathbb{A}, V) \mid V \subseteq A \text{ and } \mathbb{A}\restriction V \in \mathcal{C}\}$$

and $\mathsf{Q}_{r(\mathcal{C})}$ is the relativization of $\mathsf{Q}_{\mathcal{C}}$.

Note that a $\sigma \cup \{S\}$-structure $\mathbb{A}$ is in the class $q(\mathcal{C})$ if and only if $S(\mathbb{A})$ is a congruence relation on $\mathbb{A}$ and $(\mathbb{A}/S(\mathbb{A}))|_\sigma$ is in $\mathcal{C}$, where $|_\sigma$ denotes the operation of taking the *reduct* of a structure to the vocabulary σ (i.e., forgetting the interpretations of symbols not included in σ). Similarly, a $\sigma \cup \{V\}$-structure $\mathbb{A}$ is in the class $r(\mathcal{C})$ if and only if $(\mathbb{A}{\upharpoonright}V)|_\sigma$ is in $\mathcal{C}$.

Assume for the rest of this section that L is a logic which extends FO and which has the substitution property, i.e., L allows substituting relation symbols by formulas (see [Ebb85] for a precise definition). Note that such an L is always closed under Boolean operations and first-order quantifications. We say that L is *closed under vectorization*, if for each class $\mathcal{C}$ definable in L, the class $\mathcal{C}^k$ is also definable in L, for each $k \geq 1$. Correspondingly, L is *congruence closed*, if for each class $\mathcal{C}$ definable in L also the quotient class $q(\mathcal{C})$ is definable in L ([MS86]), and L is *closed under relativization*, if for each class $\mathcal{C}$ definable in L also the relativized class $r(\mathcal{C})$ is definable in L. The *extended vectorization* $L_*^{(\omega)}$ of the logic L is the least logic extending L with the substitution property which is congruence closed and closed under vectorization and relativization.

Consider then a Lindström logic $\mathrm{FO}(\mathbf{Q})$. Let $\mathrm{FO}^{(\omega)}(\mathbf{Q})$ be the logic obtained by vectorizing all the quantifiers in $\mathbf{Q}$, i.e.,

$$\mathrm{FO}^{(\omega)}(\mathbf{Q}) = \mathrm{FO}(\{\mathsf{Q}_{\mathcal{C}^k} \mid \mathsf{Q}_{\mathcal{C}} \in \mathbf{Q},\ k \geq 1\}).$$

It is straightforward to prove by induction that $\mathrm{FO}^{(\omega)}(\mathbf{Q})$ is the least extension of $\mathrm{FO}(\mathbf{Q})$ with the substitution property that is closed under vectorization. In [MS86] it was observed that $\mathrm{FO}(\{\mathsf{Q}_{q(\mathcal{C})} \mid \mathsf{Q}_{\mathcal{C}} \in \mathbf{Q}\})$ is the least congruence closed logic which extends $\mathrm{FO}(\mathbf{Q})$ and has the substitution property. Furthermore, it was proved in [Kry88] that if $L = \mathrm{FO}(\mathbf{Q})$ or $L = \mathrm{FO}^{(\omega)}(\mathbf{Q})$, then

$$L_*^{(\omega)} \equiv \mathrm{FO}^{(\omega)}(\{\mathsf{Q}_{q(r(\mathcal{C}))} \mid \mathsf{Q}_{\mathcal{C}} \in \mathbf{Q}\}) \equiv \mathrm{FO}(\{\mathsf{Q}_{(q(r(\mathcal{C})))^k} \mid \mathsf{Q}_{\mathcal{C}} \in \mathbf{Q},\ k \geq 1\}).$$

We call the class $(q(r(\mathcal{C})))^k$ the *k-th extended vectorization* of $\mathcal{C}$ and denote it by $\mathcal{C}_*^k$; analogously we call $\mathsf{Q}_{\mathcal{C}_*^k}$ the *k-th extended vectorization* of $\mathsf{Q}_{\mathcal{C}}$.

¿From the point of view of descriptive complexity theory, it is natural to require that a logic is congruence closed and closed under vectorization and relativization. This is because of the following observation, which extends a corresponding result in [GH96] by allowing relativization.

2.6 Proposition. *Let K be one of the standard complexity classes* LOGSPACE, NLOGSPACE, PTIME, NP, PSPACE, *etc. Let L be a logic which extends FO and which has the substitution property. If L is contained in K (i.e., every L-definable class is K-computable), then $L_*^{(\omega)}$ is also contained in K.*

Proof. Note first that $L \equiv \mathrm{FO}(\mathbf{Q}_L)$, where $\mathbf{Q}_L$ is the set of all quantifiers $\mathsf{Q}_{\mathcal{C}}$ such that $\mathcal{C}$ is definable in L. Thus, by the observation above, $L_*^{(\omega)} \equiv \mathrm{FO}(\{\mathsf{Q}_{\mathcal{C}_*^k} \mid \mathsf{Q}_{\mathcal{C}} \in \mathbf{Q}_L,\ k \geq 1\})$. Furthermore, it is easy to see that if a class $\mathcal{C}$ is K-computable, then so are the classes $q(\mathcal{C})$, $r(\mathcal{C})$ and $\mathcal{C}^k$. Thus, if L is contained in K, then each of the quantifiers $\mathsf{Q}_{\mathcal{C}_*^k}$, $\mathsf{Q}_{\mathcal{C}} \in \mathbf{Q}_L$, is K-computable. It is straightforward to prove by induction that if φ is a sentence of $L_*^{(\omega)}$, then the class defined by φ is K-computable. □

3 From 1-bijective to k-bijective equivalence

In this section we give a general construction which starts from a pair of 1-bijectively equivalent structures and gives a pair of k-bijectively equivalent structures.

Let $\mathbb{A}$ and $\mathbb{B}$ be σ-structures and $c : A \to \{0, \ldots, s-1\}$ and $c' : B \to \{0, \ldots, s-1\}$ colorings of A and B with s different colors. Assume that $f : A \to B$ is a bijection. If for every $a \in A$ we have $c'(f(a)) = c(a)$, then f is *color-preserving*. The colorings c and c' give rise to partitions of the sets A and B into *color blocks* $P_j^c = \{a \in A \mid c(a) = j\}$ and $P_j^{c'} = \{b \in B \mid c'(b) = j\}$, $j < s$. We denote the *color block sizes* $|P_j^c|$ and $|P_j^{c'}|$ by t_j and t'_j, respectively. The *color-class size* of c (c') is $\max\{t_j \mid j < s\}$ ($\max\{t'_j \mid j < s\}$).

Let t be the color-class size of c, and let $\iota : A \to \{0, \ldots, t-1\}$ be a function which, for each color j, enumerates the elements of the color block P_j^c by numbers from 0 to $t_j - 1$. For each $0 < l < t$, let $F_l^\iota \subseteq A^2$ be the relation such that $(a, b) \in F_l^\iota$ if and only if $c(a) = c(b) = j$ and $\iota(b) \equiv \iota(a) + l \pmod{t_j}$. Thus the restriction of F_1^ι to each color block P_j^c is a directed cycle of length t_j. Note further that $(a, b) \in F_l^\iota$ if and only if $c(a) = c(b)$ and the F_1^ι-path starting from a and ending at b is of length l.

Given a σ-structure $\mathbb{A}$, a coloring $c : A \to \{0, \ldots, s-1\}$ and an indexing function $\iota : A \to \{0, \ldots, t-1\}$ as above, we denote the expanded structure $(\mathbb{A}, P_0^c, \ldots, P_{s-1}^c, F_1^\iota, \ldots, F_{t-1}^\iota)$ by $\mathbb{A}^{c,\iota}$.

3.1 Definition. Suppose $\mathbb{A}$ and $\mathbb{B}$ are σ-structures. Structures $\mathbb{A}$ and $\mathbb{B}$ are

- *cyclically* $\mathrm{FO}_r(\mathbf{Q}_1)$*-equivalent* if there are colorings $c : A \to \{0, \ldots, s-1\}$ and $c' : B \to \{0, \ldots, s-1\}$ and functions ι and ι' which index elements in each color block such that $\mathbb{A}^{c,\iota} \equiv \mathbb{B}^{c',\iota'}$ $(\mathrm{FO}_r(\mathbf{Q}_1))$;
- *cyclically* $\mathcal{L}^k_{\infty\omega}(\mathbf{Q}_1)$*-equivalent* if there are colorings $c : A \to \{0, \ldots, s-1\}$ and $c' : B \to \{0, \ldots, s-1\}$ and functions ι and ι' which index elements in each color block such that $\mathbb{A}^{c,\iota} \equiv \mathbb{B}^{c',\iota'}$ $(\mathcal{L}^k_{\infty\omega}(\mathbf{Q}_1))$.

This means that if $\mathbb{A}$ and $\mathbb{B}$ are cyclically $\mathrm{FO}_r(\mathbf{Q}_1)$-equivalent (cyclically $\mathcal{L}^k_{\infty\omega}(\mathbf{Q}_1)$-equivalent) with color class size t, each bijection chosen by the duplicator in $\mathrm{BEF}_1^r(\mathbb{A}^{c,\iota}, \mathbb{B}^{c',\iota'})$ $(\mathrm{BP}_1^k(\mathbb{A}^{c,\iota}, \mathbb{B}^{c',\iota'}))$ according to his winning strategy has to be color-preserving, and cyclicity means that any two different bijections chosen by the duplicator according to his winning strategy can differ in each color block only by some cyclic permutation. In other words, each bijection chosen by the duplicator has to be of the form $\bigcup_{j<s} \pi^j_{l(j)}$, where π^j_l denotes the cyclic function $P_j^c \to P_j^{c'}$ defined by the condition $\pi_l^j(a) = b \iff \iota'(b) \equiv \iota(a) + l \pmod{t_j}$.

Observe that in the case when color-class size is two, F_1^ι and $F_1^{\iota'}$ become trivial and can be ignored. Many natural graph properties can be proved to be undefinable in $\mathrm{FO}(\mathbf{Q}_1)$ by constructing suitable 1-bijectively equivalent graphs with color class size two. Such properties are considered in Section 4.

Note also that whenever $\mathbb{A}$ and $\mathbb{B}$ are cyclically $\mathrm{FO}_r(\mathbf{Q}_1)$-equivalent (cyclically $\mathcal{L}^k_{\infty\omega}(\mathbf{Q}_1)$-equivalent) and an element $a \in P_j^c$ has been selected by the spoiler in $\mathrm{BEF}_1^r(\mathbb{A}^{c,\iota}, \mathbb{B}^{c',\iota'})$ $(\mathrm{BP}_1^k(\mathbb{A}^{c,\iota}, \mathbb{B}^{c',\iota'}))$, the next bijection f played by the duplicator in the game has to preserve all the cycle relations F_l, whence the restriction of f to the color block P_j is uniquely determined. Thus, cyclic equivalence of $\mathbb{A}$ and $\mathbb{B}$ guarantees a winning strategy for the duplicator even in the modified game in which the spoiler is allowed to choose entire color blocks instead of singletons. We denote this modified game by MBEF_1^r (MBP_1^k).

We will now give the promised construction of vectorized structures which are used in the main result in this paper.

3.2 DEFINITION. Assume σ is a relational vocabulary and $\mathbb{A} = (A, (R(\mathbb{A}))_{R\in\sigma})$ is a finite σ-structure. Let c be a coloring on A and let ι be a function which enumerates elements in each color block of c. For each $k \geq 1$, we define the following $\sigma^k \cup \{V_k, S_k\}$-structure $\mathbb{A}_k^{c,\iota} = (A_k, (R_k(\mathbb{A}_k^{c,\iota}))_{R\in\sigma}, V_k(\mathbb{A}_k^{c,\iota}), S_k(\mathbb{A}_k^{c,\iota}))$:

$$A_k = A \times \{1, \ldots, k\},$$

$$V_k(\mathbb{A}_k^{c,\iota}) = \{((a_1, 1), \ldots, (a_k, k)) \in (A_k)^k \mid c(a_i) = c(a_j) \text{ for all } 1 \leq i, j \leq k\}.$$

Let $h_{\mathbb{A}}^{c,\iota} : V_k(\mathbb{A}_k^{c,\iota}) \to A$ be the function such that, for each $\overline{a} = ((a_1, 1), \ldots, (a_k, k)) \in V_k(\mathbb{A}_k^{c,\iota})$, $h_{\mathbb{A}}^{c,\iota}(\overline{a}) = a$, where a is the unique element of A such that $c(a) = c(a_1)$ and $\iota(a) \equiv \sum_{i=1}^k \iota(a_i) \pmod{t_{c(a)}}$. For each relation symbol $R \in \sigma$ with arity $\mathrm{ar}(R) = n$, we define

$$R_k(\mathbb{A}_k^{c,\iota}) = \{\overline{a}_1\frown\cdots\frown\overline{a}_n \in (A_k)^{kn} \mid \overline{a}_1, \ldots, \overline{a}_n \in V_k(\mathbb{A}_k) \text{ and } (h_{\mathbb{A}}^{c,\iota}(\overline{a}_1), \ldots, h_{\mathbb{A}}^{c,\iota}(\overline{a}_n)) \in R(\mathbb{A})\}.$$

Finally we define

$$\begin{aligned} S_k(\mathbb{A}_k^{c,\iota}) = \quad & \{\overline{a}\frown\overline{b} \in (A_k)^{k\cdot 2} \mid \overline{a}, \overline{b} \in V_k(\mathbb{A}_k^{c,\iota}) \text{ and } h_{\mathbb{A}}^{c,\iota}(\overline{a}) = h_{\mathbb{A}}^{c,\iota}(\overline{b})\} \cup \\ & \{\overline{a}\frown\overline{b} \in (A_k)^{k\cdot 2} \mid \overline{a}, \overline{b} \notin V_k(\mathbb{A}_k^{c,\iota})\}. \end{aligned}$$

¿From the definition of the relations $R_k(\mathbb{A}_k^{c,\iota})$ it follows that $\overline{S}_k(\mathbb{A}_k^{c,\iota})$ is a congruence relation on $\overline{\mathbb{A}}_k^{c,\iota}$, where $\overline{\mathbb{A}}_k^{c,\iota} = ((A_k)^k, (\overline{R}_k(\mathbb{A}_k^{c,\iota}))_{R\in\sigma}, \overline{V}_k(\mathbb{A}_k^{c,\iota}), \overline{S}_k(\mathbb{A}_k^{c,\iota}))$ is the $\sigma \cup \{V, S\}$-structure which is obtained from $\mathbb{A}_k^{c,\iota}$ by regarding k-tuples as elements (see Section 2.3).

Henceforth, we will call $\mathbb{A}_k^{c,\iota}$ the *k-th extended vectorization* of $\mathbb{A}$ *with respect to* c and ι. This name is justified by the next lemma.

3.3 LEMMA. *Let $\mathcal{C}$ be a class of σ-structures and $\mathbb{A}$, c and ι as above. Then $\mathbb{A} \in \mathcal{C}$ if and only if $\mathbb{A}_k^{c,\iota} \in \mathcal{C}_*^k$.*

Proof. By the definition of $\mathcal{C}_*^k$ we have

$$\begin{aligned} \mathbb{A}_k^{c,\iota} \in \mathcal{C}_*^k \quad & \Leftrightarrow \quad \overline{\mathbb{A}}_k^{c,\iota} \in q(r(\mathcal{C})) \\ & \Leftrightarrow \quad (\overline{\mathbb{A}}_k^{c,\iota} / \overline{S}_k(\mathbb{A}_k^{c,\iota}))|_{\sigma\cup\{V\}} \in r(\mathcal{C}) \\ & \Leftrightarrow \quad [(\overline{\mathbb{A}}_k^{c,\iota} / \overline{S}_k(\mathbb{A}_k^{c,\iota})) \upharpoonright \overline{V}_k(\mathbb{A}_k^{c,\iota})]|_\sigma \in \mathcal{C}. \end{aligned}$$

Thus the claim follows immediately from the fact that $h_{\mathbb{A}}^{c,\iota}$ induces the isomorphism

$$[(\overline{\mathbb{A}}_k^{c,\iota} / \overline{S}_k(\mathbb{A}_k^{c,\iota})) \upharpoonright \overline{V}_k(\mathbb{A}_k^{c,\iota})]|_\sigma \cong \mathbb{A}.$$

□

For the rest of this section we fix two σ-structures $\mathbb{A}$ and $\mathbb{B}$, colorings $c : A \to \{0, \dots, s-1\}$ and $c' : B \to \{0, \dots, s-1\}$, and enumerations of color blocks $\iota : A \to \{0, \dots, t-1\}$ and $\iota' : B \to \{0, \dots, t-1\}$. Let $\mathbb{A}_k$ and $\mathbb{B}_k$ be the corresponding extended vectorizations of $\mathbb{A}$ and $\mathbb{B}$ (to simplify notation, we drop the superscripts c, ι and c', ι'). A partial function $p : A_k \to B_k$ is *good* (for our purposes) if it satisfies the following conditions:

- for all $(a, i) \in \mathrm{dom}(p)$, $p((a, i)) = (b, i)$ for some $b \in B$ with $c(a) = c'(b)$;
- for all $(a, i), (a', i) \in \mathrm{dom}(p)$, if $p((a, i)) = (b, i)$ and $p((a', i)) = (b', i)$, then $(a, a') \in F_l^{\iota} \iff (b, b') \in F_l^{\iota'}$ for each $0 < l < t$.

Thus, p is good if it preserves second components, and the colors and the cycle relations on first components. Note that this holds if and only if for each $j < s$ and $1 \le i \le k$ there is an l such that $p((a, i)) = (\pi_l^j(a), i)$ for each $(a, i) \in \mathrm{dom}(p)$ with $a \in P_j^c$. (Recall that $\pi_l^j : P_j^c \to P_j^{c'}$ is the cyclic function such that $\pi_l^j(a) = b \iff \iota'(b) \equiv \iota(a) + l \pmod{t_j}$.)

As can be easily observed, any good partial function preserves the relations V_k and S_k. We denote by $p^k : V_k(\mathbb{A}_k) \to V_k(\mathbb{B}_k)$ the induced function defined as $p^k(\overline{a}) = (p((a_1, 1)), \dots, p((a_k, k)))$ for all $\overline{a} = ((a_1, 1), \dots, (a_k, k)) \in V_k(\mathbb{A}_k)$. In the sequel we work only with good partial functions.

Our aim is to prove that if $\mathbb{A}$ and $\mathbb{B}$ are cyclically $\mathrm{FO}_{kr}(\mathbf{Q}_1)$-equivalent, then $\mathbb{A}_k$ and $\mathbb{B}_k$ are $\mathrm{FO}_r(\mathbf{Q}_k)$-equivalent (Theorem 3.6). For this purpose we need the following two lemmas.

3.4 Lemma. *Assume that $g : A \to B$ is a color-preserving bijection that also preserves the cycle relations F_l, $0 < l < t$. Assume further that $p : A_k \to B_k$ is a good partial function such that $h_{\mathbb{B}} \circ p^k \subseteq g \circ h_{\mathbb{A}}$. Then there is a good bijection $f_{p,g} : A_k \to B_k$ such that $p \subseteq f_{p,g}$ and $h_{\mathbb{B}} \circ f_{p,g}^k = g \circ h_{\mathbb{A}}$.*

Proof. We start by choosing for each $j < s$ and $1 \le i \le k$ a number $l(j, i) < t_j$. The desired function $f_{p,g}$ is then defined in terms of the cyclic functions $\pi_{l(j,i)}^j$. The numbers $l(j, i)$ are chosen according to the following cases:

Case 1. Assume that $\mathrm{dom}(p) \cap P_j^c \times \{i\} \neq \emptyset$. Then, since p is good, there is $l < t_j$ such that $p((a, i)) = (\pi_l^j(a), i)$ for all $(a, i) \in \mathrm{dom}(p)$ with $a \in P_j^c$. We let $l(j, i)$ to be this l.

Case 2. Assume then that $\mathrm{dom}(p) \cap P_j^c \times \{i\} = \emptyset$. There are two subcases:

> **Subcase 2a.** If the same also holds for some $i' \le k$ such that $i < i'$ (i.e., $\mathrm{dom}(p) \cap P_j^c \times \{i'\} = \emptyset$), then we let $l(j, i) = 0$.
>
> **Subcase 2b.** Otherwise i is the largest number $\le k$ satisfying the assumption of Case 2. Since g preserves colors and the cycle relations, there is an $m < t_j$ such that $g \restriction P_j^c = \pi_m^j$. In this case we let

$$l(j, i) \equiv m - (l(j, 1) + \ldots + l(j, i-1) + l(j, i+1) + \ldots + l(j, k)) \pmod{t_j}.$$

The function $f_{p,g}$ is now defined as follows: for each $(a, i) \in A_k$, $f_{p,g}((a, i)) = (\pi_{l(j,i)}^j(a), i)$, where $j = c(a)$.

It is clear that $f_{p,g}$ is a good bijection. Furthermore, the way of choosing $l(j,i)$ in Case 1 guarantees that $p \subseteq f_{p,g}$. To prove that $f_{p,g}$ satisfies the last claim of the lemma, assume that $\overline{a} = ((a_1,1),\ldots,(a_k,k)) \in V_k(\mathbb{A}_k)$. Let $c(a_1) = j$. Our task is to show that $h_{\mathbb{B}}(f^k_{p,g}(\overline{a})) = g(h_{\mathbb{A}}(\overline{a}))$. Since

$$c'(h_{\mathbb{B}}(f^k_{p,g}(\overline{a}))) = c'(\pi^j_{l(j,1)}(a_1)) = c(a_1) = c(h_{\mathbb{A}}(\overline{a})) = c'(g(h_{\mathbb{A}}(\overline{a}))),$$

it suffices to show that $\iota'(h_{\mathbb{B}}(f^k_{p,g}(\overline{a}))) = \iota'(g(h_{\mathbb{A}}(\overline{a})))$. There are two cases:

Case A. If $\mathrm{dom}(p) \cap P^c_j \times \{i\} \neq \emptyset$ for all $1 \leq i \leq k$, then there is a tuple $\overline{b} = ((b_1,1),\ldots,(b_k,k)) \in \mathrm{dom}(p^k)$ with $c(b_1) = c(a_1) = j$, and since $h_{\mathbb{B}} \circ p^k \subseteq g \circ h_{\mathbb{A}}$, we have $\iota'(h_{\mathbb{B}}(p^k(\overline{b}))) = \iota'(g(h_{\mathbb{A}}(\overline{b})))$. Then by the choice of $l(j,i)$ in Case 1

$$\iota'(h_{\mathbb{B}}(f^k_{p,g}(\overline{a}))) - \iota(h_{\mathbb{A}}(\overline{a})) \equiv \sum_{i=1}^{k} l(j,i) \equiv \iota'(h_{\mathbb{B}}(p^k(\overline{b}))) - \iota(h_{\mathbb{A}}(\overline{b})) \pmod{t_j},$$

whence

$$\iota'(h_{\mathbb{B}}(f^k_{p,g}(\overline{a}))) - \iota'(h_{\mathbb{B}}(p^k(\overline{b}))) \equiv \iota(h_{\mathbb{A}}(\overline{a})) - \iota(h_{\mathbb{A}}(\overline{b})) \pmod{t_j}.$$

By assumption, g preserves the relations F_l. This means that

$$\iota'(g(h_{\mathbb{A}}(\overline{a}))) - \iota'(g(h_{\mathbb{A}}(\overline{b}))) \equiv \iota(h_{\mathbb{A}}(\overline{a})) - \iota(h_{\mathbb{A}}(\overline{b})) \pmod{t_j}.$$

Hence we finally get

$$\iota'(g(h_{\mathbb{A}}(\overline{a}))) \equiv \iota'(h_{\mathbb{B}}(f^k_{p,g}(\overline{a}))) - \iota'(h_{\mathbb{B}}(p^k(\overline{b}))) + \iota'(g(h_{\mathbb{A}}(\overline{b}))) \equiv \iota'(h_{\mathbb{B}}(f^k_{p,g}(\overline{a}))) \pmod{t_j}.$$

Case B. Assume then that $\mathrm{dom}(p) \cap P^c_j \times \{i\} = \emptyset$ for some $1 \leq i \leq k$. Then we have

$$\iota'(h_{\mathbb{B}}(f^k_{p,g}(\overline{a}))) \equiv \sum_{i=1}^{k} \iota'(\pi^j_{l(j,i)}(a_i)) \equiv \sum_{i=1}^{k} (\iota(a_i) + l(j,i)) \equiv \iota(h_{\mathbb{A}}(\overline{a})) + \sum_{i=1}^{k} l(j,i) \pmod{t_j}.$$

By the choice made in Case 2b, $g{\restriction}P^c_j = \pi^j_m$ for $m \equiv \sum_{i=1}^{k} l(j,i) \pmod{t_j}$. Hence we get

$$\iota'(h_{\mathbb{B}}(f^k_{p,g}(\overline{a}))) \equiv \iota(h_{\mathbb{A}}(\overline{a})) + m \equiv \iota'(g(h_{\mathbb{A}}(\overline{a}))) \pmod{t_j},$$

which completes the proof. □

3.5 Lemma. *Let $p : A_k \to B_k$ be a good partial function. Assume that $q : \mathbb{A} \to \mathbb{B}$ is a partial isomorphism and $h_{\mathbb{B}} \circ p^k \subseteq q \circ h_{\mathbb{A}}$. Then also p is a partial isomorphism $\mathbb{A}_k \to \mathbb{B}_k$.*

Proof. As a good partial function p preserves the relations V_k and S_k. Hence we only need to check the relations $R_k \in \sigma^k$. Let $R \in \sigma$ be a relation with $\mathrm{ar}(R) = n$ and $\overline{a}_1,\ldots,\overline{a}_n \in V_k(\mathbb{A}_k) \cap \mathrm{dom}(p^k)$. By the definitions of the relations R_k and the assumptions that q is a partial isomorphism and $h_{\mathbb{B}} \circ p^k \subseteq q \circ h_{\mathbb{A}}$, we have the following chain of equivalences:

$$\begin{aligned}
\overline{a}_1{}^\frown\cdots{}^\frown\overline{a}_n \in R_k(\mathbb{A}_k) &\Leftrightarrow (\overline{a}_1, \ldots, \overline{a}_n) \in \overline{R}_k(\mathbb{A}_k) \\
&\Leftrightarrow (h_{\mathbb{A}}(\overline{a}_1), \ldots, h_{\mathbb{A}}(\overline{a}_n)) \in R(\mathbb{A}) \\
&\Leftrightarrow (q(h_{\mathbb{A}}(\overline{a}_1)), \ldots, q(h_{\mathbb{A}}(\overline{a}_n))) \in R(\mathbb{B}) \\
&\Leftrightarrow (h_{\mathbb{B}}(p^k(\overline{a}_1)), \ldots, h_{\mathbb{B}}(p^k(\overline{a}_n))) \in R(\mathbb{B}) \\
&\Leftrightarrow (p^k(\overline{a}_1), \ldots, p^k(\overline{a}_n)) \in \overline{R}_k(\mathbb{B}_k) \\
&\Leftrightarrow p\overline{a}_1{}^\frown\cdots{}^\frown\overline{a}_n \in R_k(\mathbb{B}_k),
\end{aligned}$$

where $p\overline{a}_1{}^\frown\cdots{}^\frown\overline{a}_n$ denotes the image of the tuple $\overline{a}_1{}^\frown\cdots{}^\frown\overline{a}_n$ under the function p. Thus p is a partial isomorphism $\mathbb{A}_k \to \mathbb{B}_k$. □

3.6 THEOREM. *Let k and r be positive integers.*

(1) *If $\mathbb{A}^{c,\iota} \equiv \mathbb{B}^{c',\iota'}$ ($FO_{kr}(\mathbf{Q}_1)$), then $\mathbb{A}_k \equiv \mathbb{B}_k$ ($FO_r(\mathbf{Q}_k)$).*

(2) *If $\mathbb{A}^{c,\iota} \equiv \mathbb{B}^{c',\iota'}$ ($\mathcal{L}^r_{\infty\omega}(\mathbf{Q}_1)$), then $\mathbb{A}_k \equiv \mathbb{B}_k$ ($\mathcal{L}^r_{\infty\omega}(\mathbf{Q}_k)$).*

Proof. (1) Assume that $\mathbb{A}^{c,\iota} \equiv \mathbb{B}^{c',\iota'}$ ($\mathrm{FO}_{kr}(\mathbf{Q}_1)$). As we noted after Definition 3.1, the duplicator has a winning strategy μ in the modified game $\mathrm{MBEF}_1^{kr}(\mathbb{A}^{c,\iota}, \mathbb{B}^{c',\iota'})$ in which the spoiler chooses color blocks instead of single elements in his moves. We prove the claim by first describing a strategy ν for the duplicator in the game $\mathrm{BEF}_k^r(\mathbb{A}_k, \mathbb{B}_k)$, and then showing that ν is a winning strategy.

The desired strategy ν makes use of the bijections given by μ, and it is defined as follows:

(i) The first move f_1 of the duplicator is the good bijection $f_{p,g} : A_k \to B_k$ given by Lemma 3.4, where $p = \emptyset$ and $g : A \to B$ is the first move of the duplicator in the game $\mathrm{MBEF}_1^{kr}(\mathbb{A}^{c,\iota}, \mathbb{B}^{c'\iota'})$ according to μ.

(ii) Assume that the spoiler has chosen sets $C_1, \ldots, C_n \subseteq A_k$ in the n first rounds, where $n < r$, and the moves $f_1, \ldots, f_n$ of the duplicator according to ν have already been defined. In order to determine the next move f_{n+1} of the duplicator according to ν, we first define a sequence of auxiliary sets.

For each $1 \leq u \leq n$, let $O_1^u, \ldots, O_{v_u}^u$ be a list of those color blocks in A that correspond to the elements in C_u, i.e., P_j^c is in the list if and only if there is a pair $(a, i) \in C_u$ such that $a \in P_j^c$. Note that $v_u \leq k$, since $|C_u| \leq k$. Thus, the sequence $O_1^1, \ldots, O_{v_1}^1, \ldots, O_1^n, \ldots, O_{v_n}^n$ is of length at most $kn < kr$, whence it is a possible sequence of moves for the spoiler in the game $\mathrm{MBEF}_1^{kr}(\mathbb{A}^{c,\iota}, \mathbb{B}^{c',\iota'})$. Let $g : A \to B$ be the bijection given by μ in the position in which the spoiler has played this sequence of moves, and let $p = \bigcup_{1 \leq u \leq n} f_u \restriction C_u$. In principle, there are now two possibilities:

(a) If p is a good partial function and $h_{\mathbb{B}} \circ p^k \subseteq g \circ h_{\mathbb{A}}$, then we let f_{n+1} be the good bijection $f_{p,g} : A_k \to B_k$ given by Lemma 3.4.

(b) Otherwise we let $f_{n+1} = f_n$.

(We will soon see that, in fact, alternative (b) is never realized.)

We prove now that ν is indeed a winning strategy. Thus, assume that $C_1, \ldots, C_r$ is a legal sequence of moves for the spoiler in the game $\mathrm{BEF}_k^r(\mathbb{A}_k, \mathbb{B}_k)$ and $f_1, \ldots, f_r$ is the corresponding sequence of moves of the duplicator when he plays according to ν. Furthermore, let $O_1^1, \ldots, O_{v_1}^1, \ldots, O_1^r, \ldots, O_{v_r}^r$ be the associated list of color blocks of A (see (ii) above). For each $1 \leq n \leq r$ and $1 \leq v \leq v_n$, let g_v^n be the bijection given by the strategy μ in the position in which the spoiler has moved the color blocks $O_1^1, \ldots, O_{v_1}^1, \ldots, O_1^n, \ldots, O_{v-1}^n$, and let q_v^n be the restriction of g_v^n to the union of the sets $O_1^1, \ldots, O_{v_1}^1, \ldots, O_1^n, \ldots, O_v^n$. Note that since μ is a winning strategy in $\mathrm{MBEF}_1^{kr}(\mathbb{A}^{c,\iota}, \mathbb{B}^{c',\iota'})$, each q_v^n is a partial isomorphism $\mathbb{A}^{c,\iota} \to \mathbb{B}^{c',\iota'}$. We need to show that $p_n = \bigcup_{1 \leq u \leq n} f_u \restriction C_u$ is a partial isomorphism $\mathbb{A}_k \to \mathbb{B}_k$ for each $1 \leq n \leq r$. Note first that each p_n is a good partial function. This is because the definition of ν guarantees that f_n is always a good bijection and $p_n \subseteq f_n$. Thus, by Lemma 3.5, it suffices to show that $h_{\mathbb{B}} \circ p_n^k \subseteq q_1^n \circ h_{\mathbb{A}}$ for all $1 \leq n \leq r$. We do this by induction on n.

(i) By the definition of ν, f_1 is a good bijection such that $h_{\mathbb{B}} \circ f_1^k = g_1^1 \circ h_{\mathbb{A}}$. Since $|C_1| \leq k$, the intersection of $(C_1)^k$ and $V_k(\mathbb{A}_k)$ is either empty, or contains exactly one tuple. Let $\overline{a} \in (C_1)^k \cap V_k(\mathbb{A}_k)$ be this unique tuple, if it exists; otherwise we let $\overline{a} \in V_k(\mathbb{A}_k)$ be an arbitrary tuple such that $h_{\mathbb{A}}(\overline{a}) \in O_1^1$. Note that $h_{\mathbb{A}}(\overline{a}) \in O_1^1$ also in the former case. Then we have

$$h_{\mathbb{B}} \circ p_1^k \subseteq (h_{\mathbb{B}} \circ f_1^k) \restriction \{\overline{a}\} = (g_1^1 \circ h_{\mathbb{A}}) \restriction \{\overline{a}\} \subseteq q_1^1 \circ h_{\mathbb{A}}.$$

(ii) Let $n < r - 1$, and assume as induction hypothesis that $h_{\mathbb{B}} \circ p_n^k \subseteq q_1^n \circ h_{\mathbb{A}}$. Since $q_1^n \subseteq q_{v_n}^n \subseteq g_1^{n+1}$, we have $h_{\mathbb{B}} \circ p_n^k \subseteq g_1^{n+1} \circ h_{\mathbb{A}}$, and hence, by the definition of ν, f_{n+1} is the good bijection $f_{p_n, g_1^{n+1}}$ given by Lemma 3.4. In particular, $h_{\mathbb{B}} \circ f_{n+1}^k = g_1^{n+1} \circ h_{\mathbb{A}}$. Let M be the set of all elements of A_k that correspond to the color blocks $O_1^1, \ldots, O_{v_1}^1, \ldots, O_1^n, \ldots, O_{v_n}^n$:

$$M = (O_1^1 \cup \ldots \cup O_{v_1}^1 \cup \ldots \cup O_1^n \cup \ldots \cup O_{v_n}^n) \times \{1, \ldots, k\}.$$

There is at most one tuple $\overline{a}$ in the intersection $(\bigcup_{1 \leq u \leq n+1} C_u)^k \cap V_k(\mathbb{A}_k)$ which is not already contained in M^k. Let $\overline{a}$ be this unique tuple if it exists; otherwise we let $\overline{a} \in V_k(\mathbb{A}_k)$ be an arbitrary tuple such that $h_{\mathbb{A}}(\overline{a}) \in O_1^1$. Note that in any case $h_{\mathbb{A}}(\overline{a}) \in O_1^1 \cup \ldots \cup O_{v_n}^n \cup O_1^{n+1}$. Thus, we have

$$h_{\mathbb{B}} \circ p_{n+1}^k \subseteq (h_{\mathbb{B}} \circ f_{n+1}^k) \restriction (M^k \cup \{\overline{a}\}) = (g_1^{n+1} \circ h_{\mathbb{A}}) \restriction (M^k \cup \{\overline{a}\}) \subseteq q_1^{n+1} \circ h_{\mathbb{A}}.$$

(2) The analogous claim for the finite variable logics $\mathcal{L}_{\infty\omega}^r(\mathbf{Q}_1)$ and $\mathcal{L}_{\infty\omega}^r(\mathbf{Q}_k)$ is proved with a similar argument: The duplicator uses the bijections given by a winning strategy μ in the game $\mathrm{MBP}_1^r(\mathbb{A}^{c,\iota}, \mathbb{B}^{c'\,\iota'})$ in order to determine his moves f_n in the game $\mathrm{BP}_k^r(\mathbb{A}_k, \mathbb{B}_k)$.

The main difference is that this time the duplicator maintains the list $O_1^n, \ldots, O_{v_n}^n$ of those color blocks which after the n-th move of the spoiler correspond to elements in $\mathrm{dom}(p_n)$. Thus, if the spoiler chooses sets $C_{n+1} \subseteq A_k$ and $D_n \subseteq \mathrm{dom}(p_n)$ in his next move, the list is updated by first removing all color blocks P_j^c with $D_n \cap (P_j^c \times \{1, \ldots, k\}) = \emptyset$, and then adding all color blocks $P_{j'}^c$ with $C_{n+1} \cap (P_{j'}^c \times \{1, \ldots, k\}) \neq \emptyset$. This can be seen as a sequence (of length $\leq k$) of legal moves for the spoiler in $\mathrm{MBP}_1^r(\mathbb{A}^{c,\iota}, \mathbb{B}^{c'\iota'})$. Let $g : A \to B$ be the bijection given by μ as an answer to the last move of this sequence. As in the proof of claim (1), f_{n+1} is then defined to be the good bijection $f_{p_n,g}$.

The rest of the proof follows that of claim (1): Observing that there is at most one tuple $\overline{a} \in (C_{n+1} \cup D_n)^k \cap V_k(\mathbb{A}_k)$ that is not already contained in $(\bigcup_{1 \leq u \leq v_n} O_u^n \times \{1, \ldots, k\})^k$, it is easy to show that $h_{\mathbb{B}} \circ p_{n+1}^k \subseteq q \circ h_{\mathbb{A}}$, where q is a partial isomorphism $\mathbb{A} \to \mathbb{B}$ obtained as a suitable restriction of g. Hence, by Lemma 3.5, p_n is a partial isomorphism $\mathbb{A}_k \to \mathbb{B}_k$. □

3.7 COROLLARY. *Let $\mathcal{C}$ be a class of finite structures and k a positive integer. If for every q there are cyclically $FO_q(\mathbf{Q}_1)$-equivalent structures $\mathbb{A} \in \mathcal{C}$ and $\mathbb{B} \notin \mathcal{C}$, then the k-th extended vectorization of $\mathsf{Q}_{\mathcal{C}}$ is not definable in $FO(\mathbf{Q}_k)$.*

Proof. Assume on the contrary that $\varphi \in \mathrm{FO}(\mathbf{Q}_k)$ defines the class $\mathcal{C}_*^k$. Let $\mathrm{qr}(\varphi) = r$. By assumption there exist $\mathbb{A} \in \mathcal{C}$ and $\mathbb{B} \notin \mathcal{C}$ together with colorings c and c' and enumerations ι and ι' such that $\mathbb{A}^{c,\iota} \equiv \mathbb{B}^{c',\iota'}$ $(\mathrm{FO}_{kr}(\mathbf{Q}_1))$. Thus, by Theorem 3.6, $\mathbb{A}_k \equiv \mathbb{B}_k$ $(\mathrm{FO}_r(\mathbf{Q}_k))$. In particular, $\mathbb{A}_k \models \varphi$ if and only if $\mathbb{B}_k \models \varphi$. On the other hand, by Lemma 3.3 we have $\mathbb{A}_k \in \mathcal{C}_*^k$ and $\mathbb{B}_k \notin \mathcal{C}_*^k$, which is a contradiction. □

A logic L is called *finitely generated* if there is a finite set $\mathbf{Q}$ of generalized quantifiers such that $\mathrm{FO}(\mathbf{Q})$ and L have the same expressive power. It is obvious that if for each k the k-th extended vectorized of $\mathsf{Q}_{\mathcal{C}}$ is not definable in $\mathrm{FO}(\mathbf{Q}_k)$ then $\mathrm{FO}_*^{(\omega)}(\mathsf{Q}_{\mathcal{C}})$ is not finitely generated.

3.8 COROLLARY. *Let $\mathcal{C}$ be a class of finite structures. If for every q there are cyclically $FO_q(\mathbf{Q}_1)$-equivalent structures $\mathbb{A} \in \mathcal{C}$ and $\mathbb{B} \notin \mathcal{C}$, then the extended vectorization of the logic $FO(\mathsf{Q}_{\mathcal{C}})$ is not finitely generated.*

4 Applications to graph quantifiers

In this section we construct cyclically $\mathrm{FO}_r(\mathbf{Q}_1)$-equivalent graphs. By applying Corollary 3.7 we get inexpressibility results for extended vectorizations of certain classes of graphs. We also consider some natural graph properties for which this construction can be applied.

4.1 The general construction

Let $\mathbb{G} = (G, E(\mathbb{G}))$ be an undirected graph. A *stretching* $\mathbb{G}_e = (G_e, E(\mathbb{G}_e))$ *of order* e of $\mathbb{G}$ is the undirected graph obtained when each edge (a, b) of $\mathbb{G}$ is replaced by a path $p_{a,b}$ of length e; thus, in addition to the end points a and b, $p_{a,b}$ contains $e - 1$ new vertices. We consider G as a subset of G_e and call the vertices of $\mathbb{G}$ the original vertices of $\mathbb{G}_e$. For each original vertex a we denote by A_a the set of vertices on paths adjacent to a, i.e., $A_a = \{u \in G_e \mid \exists b \in G, u \in p_{a,b}\}$.

Our aim is to prove that if $\mathbb{G}$ and $\mathbb{H}$ are 2-equivalent, then for every r there is an e such that $\mathbb{G}_e$ and $\mathbb{H}_e$ are cyclically $\mathrm{FO}_r(\mathbf{Q}_1)$-equivalent. (Recall the concept of e-equivalence from Section 2.2.) We start with the following simple observation.

Suppose that $\mathbb{G}$ and $\mathbb{H}$ are undirected 2-equivalent graphs. Then for each k in both graphs there are equally many vertices with degree k, and it is easy to see that the duplicator has a winning strategy in the game $\mathrm{BEF}_1^2(\mathbb{G}, \mathbb{H})$. Indeed, as his first move the duplicator can give any bijection g that preserves the degrees of vertices, i.e., for every $a \in G$, $\deg(a) = \deg(g(a))$. For any move $a \in G$ of the spoiler, the duplicator can then choose a bijection $g_a : G \to H$ such that

$(*)$ $g_a(a) = g(a)$ and for every $b \in G$, $(g(a), g_a(b)) \in E(\mathbb{H})$ if and only if $(a, b) \in E(\mathbb{G})$.

A class $\mathcal{C}$ of graphs is *closed under stretching* if $\mathbb{G}_e \in \mathcal{C}$ for every $\mathbb{G} \in \mathcal{C}$ and positive integer e. For instance, the class of connected graphs and the class of rigid graphs are closed under stretching.

The next result gives an easy way to prove inexpressibility results in $\mathrm{FO}(\mathbf{Q}_1)$ for classes closed under stretching.

4.1 PROPOSITION. *Let $\mathbb{G}$ and $\mathbb{H}$ be finite graphs and e a positive integer. If $\mathbb{G}$ and $\mathbb{H}$ are 2-equivalent, then $\mathbb{G}_{2e-1}$ and $\mathbb{H}_{2e-1}$ are e-equivalent. Especially, for every positive integer r there is a positive integer e such that $\mathbb{G}_{2e-1} \equiv \mathbb{H}_{2e-1}$ $(FO_r(\mathbf{Q}_1))$.*

Proof. Assume that $\mathbb{G}$ and $\mathbb{H}$ are 2-equivalent. Let g and g_a, $a \in G$, be bijections $G \to H$ as in the observation above. We show that g can be extended to a bijection $f : G_{2e-1} \to H_{2e-1}$ which preserves e-types.

For each $a \in G$, let f_a be the bijection $A_a \to A_{g(a)}$ which maps each $u \in p_{a,b}$ to the unique $v \in p_{g(a),g_a(b)}$ such that $d_E(a, u) = d_E(g(a), v)$, where $d_E(x, y)$ is the distance between x and y with respect to the edge relation E. The bijection f is defined in terms of these functions:

$$f = \bigcup_{a \in G} (f_a \restriction N(e, a)).$$

Note that for each $u \in G_{2e-1}$ $(u \in H_{2e-1})$ there is exactly one original vertex a such that $d(a, u) < e$; we denote this a by $a(u)$. Thus, $f(u) = f_{a(u)}(u)$ for all $u \in G_{2e-1}$.

To see that f preserves e-types, observe that for each $u \in G_{2e-1}$, the isomorphism type of $(\mathbb{G}_{2e-1}, u) \restriction N(e, u)$ is completely determined by the distance between u and $a(u)$ and the degree of $a(u)$. Clearly these two parameters are equal for u and $f(u)$.

The second claim follows immediately from Theorem 2.5. □

In the following theorem we strengthen Proposition 4.1 by showing that, for large enough d, $\mathbb{G}_d$ and $\mathbb{H}_d$ are even cyclically $\mathrm{FO}_r(\mathbf{Q}_1)$-equivalent.

4.2 THEOREM. *Let $\mathbb{G}$ and $\mathbb{H}$ be finite 2-equivalent graphs. Then for every positive integer r there is a positive integer d such that $\mathbb{G}_d$ and $\mathbb{H}_d$ are cyclically $FO_r(\mathbf{Q}_1)$-equivalent.*

Proof. Suppose r is given and let e be such that $\mathbb{A} \equiv \mathbb{B}$ $(\mathrm{FO}_r(\mathbf{Q}_1))$ holds, whenever $\mathbb{A}$ and $\mathbb{B}$ are e-equivalent structures; such an e exists by Theorem 2.5. Let $d = 4e - 1$. It suffices to show that there are colorings c and c' and enumerations ι and ι' such that the expanded structures $\mathbb{G}_d^{c,\iota}$ and $\mathbb{H}_d^{c',\iota'}$ are e-equivalent.[1]

Let again g and g_a, $a \in G$, be bijections $G \to H$ satisfying condition $(*)$. Furthermore, let $f : G_d \to H_d$ be the extension of g as defined in the proof of Proposition 4.1; thus, $f(u) = f_{a(u)}(u)$, where $a(u)$ is the unique original vertex such that $d_E(a(u), u) < 2e$. Note

[1] We could prove this even for $d = 2e - 1$ (as in Proposition 4.1), but the proof would be more complicated.

that for each vertex $u \in G_d \setminus G$ ($u \in H_d \setminus H$) there is exactly one original vertex $b \neq a(u)$ such that $d_E(b,u) < 4e-1$ (b is the unique original vertex such that $u \in p_{a(u),b}$); we denote this b by $b(u)$. Replacing $a(u)$ by $b(u)$ in the definition of f, we obtain another closely related bijection $\hat{f}$ from G_d to H_d:

$$\hat{f}(u) = \begin{cases} g(u), & \text{if } u \in G \\ f_{b(u)}(u), & \text{if } u \in G_d \setminus G. \end{cases}$$

The desired colorings c and c' and enumerations ι and ι' are now defined in terms of the bijections f and $\hat{f}$ as follows. Let F be the binary relation on G_d defined by the condition

$$(u,v) \in F \iff v = \hat{f}^{-1}(f(u)).$$

Similarly, let F' be the binary relation on H_d such that

$$(u,v) \in F' \iff v = f(\hat{f}^{-1}(u)).$$

Since F is the graph of the permutation $\hat{f}^{-1} \circ f : G_d \to G_d$, it is a disjoint union of directed cycles. Note further that F' is the image of F under the bijection f, i.e., for all $u,v \in G_d$, $(f(u),f(v)) \in F' \iff (u,v) \in F$. Hence, F and F' are isomorphic.

We choose now a coloring c of G_d such that for all $u,v \in G_d$, $c(u) = c(v)$ if and only if u and v belong to the same cycle of F. Furthermore, we let c' be the coloring of H_d such that $c'(f(u)) = c(u)$ for all $u \in G_d$. To define the enumeration ι, we fix a vertex u_j from each color block P^c_j of c, and let $\iota(u_j) = 0$. For other vertices $u \in P^c_j$, we let $\iota(u)$ to be the length of the F-path from u_j to u. Finally, we let ι' be the function such that for all $u \in G_d$, $\iota'(f(u)) = \iota(u)$. Observe that the corresponding cycle relations F^ι_1 and $F^{\iota'}_1$ are then equal to F and F', respectively.

To complete the proof, we show that the bijection f preserves e-types in the expanded structures $\mathbb{G}^{c,\iota}_d$ and $\mathbb{H}^{c',\iota'}_d$, i.e., $(\mathbb{G}^{c,\iota}_d, u)\restriction N(e,u) \cong (\mathbb{H}^{c',\iota'}_d, f(u))\restriction N(e,f(u))$ for all $u \in G_d$. We divide this task into two cases according to the distance (with respect to E) between u and the closest original vertex $a(u)$.

Case A. Assume that $d_E(u,a(u)) < e$. We show that for any original vertex a, the restriction of f to $N(2e-1,a)$ is an isomorphism $(\mathbb{G}^{c,\iota}_d, a)\restriction N(2e-1,a) \to (\mathbb{H}^{c',\iota'}_d, f(a))\restriction N(2e-1, f(a))$. Since $N(e,u) \subseteq N(2e-1,a(u))$, it follows then that $f\restriction N(e,u)$ is an isomorphism $(\mathbb{G}^{c,\iota}_d, u)\restriction N(e,u) \to (\mathbb{H}^{c',\iota'}_d, f(u))\restriction N(e,f(u))$.

Note first that f (as well as $\hat{f}$) is an isomorphism with respect to the coloring relations P_j and the cycle relations F_l; this is because we defined c' and ι' to be the images of c and ι under f. Moreover, f preserves the edge relation E everywhere else, except at the middle of paths $p_{b,c}$ (and $p_{g(b),g(c)}$) such that $g_b(c) \neq g(c)$. But the two middle elements on a path $p_{b,c}$ are not contained in $N(2e-1,a)$ for any $a \in G$ (and the same holds for $p_{g(b),g(c)}$ and $N(2e-1,f(a))$), whence $f\restriction N(2e-1,a)$ is an isomorphism also with respect to E.

Case B. If $d_E(u,a(u)) \geq e$, then $N(e,u)$ does not contain any original vertices. Since f preserves the distance from the closest original vertex, the same holds for $N(e,f(u))$. It is then easy to see that, up to a renaming of the coloring relations

P_l, both $(\mathbb{G}_d^{c,\iota}, u)\restriction N(e,u)$ and $(\mathbb{H}_d^{c',\iota'}, f(u))\restriction N(e,f(u))$ are isomorphic to the structure $(A, E, P_0, \ldots, P_{s-1}, F_1, \ldots, F_{t-1}, c)$, where $A = \{0,\ldots,t-1\} \times \{1,\ldots,2e-3\} \cup \{(0,0),(0,2e-2)\}$, $E = \{((i,j),(i',j')) \mid i = i', |j-j'| = 1\}$, $P_l = \{(i,j) \mid j = l\}$ for $l < 2e-1$ and $P_l = \emptyset$ for $2e-1 \le l < s$, $F_l = \{((i,j),(i',j')) \mid j = j', i'-i \equiv l \pmod t\}$, and $c = (0, e-1)$. □

The following corollary is an immediate consequence of Corollary 3.7, Proposition 4.1 and Theorem 4.2.

4.3 COROLLARY. *Let $\mathcal{C}$ be a class of finite graphs such that $\mathcal{C}$ and its complement are closed under stretching. If there are 2-equivalent graphs $\mathbb{G} \in \mathcal{C}$ and $\mathbb{H} \notin \mathcal{C}$ then $\mathcal{C}$ is not definable in FO($\mathbf{Q}_1$). Moreover the k-th extended vectorization of $\mathsf{Q}_\mathcal{C}$ is not definable in FO($\mathbf{Q}_k$) for any $k \in \omega$. In particular, the extended vectorization of FO($\mathsf{Q}_\mathcal{C}$) is not finitely generated.*

4.4 REMARK. In the results above we assumed that the graphs $\mathbb{G}$ and $\mathbb{H}$ are 2-equivalent. However, the proofs of the results show that a slightly weaker assumption, namely that the duplicator has a winning strategy in $\mathrm{BEF}_1^2(\mathbb{G},\mathbb{H})$, suffices.

4.2 Examples of graph quantifiers

In [Nur96] it was proved that the class of connected finite graphs CONN, and also the class of rigid graphs RIG, are not definable in FO($\mathbf{Q}_1$). The proofs of these results are based on counterexamples which are of color class size two. More precisely, if $\mathcal{C}$ = CONN or $\mathcal{C}$ = RIG then for each r there are structures $\mathbb{A} \in \mathcal{C}$ and $\mathbb{B} \notin \mathcal{C}$ such that $\mathbb{A}$ and $\mathbb{B}$ are cyclically $\mathrm{FO}_r(\mathbf{Q}_1)$-equivalent. As a consequence to Corollary 3.7 we get the following result.

4.5 THEOREM. *For every positive integer k, the k-th extended vectorizations of $\mathsf{Q}_{\mathrm{CONN}}$ and $\mathsf{Q}_{\mathrm{RIG}}$ are not definable in FO($\mathbf{Q}_k$). Consequently, the extended vectorizations of FO($\mathsf{Q}_{\mathrm{CONN}}$) and FO($\mathsf{Q}_{\mathrm{RIG}}$) are not finitely generated.*

Note also that for this theorem we could use Corollary 4.3 and the obvious fact that the graph consisting of two cycles of length four and the graph consisting of one cycle of length eight, are 2-equivalent and the class of connected graphs and the class of non-connected graphs are closed under stretching. It is also well known that CONN is definable in FO(Q_{TC}) and thus it follows from the construction that $\mathsf{Q}_{\mathrm{TC}^k}$ is not definable in FO($\mathbf{Q}_k$) for any $k \in \omega$. On the other hand, Grohe and Hella [GH96] proved a hierarchy theorem for transitive closure logics which implies that $\mathsf{Q}_{\mathrm{TC}^k}$ is not definable even in FO($\mathbf{Q}_{2k-1}$). Hence our general construction gives as a direct corollary a slightly weaker version of this result of [GH96].

As a second type of examples, consider classes of graphs defined by forbidden minors. A graph $\mathbb{H}$ is a minor of a graph $\mathbb{G}$, if $\mathbb{H}$ can be obtained from a (not necessarily induced) subgraph of $\mathbb{G}$ by contracting edges. (Contracting an edge $(a_1,a_2) \in E(\mathbb{G})$ means removing the vertices a_1, a_2, adding a new vertex b, and replacing all edges (u,a_i), $i = 1,2$, by (u,b); for more details, see e.g. [Tut84].) A class $\mathcal{C}$ is defined by forbidden minors $\mathbb{H}_1,\ldots,\mathbb{H}_n$, $\mathcal{C} = \mathrm{FORB}(\mathbb{H}_1,\ldots\mathbb{H}_n)$, if for every graph $\mathbb{G}$, $\mathbb{G} \in \mathcal{C}$ if and only if no $\mathbb{H}_i$ is a minor of $\mathbb{G}$, for $1 \le i \le n$.

In [Nur96] the following notion was used to classify certain classes of graphs defined by forbidden minors.

4.6 DEFINITION. A finite graph $\mathbb{H}$ is *trivial*, if it is

- acyclic, and
- every vertex has degree at most three, and
- each connected component contains at most one vertex with degree exactly three.

A class $\mathcal{C} = \mathrm{FORB}(\mathcal{H})$ is *trivial* if all graphs in $\mathcal{H}$ are trivial.

Note that a finite graph $\mathbb{H}$ is trivial, if each connected component of $\mathbb{H}$ is either an isolated vertex, or a (chordless) path, or consists of three (chordless) paths with one endpoint in common.

It is easy to see that every trivial class $\mathrm{FORB}(\mathcal{H})$ is first-order definable. On the other hand, we have the following result.

4.7 LEMMA. ([NUR96]) *For every non-trivial graph $\mathbb{H}$ there are 2-equivalent graphs $\mathbb{G}$ and $\mathbb{G}'$ such that $\mathbb{G}$ has $\mathbb{H}$ as a minor and all minors of $\mathbb{G}'$ are trivial.*

As an immediate consequence we get the following theorem.

4.8 THEOREM. *For every positive integer k, if $\mathcal{C} = FORB(\mathbb{H}_1, \ldots, \mathbb{H}_n)$ and no $\mathbb{H}_i$ is trivial, then the k-th extended vectorization of $\mathsf{Q}_\mathcal{C}$ is not definable in $FO(\mathbf{Q}_k)$. Hence the extended vectorization of $FO(\mathsf{Q}_\mathcal{C})$ is not finitely generated.*

Proof. By Lemma 4.7, there are 2-equivalent graphs $\mathbb{G}$ and $\mathbb{G}'$ such that $\mathbb{H}_1$ is a minor of $\mathbb{G}$, but $\mathbb{G}'$ has no non-trivial minors. Since $\mathbb{G}$ can be obtained from any of its stretchings $\mathbb{G}_e$ by contracting edges, it follows that $\mathbb{H}_1$ is also a minor of each $\mathbb{G}_e$. On the other hand, it is clear that no stretching $\mathbb{G}'_e$ of $\mathbb{G}'$ has non-trivial minors. Thus we conclude that for all e, $\mathbb{G}'_e \in \mathcal{C}$ and $\mathbb{G}_e \notin \mathcal{C}$. The claim follows now from Corollary 3.7 and Theorem 4.2. □

Especially Theorem 4.8 holds for the class PLAN of planar graphs and the class ACYC of acyclic graphs. This is because $\mathrm{PLAN} = \mathrm{FORB}(\mathbb{K}_5, \mathbb{K}_{3,3})$ (see e.g. [Tut84]) and $\mathrm{ACYC} = \mathrm{FORB}(\mathbb{K}_3)$.

4.9 COROLLARY. *For every positive integer k, the k-th extended vectorizations of $\mathsf{Q}_{\mathrm{PLAN}}$ and $\mathsf{Q}_{\mathrm{ACYC}}$ are not definable in $FO(\mathbf{Q}_k)$. Thus the extended vectorizations of $FO(\mathsf{Q}_{\mathrm{PLAN}})$ and $FO(\mathsf{Q}_{\mathrm{ACYC}})$ are not finitely generated.*

5 Conclusion

We conclude the paper by considering a couple of open problems. First of all, it is clear that Corollary 3.7 is far from being a complete characterization for the k-th extended vectorization of a quantifier $\mathsf{Q}_\mathcal{C}$ to be inexpressible in $\mathrm{FO}(\mathbf{Q}_k)$ (or $\mathcal{L}^\omega_{\infty\omega}(\mathbf{Q}_k)$). In particular, the notion of cyclical $\mathrm{FO}_q(\mathbf{Q}_1)$-equivalence is quite technical and has an *ad hoc* flavor. Thus, it is natural to ask whether cyclical equivalence could be replaced by a more general and/or simpler condition. One could even ask whether just mere $\mathrm{FO}_q(\mathbf{Q}_1)$-equivalence is enough:

5.1 Problem. Assume that for every q there are $\mathrm{FO}_q(\mathbf{Q}_1)$-equivalent structures $\mathbb{A} \in \mathcal{C}$ and $\mathbb{B} \notin \mathcal{C}$. Does it follow that, for each $k \geq 1$, the k-th extended vectorization of $\mathsf{Q}_\mathcal{C}$ is not definable in $\mathrm{FO}(\mathbf{Q}_k)$?

Secondly, all our results concern inexpressibility of extended vectorizations of quantifiers. Looking at the proof of Theorem 3.6 the reader should be convinced that taking quotients cannot be avoided in our approach. (Relativization is probably not necessary, but this is of lesser importance.) It would be interesting to know whether it is possible to prove some analogue of Corollary 3.7 also for the plain vectorizations $\mathsf{Q}_{\mathcal{C}^k}$ of a quantifier $\mathsf{Q}_\mathcal{C}$.

Finally, let us discuss possible applications to the second part of Theorem 3.6.

5.2 Problem. Find examples of cyclically $\mathcal{L}^k_{\infty\omega}(\mathbf{Q}_1)$-equivalent non-isomorphic pairs of finite models.

So far we have not been able to find any such examples. Note in particular that in the concrete applications of the previous section we could use counterexample models of color class size two. By the following result of Immerman and Lander, in the case of $\mathcal{L}^\omega_{\infty\omega}(\mathbf{Q}_1)$ it is impossible to use counterexamples of this type. (Recall that $\mathcal{L}^\omega_{\infty\omega}(\mathbf{Q}_1) \equiv \mathcal{L}^\omega_{\infty\omega}(\mathbf{C})$, where $\mathbf{C}$ is the set of all counting quantifiers.)

5.3 Theorem. ([IL90]) *Every class of graphs of color class size at most three is definable in* $\mathcal{L}^\omega_{\infty\omega}(\mathbf{C})$.

Moreover, the class of connected graphs is definable in least fixed point logic LFP, whence the analogue of Theorem 4.5 for $\mathsf{Q}_{\mathrm{CONN}}$ and $\mathcal{L}^\omega_{\infty\omega}(\mathbf{Q}_k)$ fails. The same holds for the planarity quantifier $\mathsf{Q}_{\mathrm{PLAN}}$ as well, since by a recent result of Grohe [Gro98], the class of planar graphs is definable in LFP. On the other hand, the class of rigid graphs is not definable in $\mathcal{L}^\omega_{\infty\omega}(\mathbf{Q}_1)$ (this follows from the main result of [GS96]), so that it is conceivable that the analogue of Theorem 4.5 for $\mathsf{Q}_{\mathrm{RIG}}$ and $\mathcal{L}^\omega_{\infty\omega}(\mathbf{Q}_k)$ holds.

In [CFI92], Cai, Fürer and Immerman gave examples of $\mathcal{L}^k_{\infty\omega}(\mathbf{C})$-equivalent graphs $\mathbb{G}$ and $\mathbb{H}$. They used these graphs for proving that there are PTIME computable properties of graphs which are not definable in $\mathcal{L}^\omega_{\infty\omega}(\mathbf{C})$. It is not possible to define colorings c, c' and indexing functions ι, ι' of these graphs such that the expanded graphs $\mathbb{G}^{c,\iota}$ and $\mathbb{H}^{c',\iota'}$ would still be $\mathcal{L}^\omega_{\infty\omega}(\mathbf{C})$-equivalent. However, the graphs $\mathbb{G}$ and $\mathbb{H}$ are of color class size four, and it is possible to impose a slightly different additional structure to the color blocks in such a way that a modification of Theorem 3.6 goes through. In this way it is possible to get an alternative proof to the result of [Hel96] stating that for each k there is a PTIME computable property of models which is not expressible in $\mathcal{L}^\omega_{\infty\omega}(\mathbf{Q}_k)$.

Acknowledgement

We are grateful to an anonymous referee for useful comments and suggestions that helped improving the readability of the paper.

References

[CFI92] J. Cai, M. Fürer, and N. Immerman. On optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12:389–410, 1992.

[Daw95] A. Dawar. Generalized quantifiers and logical reducibilities. *Journal of Logic and Computation*, 5(2):213–226, 1995.

[DH95] A. Dawar and L. Hella. The expressive power of finitely many generalized quantifiers. *Information and Computation*, 123:172–184, 1995.

[Ebb85] H.-D. Ebbinghaus. Extended logic: the general framework. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, pages 25–76. Springer-Verlag, 1985.

[EF95] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Perspectives in Mathematical Logic. Springer, Berlin, 1995.

[Fag74] R. Fagin. Generalized first-order spectra and polynomial-time recognizable set. In R. M. Karp, editor, *Complexity of Computation, SIAM-AMS Proceedings*, volume 7, pages 43–73, 1974.

[FSV95] R. Fagin, L. Stockmeyer, and M. Vardi. On monadic NP vs. monadic co-NP. *Information and Computation*, 120(1):78–92, July 1995.

[Gro98] M. Grohe. Fixed-point logics on planar graphs. In *Proceedings, 13th IEEE Symposium on Logic in Computer Science*, pages 6–15, 1998.

[GH96] M. Grohe and L. Hella. Double arity hierarchies for transitive closure logics. *Archive for Mathematical Logic*, 35(3):157–172, 1996.

[GS96] Y. Gurevich and S. Shelah. On finite rigid structures. *Journal of Symbolic Logic*, 61(2):549–562, June 1996.

[Han65] W. Hanf. Model-theoretic methods in the study of elementary logic. In J. Addison, L. Henkin, and A. Tarski, editors, *The Theory of Models*, pages 132–145. North Holland, 1965.

[Hel89] L. Hella. Definability hierarchies of generalized quantifiers. *Annals of Pure and Applied Logic*, 43:235–271, 1989.

[Hel96] L. Hella. Logical hierarchies in PTIME. *Information and Computation*, 129(1):1–19, August 1996.

[IL90] N. Immerman and E. Lander. Describing graphs: a first-order approach to graph canonization. In A. Selman, editor, *Complexity Theory Retrospective*, pages 59–81. Springer-Verlag, Berlin/New York, 1990.

[Imm86] N. Immerman. Relational queries computable in polynomial time. *Information and Control*, 68:86–104, 1986.

[Imm87] N. Immerman. Languages which capture complexity classes. *SIAM Journal on Computing*, 16(4):760–778, 1987.

[Kry88] M. Krynicki. Notion of interpretation and nonelementary languages. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 34:541–552, 1988.

[KV95] Ph. G. Kolaitis and J. A. Väänänen. Generalized quantifiers and pebble games on finite structures. *Annals of Pure and Applied Logic*, 74(1):23–75, June 1995.

[Lin66] P. Lindström. First order predicate logic with generalized quantifiers. *Theoria*, 32:186–195, 1966.

[Mos] M. Mostowski. The logic of divisibility. To appear.

[MP95] J. Makowsky and Y. Pnueli. Computable quantifiers and logics over finite structures. In M. Krynicki, M. Mostowski, and L. Szczerba, editors, *Quantifiers: Logics, Models and Computation*, volume I, pages 313–358. Kluwer Academic Publishers, Dordrecht/Boston/London, 1995.

[MS86] A. Mekler and S. Shelah. Stationary logic and its friends II. *Notre Dame Journal of Formal Logic*, 27:39–50, 1986.

[Nur96] J. Nurmonen. On winning strategies with unary quantifiers. *Journal of Logic and Computation*, 6(6):779–798, 1996.

[Ste91] I. Stewart. Comparing the expressibility of languages formed using NP-complete operators. *Journal of Logic and Computation*, 1:305–330, 1991.

[Ste92] I. Stewart. Using the hamiltonian path operator to capture NP. *Journal of Computer and System Sciences*, 45:127–151, 1992.

[Tut84] W. Tutte. *Graph Theory.* Addison-Wesley, 1984.

[Var82] M. Vardi. The complexity of relational query languages. In *Proceedings, 14th ACM Symposium on Theory of Computing*, pages 137–146, 1982.