On vectorizations of unary generalized quantifiers

Kerkko Luosto

1. Introduction

The concept of a generalized quantifier is, in its present generality, due to Lindström [13]. Given a class of structures K and a quantifier symbol Q for K, $\mathrm{FO}(Q)$ is a minimal extension of the first order logic FO where K is definable and which has some of the basic closure properties: $\mathrm{FO}(Q)$ is closed under Boolean combinations, existential quantification and under substitution of formulas for predicates. Much of the research in abstract and finite model theory has been devoted to the search of logics sharing some nice properties of FO, such as compactness, interpolation and its variants, and Löwenheim–Skolem properties (for an overview, see [2]). For some of these properties, it is always possible to extend a logic to another one having the property under consideration, and this extension is unique up to expressive power. This gives rise to some closures of logics, such as the Δ-closure or the Beth closure first systematically studied in [17]. Later, similar closures have been studied also in the context of finite structures (see, e.g., [10]). Due to the minimal nature of quantifier logics, it is no wonder that $\mathrm{FO}(Q)$ does not usually satisfy properties that are related to these interesting closure notions.

Our concern here is the closure referred as Cartesian closure [12] or vectorization of a logic. Informally, vectorization just means a transformation process of classes of structures (or generalized quantifiers) in which elements are replaced by k-tuples, for some fixed $k \in \mathbb{Z}_+$. For example, if K is the class of $\{E\}$-structures $\mathfrak{A}$ such that $E^{\mathfrak{A}}$ is an equivalence relation with prime number of classes, then the k-th vectorization of K, say K^*, is the class of $\{E^*\}$-structures $\mathfrak{M}$ where $(E^*)^{\mathfrak{M}}$ is an equivalence on k-tuples with prime number of equivalence classes. From the complexity-theoretical point of view, membership problem for K^* can easily be reduced to K in this, and similar cases. Even better, the importance of the vectorization in the context of generalized quantifiers is manifested by some results in descriptive complexity theory. Gurevich [4] proposed the question if there is a logic for the complexity class PTIME, not only in some abstract sense, but one with a natural recursive syntax. Dawar [1] showed that such a logic exists iff there is a single generalized quantifier Q such that PTIME is captured by the logic $\mathrm{FO}(\mathcal{Q})$ where $\mathcal{Q}$ is the set of all vectorizations $Q^{(n)}$, $n \in \mathbb{N}^*$, of Q.

Dawar's result seems to call for a systematic study of vectorizations of quantifiers. Not much is known about this subject apart from the result by Hella and Krynicki [6] that $\mathrm{FO}(\mathcal{Q}_n)$ is not Cartesian closed where $\mathcal{Q}_n$ is the collection of all n-ary quantifiers. Also to mention are the results of Hella and Nurmonen [8] in a similar vain, and Kaila's result about the vectorization hierarchy of the Härtig quantifier I in [9, Section 7]. Perhaps the simplest kind of a question one might ask about vectorizing a quantifier is the following:

> Given a quantifier Q, determine if its vectorization hierarchy totally collapses, i.e., if $\mathrm{FO}(\{ Q^{(n)} \mid n \in \mathbb{N} \}) \equiv \mathrm{FO}(Q)$.

The results in this paper seem to suggest that this is quite a rare event. Starting with the simplest case, we study cardinality quantifiers. Let S be a class of cardinals; then C_S is cardinality quantifier for which

$$\mathfrak{M} \models \mathsf{C}_S\, x\, U(x) \iff \left|U^{\mathfrak{M}}\right| \in S.$$

On the one hand it is well-known that the vectorization hierarchy of C_S collapses when S is an eventually periodic set of natural numbers. On the other hand, we shall show that, under strong non-periodicity condition on $S \cap \mathbb{N}$, the vectorization hierarchy of S does not collapse; in fact, even the sublogic $\mathrm{FO}(\mathsf{C}_S^{(2)})$ is unary hierarchical in the class of finite structures.

The material is organized as follows: In the next section, the basic tools and results form [16], as well as the key concepts related to generalized quantifiers, are sketched. In Section 3, periodicity of sequences is briefly discussed. In the last section, the main result is proved using topological machinery.

2. Unary dimension

In this section, we briefly review the central concepts and results from [16], on which our results are based. Indeed, the theorem that $\mathrm{FO}(\mathsf{C}_S^{(2)})$ is unary hierarchical for S strongly non-periodic can be seen as a modification of the result of [16] to the effect that $\mathrm{FO}(\mathsf{I}^{(2)})$ is unary hierarchical. Note that here we use the terms "unary dimension" and "unary hierarchical" instead of "monadic dimension" and "monadically hierarchical", as in [16], because of the confusion with the notions related to second order logic.

A *Lindström quantifier* is a name for a class of structures $K_Q \subseteq \mathrm{Str}(\tau_Q)$ closed under isomorphism such that τ_Q is a finite relational vocabulary. K_Q is called *the defining class of* Q and τ_Q *the vocabulary of* Q. A logic $\mathcal{L}$ is *closed under the* Q*-introduction rule*, if for every vocabulary σ and sequence $\big(\psi_R(\boldsymbol{x}_R)\big)_{R\in\tau_Q}$ of σ-formulas of $\mathcal{L}$ such that $n_R = |\boldsymbol{x}_R|$, for every $R \in \tau_Q$, there is a sentence

$$\varphi = Q(\boldsymbol{x}_R \psi_R(\boldsymbol{x}_R))_{R\in\tau_Q}$$

such that for every $\mathfrak{A} \in \mathrm{Str}(\sigma)$, we have

$$\mathfrak{A} \models \varphi \text{ iff } F(\mathfrak{A}) \in K_Q$$

where the interpreted structure $F(\mathfrak{A})$ has the universe $\mathrm{Dom}(F(\mathfrak{A})) = \mathrm{Dom}(\mathfrak{A})$ and for every $R \in \tau_Q$, it holds that $R^{F(\mathfrak{A})} = \psi_R^{\mathfrak{A}} = \{\, \boldsymbol{a} \in \mathrm{Dom}(\mathfrak{A})^{|\boldsymbol{x}_R|} \mid \mathfrak{A} \models \psi_R[\boldsymbol{a}] \,\}$.

The *arity* of the quantifier Q is $\sup\{ n_R \mid R \in \tau_Q \}$ where n_R is the arity of R, for each $R \in \tau_Q$. The *width* of Q is $\mathrm{wd}(Q) = |\tau_Q|$. Q is *unary*, if it is of arity one, and *simple*, if it is of width one.

If $\mathcal{Q}$ is a set of quantifiers, $\mathrm{FO}(\mathcal{Q})$ is the smallest logic closed under first order construction rules and every Q-introduction rule where $Q \in \mathcal{Q}$. $\mathcal{L}_{\infty\omega}(\mathcal{Q})$ is defined similarly, but also closure under arbitrary disjunctions is required. $\mathrm{FVL}\,(\mathcal{Q})$ is the fragment of $\mathcal{L}_{\infty\omega}(\mathcal{Q})$ of sentences with only finitely many variables. See [11] for more details.

Definition 2.1. Let τ and σ be relational vocabularies.

a) The mapping $\Gamma\colon \mathrm{Str}(\sigma) \to \mathrm{Str}(\tau)$ is called a *Cartesian interpretation of order* $n \in \mathbb{Z}_+$, if the following conditions hold:

   1) There is a bijection $f\colon \tau \to \sigma$ such that if $R \in \tau$ is k-ary, then $f(R)$ is nk-ary.
   2) For every $\mathfrak{A} \in \mathrm{Str}(\tau)$, it holds that $\mathrm{Dom}(\Gamma(\mathfrak{A})) = \mathrm{Dom}(\mathfrak{A}^n)$.
   3) For every $\mathfrak{A} \in \mathrm{Str}(\tau)$ and $R \in \tau$ of arity k, we have

$$R^{\Gamma(\mathfrak{A})} = \{ (\boldsymbol{a}_0, \dots, \boldsymbol{a}_{k-1}) \in \mathrm{Dom}(\Gamma(\mathfrak{A}))^k \mid \boldsymbol{a}_0{}^\frown \cdots {}^\frown\boldsymbol{a}_{k-1} \in f(R)^{\mathfrak{A}} \}.$$

b) A quantifier $Q^{(n)}$ with vocabulary σ is an *n-th vectorization* of a quantifier Q with vocabulary τ, if there is a Cartesian interpretation $\Gamma\colon \mathrm{Str}(\sigma) \to \mathrm{Str}(\tau)$ of order n such that

$$K_{Q^{(n)}} = \{ \mathfrak{A} \in \mathrm{Str}(\tau) \mid \Gamma(\mathfrak{A}) \in K_Q \}.$$

The vectorization $Q^{(n)}$ is of course unique up to renaming of symbols in σ. In the case of vectorizations of cardinality quantifiers we have:

$$\mathfrak{A} \models \mathsf{C}_S^{(n)}\, \boldsymbol{x}(U(\boldsymbol{x})) \quad \text{iff} \quad \left|U^{\mathfrak{A}}\right| \in S$$

where U is a n-ary relation symbol and $\mathfrak{A} \in \mathrm{Str}(\{U\})$.

In quantifier definability theory, different kinds of dimensional invariants have appeared very useful. They have the following property in common:

If $\mathcal{Q}$ is a set of quantifiers of dimension at most n and Q is definable in $\mathrm{FO}(\mathcal{Q})$, then the dimension of Q is at most n.

Dimensions of quantifiers have their background in the papers [19], [5] where the authors implicitly dealt with dimensions related to arities of quantifiers. This idea was abstracted in [7] by the following definition: The *strict dimension,* $\dim(Q)$*,* of Q is the least n such that Q is definable in some $\mathrm{FO}(\mathcal{Q})$ with $\mathcal{Q}$ a set of n-ary quantifiers. Here, we are going to use the notion of unary dimension from [16] (cf. [18]) which has a similar relation to width of unary quantifiers than strict dimension has to arities.

Unary quantifiers have some special properties which are related to the fact that unary structures are characterized in a simple way by cardinal invaraints. Let τ be a finite, unary vocabulary and $\mathfrak{M} \in \mathrm{Str}(\tau)$. For every $\sigma \subseteq \tau$, let

$$U_{\mathfrak{M}}(\sigma) = \{ a \in \mathrm{Dom}(\mathfrak{M}) \mid \sigma = \{ R \in \tau \mid a \in R^{\mathfrak{M}} \} \}.$$

Observe that $\{\, U_{\mathfrak{M}}(\sigma) \mid \sigma \subseteq \tau \,\} \smallsetminus \{\emptyset\}$ is the partition of the universe according to isomorphism types of the elements. Let $n = 2^{|\tau|}$ and let us fix a bijection $f_\tau \colon \mathcal{P}(\tau) \to n$. Put

$$\kappa_{\mathfrak{M}} \colon n \to \text{Card},\ \kappa_{\mathfrak{M}}(i) = \bigl|U_{\mathfrak{M}}(f_\tau^{-1}(i))\bigr|.$$

Then $\kappa_{\mathfrak{M}}$ characterizes $\mathfrak{M}$ up to isomorphism, i.e., if $\mathfrak{N} \in \text{Str}(\tau)$ and $\kappa_{\mathfrak{N}} = \kappa_{\mathfrak{M}}$, then $\mathfrak{M} \cong \mathfrak{N}$.

Suppose Q is a generalized quantifier with the vocabulary τ. Restricting our attention to finite structures, denote

$$\mathcal{R}(Q, \omega) = \{\, \kappa_{\mathfrak{M}} \mid \mathfrak{M} \in K_Q,\ \mathfrak{M} \text{ finite} \,\}.$$

Similarly, if ϑ is a τ-sentence of any logic $\mathcal{L}$, then

$$\mathcal{R}_\tau(\vartheta, \omega) = \{\, \kappa_{\mathfrak{M}} \mid \mathfrak{M} \in \text{Str}(\tau) \text{ finite}, \mathfrak{M} \models \vartheta \,\}.$$

The subscript is needed to remove the possible ambiguity which arises because $\vartheta \in \mathcal{L}[\sigma]$, for all $\sigma \supseteq \tau$. We shall see that the definability problems of unary quantifiers can be reduced to combinatorial problems of relations on $\mathbb{N}$.

The definition of unary dimension is based on the notion of relative rank. Since the latter is rather involved, we give an intuitive explanation what it means that a relation $R \subseteq \mathbb{Z}^n$ has rank at most k. In effect, it means that you can determine if $\boldsymbol{a} \in \mathbb{Z}^n$ belongs to R using the following procedure: The width n of $\boldsymbol{a}$ can be thought of large, so you are allowed to use assistants, each of whom has a finite pack of differently coloured cards with hers/his name printed on them. Every assistant corresponds to a sequence $\boldsymbol{U} = (U_0, \ldots, U_{k-1})$ of disjoint subsets of indices. When you are given $\boldsymbol{a} = (a_0, \ldots, a_{n-1})$, you divide the problem in subtasks, giving the assistant dedicated for $\boldsymbol{U}$ the sequence $\boldsymbol{b}_{\boldsymbol{U}} = (\sum_{j \in U_0} a_j, \ldots, \sum_{j \in U_{k-1}} a_j)$. The assistants then work independently, and each returns one coloured card from one's pack. Then you can judge if $\boldsymbol{a} \in R$ only on basis of the cards you received.

Definition 2.2. Let $\langle M, + \rangle$ be a commutative monoid, $n \in \mathbb{Z}_+$ and $R \subseteq M^n$. For any disjoint family $\boldsymbol{U} = (U_i)_{i \in I}$ of subsets of n and $\boldsymbol{a} = (a_0, \ldots, a_{n-1}) \in M^n$, denote $\boldsymbol{s}(\boldsymbol{a}, \boldsymbol{U}) = (\sum_{j \in U_i} a_j)_{i \in I}$. For every $l \in \mathbb{N}$ with $1 \leq l \leq n$, let $\mathcal{U}_{n,l}$ be the set of sequences $\boldsymbol{U} = (U_0, \ldots, U_{l-1})$ of disjoint subsets of n. Then the *rank of R relative to $\langle M, + \rangle$*, in symbols $r_+(R)$, is the least $l \in \mathbb{N}$, $1 \leq l \leq n$, for which the following holds: There are finite colourings $\chi_{\boldsymbol{U}} \colon M^l \to F_{\boldsymbol{U}}$, for $\boldsymbol{U} \in \mathcal{U}_{n,l}$, such that R is congruent with the colouring $\chi \colon M^n \to \prod_{\boldsymbol{U} \in \mathcal{U}_{n,l}} F_{\boldsymbol{U}}$, $\chi(\boldsymbol{a}) = \bigl(\chi_{\boldsymbol{U}}(\boldsymbol{s}(\boldsymbol{a}, \boldsymbol{U}))\bigr)_{\boldsymbol{U} \in \mathcal{U}_{n,l}}$, i.e., $\chi(\boldsymbol{a}) = \chi(\boldsymbol{b})$ and $\boldsymbol{a} \in R$ imply $\boldsymbol{b} \in R$, for each $\boldsymbol{a}, \boldsymbol{b} \in M^n$. The function χ is denoted by $\nabla^+_{\boldsymbol{U} \in \mathcal{U}_{n,l}} \chi_{\boldsymbol{U}}$.

We state some basic properties of the relative rank (for proofs, see [16]). If $\langle N, + \rangle$ is a commutative monoid such that $\langle M, + \rangle$ is a submonoid of $\langle N, + \rangle$, then the rank of R relative to $\langle M, + \rangle$ is the same as relative to $\langle N, + \rangle$. Since we are concerned with cases $\langle M, + \rangle = \langle \mathbb{Z}, + \rangle$ of $\langle M, + \rangle = \langle \mathbb{N}, + \rangle$, it does not matter in which monoid we calculate the rank.

Proposition 2.3. *Consider the rank relative to $\langle \mathbb{Z}, + \rangle$. Let $R \subseteq \mathbb{Z}^m$ and $S \subseteq \mathbb{Z}^n$ be relations.*

a) Let R be a Boolean combination of relations $R_0, \ldots, R_{k-1} \subseteq \mathbb{Z}^m$ where $k \in \mathbb{Z}_+$. Then $r_+(R) \leq \max_{i \in k} r_+(R_i)$.

b) If $m = n$ and $|R \triangle S| < \omega$, then $r_+(R) = r_+(S)$.

c) Suppose $f \colon m \to n$ is an injection such that $S = \{ \boldsymbol{a} \in \mathbb{Z}^n \mid \boldsymbol{a} \circ f \in R \}$. Then $r_+(R) = r_+(S)$.

d) If $T = \{ \boldsymbol{a}\hat{\ }\boldsymbol{b} \mid \boldsymbol{a} \in R, \boldsymbol{b} \in S \} \subseteq \mathbb{Z}^{m+n}$ where R and S are non-empty, then $r_+(T) = \max\{r_+(R), r_+(S)\}$.

e) Let $\boldsymbol{a} \in \mathbb{Z}^n$ and $R' = \{ \boldsymbol{c} \in \mathbb{Z}^m \mid \boldsymbol{c} + \boldsymbol{a} \in R \}$ where $+$ stands for vector addition. Then $r_+(R') = r_+(R)$. □

Definition 2.4. Let Q be a unary Lindström quantifier. The *unary dimension* of Q *on finite structures* is $\operatorname{udim}_\omega(Q) = r_+(\mathcal{R}(Q, \omega))$. The *unary dimension* $\operatorname{udim}_\omega(\mathcal{L})$ of a logic $\mathcal{L}$ *on finite structures* is supremum over all $\operatorname{udim}_\omega(Q)$ with Q a unary Lindström quantifier definable in $\mathcal{L}$. $\mathcal{L}$ is *unary hierarchical*, if $\operatorname{udim}_\omega(\mathcal{L}) = \omega$.

The following theorem from [16] justifies the concept of unary dimension.

Theorem 2.5. *Let Q be a unary Lindström quantifier.*

a) If Q is expressible in $\mathcal{L}$ where $\mathcal{L} = \mathrm{FVL}\,(\mathcal{Q})$ with $\mathcal{Q}$ a finite set of unary Lindström quantifiers, then
$$\operatorname{udim}_\omega(Q) \leq \max_{q \in \mathcal{Q} \cup \{\exists\}} \operatorname{udim}_\omega(q).$$

b) If $\operatorname{udim}_\omega(Q) < 2^k$ where $k \in \mathbb{N}^*$, then there is a finite set of unary Lindström quantifiers $\mathcal{Q}$ of width k such that Q is definable in $\mathrm{FO}(\mathcal{Q})$ on finite structures. □

In [16], the described machinery was applied to proving that $\mathrm{FO}(\mathsf{I}^{(2)})$ is unary hierarchical on finite structures. We sketch the idea of the proof. The goal is to show that there are unary quantifiers of arbitrary high unary dimension on finite structures which are definable in $\mathrm{FO}(\mathsf{I}^{(2)})$, or equivalently, that for each $n \in \mathbb{N}$, there is $\varphi \in \mathrm{FO}(\mathsf{I}^{(2)})$ with τ a finite unary vocabulary such that $r_+(\mathcal{R}(\varphi)) \geq n$. For $\boldsymbol{c} \in \mathbb{Q}^n$, $n \in \mathbb{Z}_+$, consider the relation
$$\mathcal{R}_{\boldsymbol{c}} = \{ \boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{c} \cdot \boldsymbol{x} = 0 \}.$$
Then if $n = 2^k$ for some $k \in \mathbb{N}$, there is $\varphi_{\boldsymbol{c}} \in \mathrm{FO}(\mathsf{I}^{(2)})$ such that
$$\mathcal{R}_{\boldsymbol{c}} \cap (\mathbb{N}^n \smallsetminus \{\boldsymbol{0}\}) = \mathcal{R}(\varphi_{\boldsymbol{c}}).$$
Now it can be proved that for suitable $\boldsymbol{c}$, $r_+(\mathcal{R}_{\boldsymbol{c}} \cap (\mathbb{N}^n \smallsetminus \{\boldsymbol{0}\})) \geq n$.

We fix the notation $\mathcal{R}_{\boldsymbol{c}}$ and some notation of linear spaces for the rest of the paper. Let $X \subseteq V$ where V is a vector space over the field of coefficients K. Then $\mathrm{sp}_K(X)$ is the span of X, or the subspace generated by X. For $n \in \mathbb{N}^*$, put $X_n = \{0,1\}^n \subseteq \mathbb{Q}^n$ and

$$(*), \qquad Y_n = \bigcup \{ \mathrm{sp}_{\mathbb{Q}}(X) \mid X \in [X_n]^{n-1} \}$$

i.e., Y_n is the union of all subspaces of $\mathbb{Q}^n$ generated by $n-1$ vectors whose components are all either zero or one.

The following theorem is essentially [16, Theorem 5.3.] The difference is that in [16], the theorem was proved for $\mathcal{R}_{\boldsymbol{c}} \cap \mathbb{N}^n$ instead of $\mathcal{R}_{\boldsymbol{c}}$ and a certain condition on $\boldsymbol{c}$ was needed to ensure that $\mathcal{R}_{\boldsymbol{c}} \cap \mathbb{N}^n \neq \emptyset$. The change only simplifies the labourous proof using Ramsey theory which we omit here.

Theorem 2.6. *Let $\boldsymbol{c} \in \mathbb{Q}^n \setminus Y_n$ where $n \in \mathbb{N}$ and $n \geq 2$. Then $r_+(\mathcal{R}_{\boldsymbol{c}}) = n$.* □

3. Periodicity

In this short section, we briefly review some concepts related to the periodicity of sets. The notion of strong non-periodicity is introduced and, for comparison, almost periodicity, which is a thoroughly studied notion in combinatorics of words (see [14, 15]), is discussed.

We identify a tuple $w = (w_0, \ldots, w_{n-1}) \in C^n$ with a function $w\colon \{0, \ldots, n-1\} \to C$, $w(i) = w_i$. In this context, the tuple $w \in C^n$ is often called a *word* over the alphabet C of length n. Similarly, $\chi\colon \mathbb{N} \to C$ may be called a colouring of $\mathbb{N}$ or an infinite word, depending on the context.

Definition 3.1. Let $\chi\colon \mathbb{N} \to C$ be a colouring of $\mathbb{N}$, $k \in \mathbb{N}$, and $w\colon \{0, \ldots, k-1\} \to C$. Denote by T^χ_w the set of all $n \in \mathbb{N}$ such that for all $i \in \{0, \ldots, k-1\}$, we have that $\chi(m+1) = w(i)$. We call T^χ_w *the set of occurences of the word w*.

a) *Word w occurs almost periodically in χ*, if there exists $r \in \mathbb{N}^*$ such that for every $m \in \mathbb{N}$, it holds that $T^\chi_w \cup \{m, m+1, \ldots, m+r-1\} \neq \emptyset$.

b) *Word w occurs strongly non-periodically in χ*, if T^χ_w is infinite and for every $r \in \mathbb{N}^*$ there is $m \in \mathbb{N}$ such that $m - r \geq 0$ and $T^\chi_w \cup [m-r, m+r] \cap \mathbb{N} = \{m\}$.

c) The colouring χ is *almost periodic* if every word $w\colon \{0, \ldots, k-1\} \to C$ ($k \in \mathbb{N}^*$) occurs almost periodically in χ. The colouring χ is *strongly non-periodic* if there is a word $w\colon \{0, \ldots, k-1\} \to C$ ($k \in \mathbb{N}^*$) occuring strongly non-periodically in χ. A set $S \subseteq \mathbb{N}$ is *almost periodic* (resp. *strongly non-periodic*) is its characteristic function $\chi_S : \mathbb{N} \to \{0,1\}$ is almost periodic (resp. strongly non-periodic).

Note that in literature of combinatorics of words, one often uses terminology of dynamical systems. Then χ is called *uniformly recurrent* iff it is almost periodic, and χ is called *recurrent* iff every word w occuring in χ occurs infinitely often in χ.

It is immediate that a word $w \in C^n$ cannot occur simultaneously almost periodically and strongly non-periodically in a colouring $\chi\colon \mathbb{N} \to C$. Thus, if χ is strongly non-periodic, it is not almost periodic. However, it may be neither strongly non-periodic nor almost periodic, and we provide examples from each of the three categories.

Example 3.2. a) The best-known example of an almost periodic sequence is the *Thue–Morse sequence*, which is the characteristic function $\chi_M\colon \mathbb{N} \to \{0,1\}$ of the Thue–Morse set

$$M = \{ \sum_{\alpha \in A} 2^\alpha \mid |A| \text{ odd} \}.$$

The Thue–Morse sequence has been thoroughly studied, and the proof of its almost periodicity and other basic properties can be found in many combinatorial textbooks, e.g. [14, 15].

b) Consider what could be called the *discrete Cantor set* C:

$$C = \{ 2 \sum_{\alpha \in A} 3^\alpha \mid A \subseteq \mathbb{N} \text{ finite} \}.$$

In other words, C consists of $n \in \mathbb{N}$ such that the ternary expansion of n does not contain the digit 1. We observe that C has the following periodicity property: For every $x \in \mathbb{N}$ and $c \in C$ with $x < 3^k \mid c$ we have $\chi_C(x) = \chi_C(x+c)$.

Let $w \in \{0,1\}^s$ be a word occurring in χ_C. If w consists of zeros only, then w occurs almost periodically in χ_C: Pick $l \in \mathbb{N}$ with $3^l \geq s$, then for every $x \in \mathbb{N}$ with $3^{l+1} \mid x$ we have $[x+3^l, x+2\cdot 3^l - 1] \cap C = \emptyset$ implying $x \in T_w^\chi$. On the other hand, suppose that 1 occurs in the word w. Then the same argument that there are large gaps in C implies that w cannot occur almost periodically in χ_C. Now fix $x_0 \in T_w^\chi$ and $k \in \mathbb{N}$ such that $x_0 + s \leq 3^k$. For every $x \in T_w^{\chi_C}$, there exists $x_1 \in C$ such that $|x - x_1| \leq 3^k$ and $3^k \mid x_1$, as 1 occurs in w. Now the periodicity property we observed in the beginning implies that $x_0 + x_1, x_0 + x_1 + 3^k \in T_w^{\chi_C}$, so $|T_w^{\chi_C} \cap [x - 3^{k+1}, x + 3^{k+1}]| \geq 2$. Consequently, w does not occur strongly non-periodically in χ_C, either. For example, the word (1) is such. Hence, C is neither almost periodic nor strongly non-periodic.

c) It is easy to construct concrete examples of strongly non-periodic sets. To have one, let $P\colon \mathbb{N} \to \mathbb{N}$ be a polynomial function of degree at least 2 with coeffiecents in $\mathbb{N}$, and let $S = \operatorname{rg}(P)$. Then $\lim_{n\to\infty}(P(n+1) - P(n)) = \infty$, and so (1) occurs strongly non-periodically in χ_S, i.e., S is strongly non-periodic.

In fact, most $S \subseteq \mathbb{N}$ are strongly non-periodic, in the following probabilistic sense.

Proposition 3.3. *Select $S \subseteq \mathbb{N}$ randomly from uniform distribution, i.e., the events $A_n = \{n \in S\}$, $n \in \mathbb{N}$, are independent and have probability $\frac{1}{2}$. Then almost surely S is strongly non-periodic, i.e., $\mathbb{P}\{S \text{ is strongly non-periodic}\} = 1$.*

Proof. It suffices to show that almost surely (1) occurs strongly non-periodically in χ_S. Let $r \in \mathbb{Z}_+$. Consider the events $A_k = \{S \cap [3kr, (3k+2)r] = \{3(k+1)r\}\}$, for $k \in \mathbb{N}$. Then the events A_k are independent and $\mathbb{P}(A_k) = (\frac{1}{2})^{2r+1}$. Hence

$$\mathbb{P}(\bigcup_{j\in\mathbb{N}} A_j) = 1 - \mathbb{P}(\bigcap_{j\in\mathbb{N}} \overline{A_j}) = 1 - \lim_{k\to\infty} \mathbb{P}(\bigcap_{j=0}^{k} \overline{A_j})$$
$$= 1 - \lim_{k\to\infty} \left(1 - \left(\frac{1}{2}\right)^{2r+1}\right)^{k+1} = 1 - 0 = 1.$$

Consequently, almost surely there is $m \in \mathbb{N}$ such that $m \geq r$ and $S \cap [m-r, m+r] = \{m\}$. By σ-additivity of probability, almost surely this holds for all $r \in \mathbb{Z}_+$ simultaneously, i.e., S is strongly non-periodic. □

4. Topological dynamics

Our aim is to make the best use of the techniques for proving that $\mathrm{FO}(\mathsf{I}^{(2)})$ is unary hierarchical on finite structures [16, Theorem 5.4] in order to prove that $\mathrm{FO}(Q^{(2)})$ it that, too, for certain cardinality quantifiers. The former result used Theorem 2.6 directly in the sense that the appropriate model classes are definable in $\mathrm{FO}(\mathsf{I}^{(2)})$. Here, we cannot rely on direct definability, but we are able to approximate the circumstances in [16, Theorem 5.4], getting the desired result in an indirect way. The tools are topological and similar than some soft tools used in Ramsey theory, hence the section title Topological dynamics (cf. [3]).

For $\boldsymbol{c} \in \mathbb{Q}^n$ and $S \subseteq \mathbb{Z}$, set

$$\mathcal{R}_{\boldsymbol{c},S} = \{\, \boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{c} \cdot \boldsymbol{x} \in S \,\}.$$

In rough terms, we are going to show that if S is sparse, but infinite, in $\mathbb{N}$, then it is possible to approximate $\mathcal{R}_{\boldsymbol{c}}$ by $\mathcal{R}_{\boldsymbol{c},S}$ and therefore $\mathcal{R}_{\boldsymbol{c},S}$ is at least as complex as $\mathcal{R}_{\boldsymbol{c}}$. To make this precise, we need to introduce concepts from topological dynamics. The logical definability part is easy:

Lemma 4.1. *Let $S \subseteq \mathbb{N}$, $n = 2^k$ for some $k \in \mathbb{Z}_+$ and $\boldsymbol{c} = (c_0, \ldots, c_{n-1}) \in \mathbb{N}^n$. Then there is $\varphi_{\boldsymbol{c},S} \in \mathrm{FO}(\mathsf{C}_S^{(2)})[\tau]$ with τ unary such that $\mathcal{R}_\tau(\varphi_{\boldsymbol{c},S}) = (\mathcal{R}_{\boldsymbol{c},S} \cap \mathbb{N}^n) \setminus \{\boldsymbol{0}\}$.*

Proof. Fix unary relation symbols $U_0, \ldots, U_{k-1}$ and put $\tau = \{U_0, \ldots, U_{k-1}\}$. Denote $m = \max\{c_0, \ldots, c_{n-1}\}$. For each $i \in \{0, \ldots, n-1\}$, we know that we can write a quantifier-free formula $\alpha_i(x)$ of the vocabulary τ such that $\alpha_i^{\mathfrak{M}} = U_{\mathfrak{M}}(f_\tau^{-1}(i))$, for every τ-structure $\mathfrak{M}$. Consider the sentence φ:

$$\exists y_0 \ldots \exists y_{m-1} \big(\bigwedge_{\substack{i,j\in\{0,\ldots,m-1\}\\ i\neq j}} y_i \neq y_j \wedge C_S^{(2)} xy \bigvee_{i=0}^{n-1} \bigvee_{j=0}^{c_i-1} (\alpha_i(x) \wedge y = y_j) \big).$$

Now if $\mathfrak{M}$ is a finite τ-structure with at least m elements, then $\kappa_{\mathfrak{M}} \in \mathcal{R}_\tau(\varphi)$ iff $\mathfrak{M} \models \varphi$ iff $\boldsymbol{c} \cdot \kappa_{\mathfrak{M}} \in S$ iff $\kappa_{\mathfrak{M}} \in \mathcal{R}_{\boldsymbol{c},S}$. Since there are only finitely many isomorphism types of τ-structures with less that m elements, we can patch φ so that we get the desired sentence $\varphi_{\boldsymbol{c},S}$. □

The n-ary relations on $\mathbb{Z}$ are points of the set $\mathcal{P}(\mathbb{Z}^n)$. We endow $\mathcal{P}(\mathbb{Z}^n)$ with its natural topology $\mathcal{T}_n$, i.e., thinking of $2 = \{0, 1\}$ as a discrete space and ${}^2(\mathbb{Z}^n)$ as a product space, $\mathcal{T}_n$ is the topology induced by the canonical bijection

$$f\colon \mathcal{P}(\mathbb{Z}^n) \to {}^2(\mathbb{Z}^n),\ f(R)(\boldsymbol{x}) = \begin{cases} 1, & \text{for } \boldsymbol{x} \in R \\ 0, & \text{for } \boldsymbol{x} \in \mathbb{Z} \smallsetminus R. \end{cases}$$

By Tychonoff's theorem, $(\mathcal{P}(\mathbb{Z}^n), \mathcal{T}_n)$ is a compact topological space. For $R \in \mathcal{P}(\mathbb{Z}^n)$ and $C \subseteq \mathbb{Z}$, denote

$$U(R, C) = \{\, S \subseteq \mathbb{Z}^n \mid S \cap C^n = R \cap C^n \,\}.$$

It is routine to verify that the family

$$\{\, U(R, C) \mid C \subseteq \mathbb{Z} \text{ finite} \,\}$$

or even

$$\{\, U(R, C) \mid C = [-n, n] \cap \mathbb{Z} \text{ for some } n \in \mathbb{N} \,\}$$

forms a neighbourhood basis for R. The next result links relative ranks into discussion.

Lemma 4.2. *Let $n, k \in \mathbb{Z}_+$. For $\boldsymbol{U} \in \mathcal{U}_{n,k}$, endow a set $F_{\boldsymbol{U}}$ with discrete topology and $X_{\boldsymbol{U}} = {}^{(\mathbb{Z}^k)}F_{\boldsymbol{U}}$, $F = \prod_{\boldsymbol{U} \in \mathcal{U}_{n,k}} F_{\boldsymbol{U}}$, $X = {}^{(M^n)}F$ with corresponding product topologies. Then the function*

$$\nabla^+\colon \prod_{\boldsymbol{U} \in \mathcal{U}_{n,k}} X_{\boldsymbol{U}} \to X,\ \big((\chi_{\boldsymbol{U}})_{\boldsymbol{U} \in \mathcal{U}_{n,k}}\big) \mapsto \nabla^+_{\boldsymbol{U} \in \mathcal{U}_{n,l}} \chi_{\boldsymbol{U}}$$

is continuous.

Proof. For $\boldsymbol{V} \in \mathcal{U}_{n,k}$, $\boldsymbol{b} \in M^k$ and $\boldsymbol{a} \in M^n$, the projections

$$p_{\boldsymbol{V}}\colon \prod_{\boldsymbol{U} \in \mathcal{U}_{n,k}} X_{\boldsymbol{U}} \to X_{\boldsymbol{V}}, \quad p\big((\chi_{\boldsymbol{U}})_{\boldsymbol{U} \in \mathcal{U}_{n,k}}\big) = \chi_{\boldsymbol{V}}$$

$$q_{b,\boldsymbol{U}}\colon X_{\boldsymbol{U}} \to F_{\boldsymbol{U}}, \quad q_{b,\boldsymbol{U}}(\xi) = \xi(\boldsymbol{b}),$$

$$p^*_{\boldsymbol{V}}\colon F \to F_{\boldsymbol{V}}, \quad p^*_{\boldsymbol{V}}\big((\boldsymbol{c}_{\boldsymbol{U}})_{\boldsymbol{U} \in \mathcal{U}_{n,k}}\big) = c_{\boldsymbol{V}},$$

and

$$q^*_{\boldsymbol{a}}\colon X \to F, \quad q^*(\chi) = \chi(\boldsymbol{a})$$

are continuous. Let $\boldsymbol{a} \in M^n$, $\boldsymbol{V} \in \mathcal{U}_{n,k}$ and $\bar{\chi} = (\chi_{\boldsymbol{U}})_{\boldsymbol{U} \in \mathcal{U}_{n,k}} \in \prod_{\boldsymbol{U} \in \mathcal{U}_{n,k}} X_{\boldsymbol{U}}$. Then

$$\begin{aligned}(q_{\boldsymbol{s}(\boldsymbol{a},\boldsymbol{V}),\boldsymbol{V}} \circ p_{\boldsymbol{V}})(\bar{\chi}) &= q_{\boldsymbol{s}(\boldsymbol{a},\boldsymbol{V}),\boldsymbol{V}}(\chi_{\boldsymbol{V}}) \\ &= \chi_{\boldsymbol{V}}(\boldsymbol{s}(\boldsymbol{a},\boldsymbol{V})) = p^*_{\boldsymbol{V}}\big((\nabla^+_{\boldsymbol{U} \in \mathcal{U}_{n,k}} \chi_{\boldsymbol{U}})(\boldsymbol{a})\big) \\ &= p^*_{\boldsymbol{V}}(q^*_{\boldsymbol{a}}(\nabla^+(\bar{\chi}))) = (p^*_{\boldsymbol{V}} \circ q^*_{\boldsymbol{a}} \circ \nabla^+)(\bar{\chi}),\end{aligned}$$

implying $q_{\boldsymbol{s}(\boldsymbol{a},\boldsymbol{V}),\boldsymbol{V}} \circ p_{\boldsymbol{V}} = p^*_{\boldsymbol{V}} \circ q^*_{\boldsymbol{a}} \circ \nabla^+$. The composition of projections on the left is continuous, for every $\boldsymbol{V} \in \mathcal{U}_{n,k}$, so $q^*_{\boldsymbol{a}} \circ \nabla^+$ is continuous. Since all the component functions $q^*_{\boldsymbol{a}} \circ \nabla^+$, $\boldsymbol{a} \in M^n$, are continuous, we get that ∇^+ is continuous. □

Note that for $n, k \in \mathbb{Z}_+$, $k < n$, the set

$$K_{n,k} = \{\, R \in \mathcal{P}(\mathbb{Z}^n) \mid r_+(R) \leq k \,\}$$

contains all the finite relations, so that it is dense in $(\mathcal{P}(\mathbb{Z}^n), \mathcal{T}_n)$. However, it is known to be a proper subset of $\mathcal{P}(\mathbb{Z}^n)$, so it cannot be closed. Fortunately, there are appropriate subsets of $K_{n,k}$ which are closed. For $n, k, s \in \mathbb{Z}_+$, $n \geq k$, let $K(n,k,s)$ be the set of all relations $R \subseteq \mathbb{Z}^n$ which are congruent with $\chi = \nabla^+_{\boldsymbol{U} \in \mathcal{U}_{n,k}} \chi_{\boldsymbol{U}}$ for some $\chi_{\boldsymbol{U}} \colon \mathbb{Z}^k \to s$, $\boldsymbol{U} \in \mathcal{U}_{n,k}$ (so that the number of colours is restricted to s).

Lemma 4.3. *Let $n, k \in \mathbb{Z}_+$, $n \geq k$.*

a) $\bigcup_{s \in \mathbb{N}} K(n,k,s) = K_{n,k}$,

b) Let $R \subseteq \mathbb{Z}^n$, $\boldsymbol{a} \in \mathbb{Z}^n$ and $s \in \mathbb{Z}_+$. Then $R \in K(n,k,s)$ iff $R^* = \{\, \boldsymbol{x} + \boldsymbol{a} \mid \boldsymbol{x} \in R \,\} \in K(n,k,s)$.

c) $K(n,k,s)$ is closed in the topological space $(\mathcal{P}(\mathbb{Z}^n), \mathcal{T}_n)$.

Proof. a) Clearly, $\bigcup_{s \in \mathbb{N}} K(n,k,s) \subseteq K_{n,k}$. In the other direction, if $R \in K_{n,k}$, then there are $\chi_{\boldsymbol{U}} \colon \mathbb{Z}^k \to F_{\boldsymbol{U}}$, $\boldsymbol{U} \in \mathcal{U}_{n,k}$, with finite $F_{\boldsymbol{U}}$'s such that R is congruent with $\nabla^+_{\boldsymbol{U} \in \mathcal{U}_{n,k}} \chi_{\boldsymbol{U}}$. Choose $s \in \mathbb{N}$ and injections $f_{\boldsymbol{U}} \colon F_{\boldsymbol{U}} \to s$. Then it is easy to see that R is congruent with $\nabla^+_{\boldsymbol{U} \in \mathcal{U}_{n,k}} (f_{\boldsymbol{U}} \circ \chi_{\boldsymbol{U}})$ implying that $R \in K(n,k,s)$.

b) As $R = \{\, \boldsymbol{x} - \boldsymbol{a} \mid \boldsymbol{x} \in R^* \,\}$, it is enough to show the implication left to right. Suppose $R \in K(n,k,s)$. Choose $\chi_{\boldsymbol{U}} \colon \mathbb{Z}^k \to s$, $\boldsymbol{U} \in \boldsymbol{U}_{n,k}$ such that R is congruent with $\nabla^+_{\boldsymbol{U} \in \mathcal{U}_{n,k}} \chi_{\boldsymbol{U}}$. Put $\boldsymbol{b}_{\boldsymbol{U}} = \boldsymbol{s}(\boldsymbol{a}, \boldsymbol{U})$ and $\chi^*_{\boldsymbol{U}} \colon \mathbb{Z}^k \to s$, $\chi^*_{\boldsymbol{U}}(\boldsymbol{x}) = \chi_{\boldsymbol{U}}(\boldsymbol{x} - \boldsymbol{b}_{\boldsymbol{U}})$. Then for every $\boldsymbol{x} \in \mathbb{Z}^n$, we have that $\chi^*(\boldsymbol{x}) = \chi(\boldsymbol{x} - \boldsymbol{a})$. It follows that if $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{Z}^n$, $\chi^*(\boldsymbol{x}) = \chi^*(\boldsymbol{y})$ and $\boldsymbol{x} \in R^*$, then $\chi(\boldsymbol{x} - \boldsymbol{a}) = \chi(\boldsymbol{y} - \boldsymbol{a})$ and $\boldsymbol{x} - \boldsymbol{a} \in R$, so, by congruence, we get $\boldsymbol{y} - \boldsymbol{a} \in R$, or equivalently, $\boldsymbol{y} \in R^*$. Consequently, R^* is congruent with $\nabla^+_{\boldsymbol{U} \in \mathcal{U}_{n,k}} \chi^*_{\boldsymbol{U}}$, implying $R^* \in K(n,k,s)$.

c) Let $F = {}^{\mathcal{U}_{n,k}}s$, $X = {}^{(\mathbb{Z}^n)}F$ and $X_0 = {}^{(\mathbb{Z}^k)}s$. Let $R \subseteq \mathbb{Z}^n$. By definition, $R \in K(n,k,s)$ iff there is $\bar{\chi} \in {}^{\mathcal{U}_{n,k}}X_0$ such that R is congruent with $\nabla^+(\bar{\chi})$, i.e., $\gamma \circ \nabla^+(\bar{\chi})$ is the characteristic function of R for appropriate $\gamma \colon F \to 2$. For $\gamma \colon F \to 2$, put $h_\gamma \colon X \to$

$\mathcal{P}(\mathbb{Z}^n)$, $h_\gamma(\xi) = (\gamma \circ \xi)^{-1}\{1\}$. Then h_γ is continuous. Furthermore, $R \in K(n,k,s)$ iff for some $\gamma\colon F \to 2$ and $\bar{\chi} \in {}^{\mathcal{U}_{n,k}}X_0$, we have that $R = h_\gamma(\nabla^+(\bar{\chi}))$. In other words,

$$K(n,k,s) = \bigcup_{\gamma\colon F \to 2} \operatorname{rg}(h_\gamma \circ \nabla^+).$$

The right-hand side of the equation is a finite union of compact sets, since $h_\gamma \circ \nabla^+$ is continuous (by previous lemma) and ${}^{\mathcal{U}_{n,k}}X_0$ is compact, by Tychonoff's theorem. Hence, $K(n,k,s)$ is closed in the Hausdorff space $(\mathcal{P}(\mathbb{Z}^n), \mathcal{T}_n)$. $\square$

For a relation $R \subseteq \mathbb{Z}^n$, we use the notation $\widehat{R}$ for the closure of the set of all translates $R + \boldsymbol{a} = \{\, \boldsymbol{x} + \boldsymbol{a} \mid \boldsymbol{x} \in R \,\}$, $\boldsymbol{a} \in \mathbb{Z}^n$, of R.

Proposition 4.4. *Let $R, S \subseteq \mathbb{Z}^n$ be relations such that $S \in \widehat{R}$. Then $r_+(S) \leq r_+(R)$.*

Proof. Put $k = r_+(R)$, whence $R \in K_{n,k}$. By Lemma 4.3.a, we can pick $\ell \in \mathbb{N}^*$ such that $R \in K(n,k,\ell)$. By Lemma 4.3.b, we have $\{\, R + \boldsymbol{a} \mid \boldsymbol{a} \in \mathbb{Z}^k \,\} \subseteq K(n,k,\ell)$. As $K(n,k,\ell)$ is closed in $(\mathcal{P}(\mathbb{Z}^n), \mathcal{T}_n)$, by Lemma 4.3.c, we have that $\widehat{R} \subseteq K(n,k,\ell)$. In particular, $S \in \widehat{R} \subseteq K(n,k,\ell)$, which implies $r_+(S) \leq k = r_+(R)$. $\square$

We start with a combinatorial application of the preceding proposition, which is a simplified version of a lemma needed for proving the intended result. This proposition is proved for mere presentational purposes, as it might be instructive to first see an easy result which pinpoints the ideas used in this and similar proofs in a clear setting.

Proposition 4.5. *Let $\boldsymbol{c} \in \mathbb{Q}^n \setminus Y_n$ where $n \in \mathbb{N}^*$ and Y_n is as in equation $(*)$ of Section 2. Suppose $S \subseteq \mathbb{Z}$ is such that every $s \in S$ can be represented in the form $s = \boldsymbol{c} \cdot \boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{Z}^n$, and for every $m \in \mathbb{N}$ there is $s \in S$ such that $[s-m, s+m] \cap S = \{s\}$. Then $r_+(\mathcal{R}_{\boldsymbol{c},S}) = n$.*

Proof. We aim to show that $\mathcal{R}_{\boldsymbol{c}} \in \widehat{\mathcal{R}_{\boldsymbol{c},S}}$, so we consider environments of the point $\mathcal{R}_{\boldsymbol{c}}$ in the topological space $(\mathcal{P}(\mathbb{Z}^n), \mathcal{T}_n)$. It is enough to consider basic open sets $U(\mathcal{R}_{\boldsymbol{c}}, C)$ where $C \subseteq \mathbb{Z}$ is finite. Choose $m \in \mathbb{N}$, $s \in S$ and $\boldsymbol{x}_0 \in \mathbb{Z}^n$ such that $m \geq \max\{\, |\boldsymbol{c} \cdot \boldsymbol{x}| \mid \boldsymbol{x} \in C^n \,\}$, $[s-m, s+m] \cap S = \{s\}$ and $s = \boldsymbol{c} \cdot \boldsymbol{x}_0$. For $\boldsymbol{x} \in C^n$, we have that $|\boldsymbol{c} \cdot (\boldsymbol{x} + \boldsymbol{x}_0) - s| = |\boldsymbol{c} \cdot \boldsymbol{x}| \leq m$, or equivalently, $\boldsymbol{c} \cdot (\boldsymbol{x} + \boldsymbol{x}_0) \in [s-m, s+m]$, so that $S \cap [s-m, s+m] = \{s\}$ implies

$$\begin{aligned} \boldsymbol{x} \in \mathcal{R}_{\boldsymbol{c}} &\iff \boldsymbol{c} \cdot (\boldsymbol{x} + \boldsymbol{x}_0) = s \iff \boldsymbol{c} \cdot (\boldsymbol{x} + \boldsymbol{x}_0) \in S \\ &\iff \boldsymbol{x} + \boldsymbol{x}_0 \in \mathcal{R}_{\boldsymbol{c},S} \iff \boldsymbol{x} \in \mathcal{R}_{\boldsymbol{c}} + (-\boldsymbol{x}_0). \end{aligned}$$

Hence, $\mathcal{R}_{\boldsymbol{c}} \cap C^n = (\mathcal{R}_{\boldsymbol{c},S} + (\boldsymbol{x}_0)) \cap C^n$. Consequently, $\mathcal{R}_{\boldsymbol{c}}$ is an accumulation point of the set $\{\, \mathcal{R}_{\boldsymbol{c},S} + \boldsymbol{y} \mid \boldsymbol{y} \in \mathbb{Z}^n \,\}$, i.e., $\mathcal{R}_{\boldsymbol{c}} \in \widehat{\mathcal{R}_{\boldsymbol{c},S}}$. Previous proposition and Theorem 2.6 now imply

$$n = r_+(\mathcal{R}_{\boldsymbol{c}}) \leq r_+(\mathcal{R}_{\boldsymbol{c},S}) \leq n. \quad \square$$

This result allows an easy generalization: The set S itself need not be sparse, it is enough that we can form a sparse set as a Boolean combination of translates of S. On the other hand, one has to pose some restrictions in order to make the result useful for the intended application. Lemma 4.1 dictates that the components of $\boldsymbol{c}$ have to be nonnegative, S a subset of $\mathbb{N}$, and, instead of $\mathcal{R}_{\boldsymbol{c},S}$, we need the result for the relation $(\mathcal{R}_{\boldsymbol{c},S} \cap \mathbb{N}^n) \setminus \{\boldsymbol{0}\}$. Note that whereas the preceding proposition holds for a singleton $S \subseteq \mathbb{Z}$, these new restrictions exclude finite S altogether. This is natural, since for finite $S \subseteq \mathbb{N}$, we have $\mathrm{FO}(\mathsf{C}_S) \equiv \mathrm{FO}$.

Lemma 4.6. *Let* $n \in \mathbb{N}^*$, $\boldsymbol{c} = (c_0, \ldots, c_{n-1}) \in \mathbb{N}^n \setminus Y_n$ *and let* $S \subseteq \mathbb{N}$ *be strongly non-periodic. Suppose also* $\gcd\{c_0, \ldots, c_{n-1}\} = 1$. *Then* $r_+((\mathcal{R}_{\boldsymbol{c},S} \cap \mathbb{N}^n) \setminus \{\boldsymbol{0}\}) = n$.

Proof. Denote $\mathcal{R}_{\boldsymbol{c},S} \cap \mathbb{N}^n$ by $\mathcal{R}$. (By Proposition 2.3.b, finite differences do not have any effect on r_+.) Pick a word $\xi\colon \{0, \ldots, r-1\} \to \{0,1\}$ occuring strongly non-periodically in $\chi_S\colon \mathbb{N} \to \{0,1\}$. We may fix $\boldsymbol{d} = (d_0, \ldots, d_{n-1}) \in \mathbb{Z}^n$ with $\boldsymbol{c}\cdot\boldsymbol{d} = 1$, as $\gcd\{c_0, \ldots, c_{n-1}\} = 1$. For $k \in \{0, \ldots, r-1\}$, put $R_k = \{\, \boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{x} + k\boldsymbol{d} \in \mathcal{R} \,\}$, if $\xi(k) = 1$, and $R_k = \{\, \boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{x} + k\boldsymbol{d} \notin \mathcal{R} \,\}$, if $\xi(k) = 0$. Put $R = \bigcap_{k=0}^{r-1} R_k$; then by Proposition 2.3 we have $r_+(R) \leq r_+(\mathcal{R})$.

The goal is to show that $\mathcal{R}_{\boldsymbol{c}} \in \widehat{R}$. Consider the environment $U(\mathcal{R}_{\boldsymbol{c}}, C)$ of $\mathcal{R}_{\boldsymbol{c}}$ where $C = [-t_0, t_0] \cap \mathbb{Z}$ for some $t_0 \in \mathbb{N}$.

It is routine to verify that for all $t \in \mathbb{N}$ there is $u \in \mathbb{N}$ such that every $m \in \mathbb{N}$ with $m \geq u$ can be written in form $m = \boldsymbol{c} \cdot \boldsymbol{x}$ where $\boldsymbol{x} = (x_0, \ldots, x_{n-1}) \in \mathbb{Z}^n$ and $x_i \geq t$, for each $i \in \{0, \ldots, n-1\}$. Let $\varrho = \max\{\, |\boldsymbol{c} \cdot \boldsymbol{x}| \mid \boldsymbol{x} \in C^n \,\}$ and $t = \lceil t_0 + r|\boldsymbol{d}| \rceil \in \mathbb{N}$. Choose $u \in \mathbb{N}$ as mentioned.

Since S is strongly non-periodic, there exists $m \in \mathbb{N}$ with $m \geq u$, $m \geq \varrho$ and $[m-\varrho, m+\varrho] \cap \mathbb{Z} \cap T_\xi^{\chi_S} = \{m\}$. Fix $\boldsymbol{x}_0$ such that $m = \boldsymbol{c} \cdot \boldsymbol{x}_0$ and $\boldsymbol{x}_0 \in (\mathbb{Z} \cap [t, \infty[)^n$. Let $\boldsymbol{x} \in C^n$ and $k \in \{0, \ldots, r-1\}$. Then $\boldsymbol{x}_0 + \boldsymbol{x} + k\boldsymbol{d} \in \mathbb{N}^n$, as $\boldsymbol{e}_i \cdot (\boldsymbol{x}_0 + \boldsymbol{x} + k\boldsymbol{d}) \geq t - t_0 - r|\boldsymbol{d}| \geq 0$, for each $i \in \{0, \ldots, n-1\}$, where $\boldsymbol{e}_i$ is the i-th vector in the canonical basis. So $\boldsymbol{x} + \boldsymbol{x}_0 + k\boldsymbol{d} \in \mathcal{R}$ iff $m + \boldsymbol{c} \cdot \boldsymbol{x} + k = \boldsymbol{c} \cdot (\boldsymbol{x} + \boldsymbol{x}_0 + k\boldsymbol{d}) \in S$ iff $\chi_S(m + \boldsymbol{c} \cdot \boldsymbol{x} + k) = 1$. Consequently, the following are equivalent:

$$\boldsymbol{x} + \boldsymbol{x}_0 \in R_k, \tag{1}$$
$$(\boldsymbol{x} + \boldsymbol{x}_0) + k\boldsymbol{d} \in \mathcal{R} \text{ iff } \xi(k) = 1, \tag{2}$$
$$\chi_S(m + \boldsymbol{c} \cdot \boldsymbol{x} + k) = \xi(k). \tag{3}$$

Now for $\boldsymbol{x} \in C^n$, we have that $\boldsymbol{x} \in R + (-\boldsymbol{x}_0)$ holds iff for every $k \in \{0, \ldots, r-1\}$, we have $\boldsymbol{x} + \boldsymbol{x}_0 \in R_k$. By the equivalence of cases (1) and (3), this holds iff $m + \boldsymbol{c} \cdot \boldsymbol{x} \in T_\xi^{\chi_S}$. As $[m-\varrho, m+\varrho] \cap \mathbb{Z} \cap T_\xi^{\chi_S} = \{m\}$ and $|\boldsymbol{c} \cdot \boldsymbol{x}| \leq \varrho$, we have $m + \boldsymbol{c}\cdot\boldsymbol{x} \in T_\xi^{\chi_S}$ iff $m + \boldsymbol{c}\cdot\boldsymbol{x} = m$ iff $\boldsymbol{c} \cdot \boldsymbol{x} = 0$ iff $\boldsymbol{x} \in \mathcal{R}_{\boldsymbol{c}}$. Hence, $\mathcal{R}_{\boldsymbol{c}} \cap C^n = (R + (-\boldsymbol{x}_0)) \cap C^n$, so $R + (-\boldsymbol{x}_0) \in U(\mathcal{R}_{\boldsymbol{c}}, C)$. Consequently, $\mathcal{R}_{\boldsymbol{c}} \in \widehat{R}$. Hence, by Proposition 4.4 and Theorem 2.6,

$$n = r_+(\mathcal{R}_{\boldsymbol{c}}) \leq r_+(R) \leq r_+(\mathcal{R}) \leq n. \quad \square$$

The latter combinatorial result quite easily implies the main theorem of this paper.

Theorem 4.7. *Let* $\mathcal{L} = \mathrm{FO}(\mathsf{C}_S^{(2)})$ *with* $S \subseteq \mathbb{N}$ *strongly non-periodic. Then* $\mathcal{L}$ *is unary hierarchical.*

Proof. Let $k \in \mathbb{N}^*$ and $n = 2^k$. Then there exists $\boldsymbol{c} = (c_0, \dots, c_{n-1}) \in \mathbb{N}^n \setminus Y_n$ such that $\gcd\{c_0, \dots, c_{n-1}\} = 1$ (cf. the proof of [L,Theorem 5.4]). Let $\varphi_{\boldsymbol{c},S} \in \mathrm{FO}(\mathsf{C}_S^{(2)})[\tau]$ be the sentence given by Lemma 4.1, and let Q be the quantifier defined by $\varphi_{\boldsymbol{c},S}$. Then

$$\mathcal{R}(Q) = \mathcal{R}_\tau(\varphi_{\boldsymbol{c},S}) = (\mathcal{R}_{\boldsymbol{c},S} \cap \mathbb{N}^n) \setminus \{\mathbf{0}\},$$

so by the preceding lemma, $\mathrm{udim}_\omega(Q) = r_+(\mathcal{R}(Q)) = n$. Since k was arbitrary, the supremum over all $\mathrm{udim}_\omega(Q)$ with Q unary and definable in $\mathrm{FO}(\mathsf{C}_S^{(2)})$ is ω, i.e., $\mathrm{FO}(\mathsf{C}_S^{(2)})$ is unary hierarchical. □

This result leaves some natural questions open: What happens when $S \subseteq \mathbb{N}$ is neither strongly non-periodic nor eventually periodic? Is it then possible that unary hierarchy of $\mathcal{L} = \mathrm{FO}(\mathsf{C}_S^{(2)})$ collapses, i.e., $\mathcal{L}$ is not unary hierarchical? In particular, it would be interesting to know how the vectorization hierarchy of $\mathrm{FO}(\mathsf{C}_S^{(n)})_{n\in\mathbb{N}}$ behaves when S is the Thue–Morse-set M or the Cantor set C. In general, how do the properties of $\mathrm{FO}(\{\, \mathsf{C}_S^{(n)} \mid n \in \mathbb{N}\,\})$ relate to the word combinatorial properties of S?

References

[1] Anuj Dawar. Generalized quantifiers and logical reducibilities. *J. Logic Comput.*, 5(2):213–226, 1995.

[2] H.-D. Ebbinghaus. Extended logics: the general framework. In *Model-theoretic logics*, Perspect. Math. Logic, pages 25–76. Springer, New York, 1985.

[3] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer. *Ramsey theory*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., New York, second edition, 1990. A Wiley-Interscience Publication.

[4] Yuri Gurevich. Logic and the challenge of computer science. In *Trends in theoretical computer science (Udine, 1984)*, volume 12 of *Principles Comput. Sci. Ser.*, pages 1–57. Computer Sci. Press, Rockville, MD, 1988.

[5] Lauri Hella. Definability hierarchies of generalized quantifiers. *Ann. Pure Appl. Logic*, 43(3):235–271, 1989.

[6] Lauri Hella and Michał Krynicki. Remarks on the Cartesian closure. *Z. Math. Logik Grundlag. Math.*, 37(6):539–545, 1991.

[7] Lauri Hella and Kerkko Luosto. Finite generation problem and n-ary quantifiers. In M. Krynicki, Mostowski M., and Szczerba L., editors, *Quantifiers: Logics, Models and Computation, Vol. I*, pages 63–104. Kluwer, 1995.

[8] Lauri Hella and Juha Nurmonen. Vectorization hierarchies of some graph quantifiers. *Arch. Math. Logic*, 39(3):183–207, 2000.

[9] Risto Kaila. On probabilistic elimination of generalized quantifiers. *Random Structures Algorithms*, 19(1):1–36, 2001.

[10] Phokion G. Kolaitis. Implicit definability on finite structures and unambiguous computations. In *Fifth Annual IEEE Symposium on Logic in Computer Science (Philadelphia, PA, 1990)*, pages 168–180. IEEE Comput. Soc. Press, Los Alamitos, CA, 1990.

[11] Phokion G. Kolaitis and Jouko A. Väänänen. Generalized quantifiers and pebble games on finite structures. *Ann. Pure Appl. Logic*, 74(1):23–75, 1995.

[12] Michał Krynicki. Notion of interpretation and nonelementary languages. *Z. Math. Logik Grundlag. Math.*, 34(6):541–552, 1988.

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

[14] M. Lothaire. *Combinatorics on words*, volume 17 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1983.

[15] M. Lothaire. *Algebraic combinatorics on words*, volume 90 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2002.

[16] Kerkko Luosto. Hierarchies of monadic generalized quantifiers. *J. Symbolic Logic*, 65(3):1241–1263, 2000.

[17] J. A. Makowsky, Saharon Shelah, and Jonathan Stavi. Δ-logics and generalized quantifiers. *Ann. Math. Logic*, 10(2):155–192, 1976.

[18] Jaroslav Nešetřil and Jouko A. Väänänen. Combinatorics and quantifiers. *Comment. Math. Univ. Carolin.*, 37(3):433–443, 1996.

[19] Jouko Väänänen. A hierarchy theorem for lindström quantifiers. In T. Wetterström M. Furberg and C. Åberg, editors, *Logic and Abstraction*, volume Acta Philosophica Gothoburgesia 1, pages 317–323. 1986.