On the Descriptional Complexity of Operations on Semilinear Sets

Simon Beier Markus Holzer Martin Kutrib

Institut für Informatik, Universität Giessen,
Arndtstr. 2, 35392 Giessen, Germany

{simon.beier,holzer,kutrib}@informatik.uni-giessen.de

We investigate the descriptional complexity of operations on semilinear sets. Roughly speaking, a semilinear set is the finite union of linear sets, which are built by constant and period vectors. The interesting parameters of a semilinear set are: (i) the maximal value that appears in the vectors of periods and constants and (ii) the number of such sets of periods and constants necessary to describe the semilinear set under consideration. More precisely, we prove upper bounds on the union, intersection, complementation, and inverse homomorphism. In particular, our result on the complementation upper bound answers an open problem from [G. J. LAVADO, G. PIGHIZZINI, S. SEKI: Operational State Complexity of Parikh Equivalence, 2014].

1 Introduction

A subset of $\mathbb{N}^k$, where $\mathbb{N}$ refers to the set of non-negative integers, of the form

$$L(C,P) = \left\{ \vec{c} + \sum_{\vec{x}_i \in P} \lambda_i \cdot \vec{x}_i \;\middle|\; \vec{c} \in C \text{ and } \lambda_i \in \mathbb{N} \right\},$$

for finite sets of periods and constants $P, C \subseteq \mathbb{N}^k$, is said to be *linear* if C is a singleton set. In this case we just write $L(\vec{c},P)$, where $\vec{c}$ is the constant vector. This can be seen as a straightforward generalization of an arithmetic progression allowing multiple differences. Moreover, a subset of $\mathbb{N}^k$ is said to be *semilinear* if it is a finite union of linear sets. Semilinear sets were extensively studied in the literature and have many applications in formal language and automata theory.

Let us recall two famous results from the very beginning of computer science, where semilinear sets play an important role. The Parikh image of a word $w \in \Sigma^*$ is defined as the function $\psi : \Sigma^* \to \mathbb{N}^{|\Sigma|}$ that maps w to a vector whose components are the numbers of occurrences of letters from Σ in w. Parikh's theorem states that the Parikh image of every context-free language L, that is, $\{\, \psi(w) \mid w \in L \,\}$, is a semilinear set [11]. A direct application of Parikh's theorem is that every context-free language is letter equivalent to a regular language. Another famous result on semilinear sets is their definability in Presburger arithmetic [6], that is, the first order theory of natural numbers with addition but without multiplication. Since Presburger arithmetic is decidable, corresponding questions on semilinear sets are decidable as well, because the conversion between semilinear set representations by vectors and Presburger formulas and *vice versa* is effectively computable.

Recently, semilinear sets appeared particularly in two different research directions from automata theory. The first research direction is that of jumping automata, a machine model for discontinuous information processing, recently introduced in [10]. Roughly speaking, a jumping finite automaton is an ordinary finite automaton, which is allowed to read letters from anywhere in the input string, not necessarily only from the left of the remaining input. Since a jumping finite automaton reads the input in a

E. Csuhaj-Varjú, P. Dömösi, Gy. Vaszil (Eds.): 15th International Conference on Automata and Formal Languages (AFL 2017)
EPTCS 252, 2017, pp. 41–55, doi:10.4204/EPTCS.252.8

discontinuous fashion, obviously, the order of the input letters does not matter. Thus, only the number of symbols in the input is important. In this way, the behavior of jumping automata is somehow related to the notions of Parikh image and Parikh equivalence. As already mentioned regular and context-free languages cannot be distinguished *via* Parikh equivalence, since both language families have semilinear Parikh images. This is in fact the starting point of the other recent research direction, the investigation of several classical results on automata conversions and operations subject to the notion of Parikh equivalence. For instance, in [8] it was shown that the cost of the conversion of an n-state nondeterministic finite automaton into a Parikh equivalent deterministic finite state device is of order $e^{\Theta(\sqrt{n \ln n})}$—this is in sharp contrast to the classical result on finite automata determinization which requires 2^n states in the worst case. A close inspection of these results reveals that there is a nice relation between Parikh images and Parikh equivalence of regular languages and jumping finite automata *via* semilinear sets. Thus one can read the above mentioned results as results on semilinear sets as well.

Here we investigate the descriptional complexity of the operation problem on semilinear sets. Recall that semilinear sets are closed under Boolean operations. The operands of the operations are semilinear sets of the form $\bigcup_{i\in I} L(\vec{c}_i, P_i) \subseteq \mathbb{N}^k$. Our resulting semilinear sets are of the form $S = \bigcup_{j\in J} L(C_j, Q_j) \subseteq \mathbb{N}^k$. We investigate upper bounds for the cardinality $|J|$ of the index set and for the norms $||Q_j||$ and $||C_j||$, these are the maximal values that appear in the vectors of periods Q_j and constants C_j. From this, one can automatically get upper bounds for the cardinalities of periods and constants through $|Q_j| \leq (||Q_j||+1)^k$ and $|C_j| \leq (||C_j||+1)^k$. One can also write the resulting set S in the form $S = \bigcup_{j\in J, \vec{c}\in C_j} L(\vec{c}, Q_j)$, which is a finite union of linear sets. In this form the index set has cardinality $\sum_{j\in J} |C_j|$. Upper bounds are proved for the Boolean operations and inverse homomorphism on semilinear sets. For instance, roughly speaking we show that intersection increases the size description polynomially, while complementation increases it double exponentially. A summary of our results can be found in Table 1. The precise bound of the former result improves a recent result shown in [9], and the latter result on the complementation answers an open question stated in [9], too.

Operation	Parameters of the resulting semilinear set $\bigcup_{j\in J} L(C_j, Q_j)$	
	$\lvert J\rvert$	$\max\{\lvert\lvert C_j\rvert\rvert, \lvert\lvert Q_j\rvert\rvert\}$
Union	$\lvert I_1\rvert + \lvert I_2\rvert$	ν
Intersection	$\lvert I_1\rvert \cdot \lvert I_2\rvert$	$O(m^2\nu^{k+2}+\nu)$
Complementation	$2^{(\nu+2)^{O(m)} \cdot \lvert I_1\rvert^{\log(3k+2)}}$	$2^{(\nu+2)^{O(m)} \cdot \lvert I_1\rvert^{\log(3k+2)}}$
Inverse Homom.	$\lvert I_1\rvert$	$O\left((\lvert\lvert H\rvert\rvert+1)^{\min(k_1,k)}(m+1)(\nu+1)^{k+1}\right)$

Table 1: Descriptional complexity results on the operation problem for semilinear subsets of $\mathbb{N}^k$. We assume k to be a constant in this table. The operands of the operations are semilinear sets of the form $\bigcup_{i\in I_\varepsilon} L(\vec{c}_i, P_i) \subseteq \mathbb{N}^k$, where $\varepsilon \in \{1,2\}$ for the first two operations and $\varepsilon = 1$ for the last two operations. The parameter ν is the maximal value that appears in the vectors of periods and constants in the operands. The parameter m is the maximal cardinality $|P_i|$ of all the period sets appearing in the operands. The inverse homomorphism is given by the matrix $H \in \mathbb{N}^{k\times k_1}$, where k_1 is also assumed to be a constant in this table. The parameter $||H||$ is the maximal value that appears in H.

It is worth mentioning that independently in [2] the operation problem for semilinear sets over the integers $\mathbb{Z}$ were studied. The obtained results there rely on involved decomposition techniques for semilinear sets. In contrast to that, our results are obtained by careful inspections of the original proofs on the closure properties. An application of the presented results on semilinear sets to the descriptional complexity of jumping automata and finite automata subject to Parikh equivalence is given in [1].

2 Preliminaries

Let $\mathbb{Z}$ be the set of integers and $\mathbb{N} = \{0,1,2,\dots\}$ be the set of non-negative integers. For the notion of semilinear sets we follow the notation of Ginsburg and Spanier [5]. For a natural number $k \geq 1$ and finite $C,P \subseteq \mathbb{N}^k$ let $L(C,P)$ denote the subset

$$L(C,P) = \left\{ \vec{c} + \sum_{\vec{x}_i \in P} \lambda_i \cdot \vec{x}_i \;\middle|\; \vec{c} \in C \text{ and } \lambda_i \in \mathbb{N} \right\}$$

of $\mathbb{N}^k$. Here the $\vec{c} \in C$ are called the *constants* and the $\vec{x}_i \in P$ the *periods*. If C is a singleton set we call $L(C,P)$ a *linear* subset of $\mathbb{N}^k$. In this case we simply write $L(\vec{c},P)$ instead of $L(\{\vec{c}\},P)$. A subset of $\mathbb{N}^k$ is said to be *semilinear* if it is a finite union of linear subsets. We further use $|P|$ to denote the size of a finite subset $P \subseteq \mathbb{N}^k$ and $||P||$ to refer to the value $\max\{\, ||\vec{x}|| \mid \vec{x} \in P \,\}$, where $||\vec{x}||$ is the maximum norm of $\vec{x}$, that is, $||(x_1,x_2,\dots,x_k)|| = \max\{\, |x_i| \mid 1 \leq i \leq k \,\}$. Observe, that

$$|P| \leq (||P||+1)^k.$$

Analogously we write $||A||$ for the maximum norm of a matrix A with entries in $\mathbb{Z}$, i.e. the maximum of the absolute values of all entries of A. The elements of $\mathbb{N}^k$ can be partially ordered by the $\leq$-relation on vectors. For vectors $\vec{x},\vec{y} \in \mathbb{N}^k$ we write $\vec{x} \leq \vec{y}$ if all components of $\vec{x}$ are less or equal to the corresponding components of $\vec{y}$. In this way we especially can speak of *minimal elements* of subsets of $\mathbb{N}^k$. In fact, due to [3] every subset of $\mathbb{N}^k$ has only a finite number of minimal elements.

Most results on the descriptional complexity of operations on semilinear sets is based on a size estimate of minimal solutions of matrix equations. We use a result due to [7, Theorem 2.6], which is based on [4], and can slightly be improved by a careful inspection of the original proof. The generalized result reads as follows:

Theorem 1 *Let $s,t \geq 1$ be integers, $A \in \mathbb{Z}^{s\times t}$ be a matrix of rank r, and $\vec{b} \in \mathbb{Z}^s$ be a vector. Moreover, let M be the maximum of the absolute values of the $r \times r$ sub-determinants of the extended matrix $(A \mid \vec{b})$, and $S \subseteq \mathbb{N}^t$ be the set of minimal elements of $\{\, \vec{x} \in \mathbb{N}^t \setminus \{\vec{0}\} \mid A\vec{x} = \vec{b} \,\}$. Then $||S|| \leq (t+1) \cdot M$.*

We will estimate the value of the above mentioned (sub)determinants with a corollary of Hadamard's inequality:

Theorem 2 *Let $r \geq 1$ be an integer, $A \in \mathbb{Z}^{r\times r}$ be a matrix, and m_i, for $1 \leq i \leq r$, be the maximum of the absolute values of the entries of the ith column of A. Then $|\det(A)| \leq r^{r/2} \prod_{i=1}^{r} m_i$.*

3 Operational complexity of semilinear sets

In this section we consider the descriptional complexity of operations on semilinear sets. We investigate the Boolean operations union, intersection, and complementation w.r.t. $\mathbb{N}^k$. Moreover, we also study the operation of inverse homomorphism on semilinear sets.

3.1 Union on semilinear sets

For the union of semilinear sets, the following result is straightforward.

Theorem 3 *Let $\bigcup_{i\in I} L(\vec{c}_i,P_i)$ and $\bigcup_{j\in J} L(\vec{c}_j,P_j)$ be semilinear subsets of $\mathbb{N}^k$, for some $k \geq 1$. Assume that I and J are disjoint finite index sets. Then the union*

$$\left(\bigcup_{i\in I} L(\vec{c}_i,P_i)\right) \cup \left(\bigcup_{j\in J} L(\vec{c}_j,P_j)\right) = \bigcup_{i\in I\cup J} L(\vec{c}_i,P_i)$$

can be described by a semilinear set with index sets size $|I|+|J|$, the maximal number of elements $m = \max_{i\in I\cup J}|P_i|$ in the period sets, and the entries in the constant vectors are at most $\ell = \max_{i\in I\cup J}||\vec{c}_i||$ and in the period vectors at most $n = \max_{i\in I\cup J}||P_i||$. □

Thus, the size increase for union on semilinear sets is only linear with respect to all parameters.

3.2 Intersection of semilinear sets

Next we consider the intersection operation on semilinear sets. The outline of the construction is as follows: we analyse the proof that semilinear sets are closed under intersection from [5, Theorem 6.1]. Due to distributivity it suffices to look at the intersection of linear sets. Those coefficients of the periods of our linear sets, which deliver a vector in the intersection, are described by systems of linear equations. For the intersection of the linear sets we get a semilinear set, where the periods and constants are built out of the minimal solutions of these systems of equations. We will estimate the size of the minimal solutions with the help of Theorems 1 and 2 in order to obtain upper bounds for the norms of the resulting periods and constants.

Theorem 4 *Let $\bigcup_{i\in I} L(\vec{c}_i,P_i)$ and $\bigcup_{j\in J} L(\vec{c}_j,P_j)$ be semilinear subsets of $\mathbb{N}^k$, for some $k \geq 1$. Assume that I and J are disjoint finite index sets. We set $n = \max_{i\in I\cup J}||P_i||$, $m = \max_{i\in I\cup J}|P_i|$, and $\ell = \max_{i\in I\cup J}||\vec{c}_i||$. Then for every $(i,j) \in I\times J$ there exist $P_{i,j}, C_{i,j} \subseteq \mathbb{N}^k$ with*

$$||P_{i,j}|| \leq 3m^2k^{k/2}n^{k+1},$$
$$||C_{i,j}|| \leq (3m^2k^{k/2}n^{k+1}+1)\ell,$$

and $\left(\bigcup_{i\in I} L(\vec{c}_i,P_i)\right) \cap \left(\bigcup_{j\in J} L(\vec{c}_j,P_j)\right) = \bigcup_{(i,j)\in I\times J} L(C_{i,j},P_{i,j})$.

Proof: We analyse the proof that semilinear sets are closed under intersection from [5, Theorem 6.1]. Let $i \in I$ and $j \in J$ be fixed and let $P_i = \{\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_p\}$, and $P_j = \{\vec{y}_1,\vec{y}_2,\ldots,\vec{y}_q\}$. Denote by X and Y the subsets of $\mathbb{N}^{p+q}$ defined by

$$X = \left\{ (\lambda_1,\ldots,\lambda_p,\mu_1,\ldots,\mu_q) \in \mathbb{N}^{p+q} \;\middle|\; \vec{c}_i + \sum_{r=1}^{p}\lambda_r\vec{x}_r = \vec{c}_j + \sum_{s=1}^{q}\mu_s\vec{y}_s \right\}$$

and

$$Y = \left\{ (\lambda_1,\ldots,\lambda_p,\mu_1,\ldots,\mu_q) \in \mathbb{N}^{p+q} \;\middle|\; \sum_{r=1}^{p}\lambda_r\vec{x}_r = \sum_{s=1}^{q}\mu_s\vec{y}_s \right\}.$$

Let C and P be the sets of minimal elements of X and $Y\setminus\{\vec{0}\}$. In the proof of [5, Theorem 6.1] it was shown that $X = L(C,P)$.

In order to estimate the size of $||C||$ and $||P||$ we use an alternative description of the vectors in X and Y in terms of matrix calculus. Let us define the matrix $H = (\vec{x}_1 \mid \vec{x}_2 \mid \cdots \vec{x}_p \mid -\vec{y}_1 \mid -\vec{y}_2 \mid \cdots \mid -\vec{y}_q)$ in $\mathbb{Z}^{k\times(p+q)}$. Then it is easy to see that

$$\vec{x} \in X \quad \text{if and only if} \quad H\vec{x} = \vec{c}_j - \vec{c}_i,$$

and

$$\vec{y} \in Y \quad \text{if and only if} \quad H\vec{y} = \vec{0}.$$

With $||P_i||, ||P_j|| \le n$, we derive from Theorem 2 that the maximum of the absolute values of any $r \times r$ sub-determinant, for $1 \le r \le k$, of the extended matrix $(H \mid \vec{0})$ is bounded from above by $k^{k/2}n^k$, because the maximum of the absolute values of the entries of the whole extended matrix $(H \mid \vec{0})$ is n. Then by Theorem 1 we conclude that

$$||P|| \le (p+q+1)k^{k/2}n^k \le 3mk^{k/2}n^k.$$

Analogously we can estimate the value of the maximum of the absolute values of any $r \times r$ sub-determinant, for $1 \le r \le k$, of the extended matrix $(H \mid \vec{c}_j - \vec{c}_i)$ by Theorem 2. It is bounded by $k^{k/2}n^k\ell$, because the maxima of the absolute values of the columns of $(H \mid \vec{c}_j - \vec{c}_i)$ are bounded by n and ℓ. Thus we have

$$||C|| \le (p+q+1)k^{k/2}n^k\ell \le 3mk^{k/2}n^k\ell$$

by Theorem 1.

Let $\tau : \mathbb{N}^{p+q} \to \mathbb{N}^k$ be the linear function given by $(\lambda_1, \ldots, \lambda_p, \mu_1, \ldots, \mu_q) \mapsto \sum_{r=1}^{p} \lambda_r \vec{x}_r$. Then we have $L(\vec{c}_i, P_i) \cap L(\vec{c}_j, P_j) = \vec{c}_i + \tau(X)$. The linearity of τ implies that $\tau(X)$ is equal to the semilinear set $L(\tau(C), \tau(P))$ (see, for example, [5]). So we get $L(\vec{c}_i, P_i) \cap L(\vec{c}_j, P_j) = L(\vec{c}_i + \tau(C), \tau(P))$. Because of $p \le m$ and $||P_i|| \le n$ we obtain

$$||\tau(P)|| \le m \cdot ||P|| \cdot n \le 3m^2k^{k/2}n^{k+1}$$

and

$$||\tau(C)|| \le m \cdot ||C|| \cdot n \le 3m^2k^{k/2}n^{k+1}\ell.$$

It follows that $||\vec{c}_i + \tau(C)|| \le \ell + ||\tau(C)|| = (3m^2k^{k/2}n^{k+1} + 1)\ell$.

Because $\left(\bigcup_{i\in I} L(\vec{c}_i, P_i)\right) \cap \left(\bigcup_{j\in J} L(\vec{c}_j, P_j)\right)$ is equal to the semilinear set $\bigcup_{(i,j)\in I\times J} L(\vec{c}_i, P_i) \cap L(\vec{c}_j, P_j)$ our theorem is proved. □

The index set of the semilinear set for the intersection has size $|I| \cdot |J|$ and the norms of the periods and constants are in $O(m^2\nu^{k+2} + \nu)$ if dimension k is constant. Here ν is the maximum of n and ℓ, which means that it is the maximum norm appearing in the two operands of the intersection. So the size increase for intersection is polynomial with respect to all parameters.

Now we turn to the intersection of more than two semilinear sets. The result is later utilized to explore the descriptional complexity of the complementation. First we have to deal with the intersection of two semilinear sets of the form $\bigcup_{i\in I} L(C_i, P_i)$ instead of $\bigcup_{i\in I} L(\vec{c}_i, P_i)$ as in the previous theorem. The following lemma is proved by writing a semilinear set of the form $L(C_i, P_i)$ as $\bigcup_{\vec{c}_i\in C_i} L(\vec{c}_i, P_i)$ and applying Theorem 4.

Lemma 5 *Let $\bigcup_{i\in I} L(C_i, P_i)$ and $\bigcup_{j\in J} L(C_j, P_j)$ be semilinear subsets of $\mathbb{N}^k$, for some $k \ge 1$. Assume that I and J are disjoint finite index sets. We set $p = \max\{|I|, |J|\}$, $n = \max_{i\in I\cup J} ||P_i||$, and $\ell = \max_{i\in I\cup J} ||C_i||$. Define $a_k = 4^{k+1}k^{k/2}$. Then there exists an index set H with*

$$|H| \le p^2(\ell+1)^{2k}$$

such that, for each $h \in H$, there are $P_h, C_h \subseteq \mathbb{N}^k$ with

$$||P_h|| \leq a_k n^{3k+1},$$
$$||C_h|| \leq (a_k n^{3k+1}+1)\ell,$$

and $\left(\bigcup_{i\in I} L(C_i,P_i)\right) \cap \left(\bigcup_{j\in J} L(C_j,P_j)\right) = \bigcup_{h\in H} L(C_h, P_h)$.

Proof: Let $i \in I$, $j \in J$, $\vec{c} \in C_i$, and $\vec{d} \in C_j$ be fixed. The proof of Theorem 4 shows that there exist $C_{i,j,\vec{c},\vec{d}}$, $P_{i,j,\vec{c},\vec{d}} \subseteq \mathbb{N}^k$ with

$$||P_{i,j,\vec{c},\vec{d}}|| \leq 3m^2 k^{k/2} n^{k+1},$$
$$||C_{i,j,\vec{c},\vec{d}}|| \leq (3m^2 k^{k/2} n^{k+1}+1)\ell,$$

and $L(\vec{c},P_i) \cap L(\vec{d},P_j) = L(C_{i,j,\vec{c},\vec{d}}, P_{i,j,\vec{c},\vec{d}})$, where m is the maximum of $|P_i|$ and $|P_j|$. Since $P_i, P_j \subseteq \mathbb{N}^k$, we have $m \leq (n+1)^k \leq (2n)^k$, for $n > 0$. This gives us

$$||P_{i,j,\vec{c},\vec{d}}|| \leq 3m^2 k^{k/2} n^{k+1} \leq 3(2n)^{2k} k^{k/2} n^{k+1} = 3\cdot 4^k k^{k/2} n^{3k+1} \leq a_k n^{3k+1}$$

and $||C_{i,j,\vec{c},\vec{d}}|| \leq (a_k n^{3k+1}+1)\ell$. With

$$\left(\bigcup_{i\in I} L(C_i,P_i)\right) \cap \left(\bigcup_{j\in J} L(C_j,P_j)\right) = \bigcup_{(i,j)\in I\times J} L(C_i,P_i)\cap L(C_j,P_j)$$

and

$$L(C_i,P_i)\cap L(C_j,P_j) = \left(\bigcup_{\vec{c}\in C_i} L(\vec{c},P_i)\right) \cap \left(\bigcup_{\vec{d}\in C_j} L(\vec{d},P_j)\right) = \bigcup_{(\vec{c},\vec{d})\in C_i\times C_j} L(\vec{c},P_i)\cap L(\vec{d},P_j)$$

our result is proven because of $|C_i \times C_j| \leq (\ell+1)^{2k}$. □

Now we present the result on the intersection of a finite number of semilinear sets.

Theorem 6 *Let $k \geq 1$ and $X \neq \emptyset$ be a finite index set. For every $x \in X$ let $\bigcup_{i\in I_x} L(C_i,P_i)$ be a semilinear subset of $\mathbb{N}^k$. Assume that I_x, I_y are disjoint finite index sets for $x, y \in X$ with $x \neq y$. We set $n = \max_{x\in X, i\in I_x} ||P_i||$ and $\ell = \max_{x\in X, i\in I_x} ||C_i||$. Define $p = \max_{x\in X} |I_x|$, $q = \lceil \log_2 |X| \rceil$, and $a_k = 4^{k+1}k^{k/2}$. Then there exists an index set J with*

$$|J| \leq p^{2^q}(\ell+1)^{k\cdot 2^{q+1}}(a_k n+1)^{4(3k+2)^{q+1}}, \tag{1}$$

such that, for each $j \in J$, there are $P_j, C_j \subseteq \mathbb{N}^k$ with

$$||P_j|| \leq (a_k n)^{(3k+1)^q},$$
$$||C_j|| \leq (a_k n+1)^{(3k+2)^q}\ell,$$

and $\bigcap_{x\in X}\left(\bigcup_{i\in I_x} L(C_i,P_i)\right) = \bigcup_{j\in J} L(C_j,P_j)$.

Proof: We prove this by induction on q. For $q = 0$ we have $|X| = 1$, so let X be the set $\{x\}$. Then we choose $J = I_x$ and get

$$|J| = p \leq p^1(\ell+1)^{2k}(a_k n+1)^{4(3k+2)^1} = p^{2^q}(\ell+1)^{k\cdot 2^{q+1}}(a_k n+1)^{4(3k+2)^{q+1}}$$

and

$$||P_j|| \leq n \leq (a_k n)^1 = (a_k n)^{(3k+1)^q},$$
$$||C_j|| \leq \ell \leq (a_k n+1)^1 \ell = (a_k n+1)^{(3k+2)^q}\ell$$

for every $j \in J = I_x$. This proves the statement for $q = 0$.

For $q = 1$ we have $|X| = 2$. In this case our statement follows directly from Lemma 5. Now let $q > 1$. We build pairs of the indices in X. This gives us $\lfloor |X|/2 \rfloor$ pairs of indices and an additional single index, if $|X|$ is odd. Due to Lemma 5 we get for each such pair (x,y) of indices an index set $H_{x,y}$ with

$$|H_{x,y}| \leq p^2(\ell+1)^{2k} \tag{2}$$

and for each $h \in H_{x,y}$ sets $C_h, P_h \subseteq \mathbb{N}^k$ with

$$||P_h|| \leq a_k n^{3k+1}, \qquad ||C_h|| \leq (a_k n^{3k+1}+1)\ell, \tag{3}$$

and

$$\left(\bigcup_{i\in I_x} L(C_i,P_i)\right) \cap \left(\bigcup_{j\in I_y} L(C_j,P_j)\right) = \bigcup_{h\in H_{x,y}} L(C_h,P_h).$$

So we have such a semilinear set for each of our pairs of indices and additionally a semilinear set for a single index out of X, if $|X|$ is odd. If we now intersect these $\lceil |X|/2 \rceil$ semilinear sets, we get $\bigcap_{x\in X}\left(\bigcup_{i\in I_x} L(C_i,P_i)\right)$. Because of $\lceil \log_2 \lceil |X|/2 \rceil \rceil = \lceil \log_2 |X| \rceil - 1 = q-1$, we can build this intersection by induction. This gives us an index set J and for each $j \in J$ sets $C_j, P_j \subseteq \mathbb{N}^k$ with

$$\bigcap_{x\in X}\left(\bigcup_{i\in I_x} L(C_i,P_i)\right) = \bigcup_{j\in J} L(C_j,P_j).$$

To get a bound for $|J|$ we use Inequality 1, where we replace q by $q-1$. Inequalities 2 and 3 give us bounds for p, ℓ, and n. So we have

$$\begin{aligned}|J| &\leq (p^2(\ell+1)^{2k})^{2^{q-1}}((a_k n^{3k+1}+1)\ell+1)^{k\cdot 2^q}(a_k^2 n^{3k+1}+1)^{4(3k+2)^q}\\ &\leq p^{2^q}(\ell+1)^{k\cdot 2^q}((a_k n^{3k+1}+1)(\ell+1))^{k\cdot 2^q}(a_k n+1)^{4(3k+1)(3k+2)^q}.\end{aligned}$$

By ordering the factors we get the upper bound

$$p^{2^q}(\ell+1)^{k(2^q+2^q)}(a_k n+1)^{(3k+1)k\cdot 2^q+4(3k+1)(3k+2)^q} = p^{2^q}(\ell+1)^{k\cdot 2^{q+1}}(a_k n+1)^{4((3k+1)k\cdot 2^{q-2}+(3k+1)(3k+2)^q)}.$$

Then

$$(3k+1)k\cdot 2^{q-2} + (3k+1)(3k+2)^q \leq (3k+2)^q + (3k+1)(3k+2)^q = (3k+2)^{q+1}$$

gives us

$$|J| \leq p^{2^q}(\ell+1)^{k\cdot 2^{q+1}}(a_k n+1)^{4(3k+2)^{q+1}}.$$

For each $j \in J$ we get
$$||P_j|| \le (a_k^2 n^{3k+1})^{(3k+1)^{q-1}} \le (a_k n)^{(3k+1)^q}$$
and
$$||C_j|| \le (a_k^2 n^{3k+1}+1)^{(3k+2)^{q-1}}(a_k n^{3k+1}+1)\ell \le (a_k n+1)^{(3k+1)(3k+2)^{q-1}+3k+1}\ell.$$
Because of $(3k+1)(3k+2)^{q-1}+3k+1 \le (3k+1)(3k+2)^{q-1}+(3k+2)^{q-1} = (3k+2)^q$ we finally obtain the bound $||C_j|| \le (a_k n+1)^{(3k+2)^q}\ell$. □

3.3 Complementation of semilinear sets

The next Boolean operation is the complementation. Our result is based on [5, Lemma 6.6, Lemma 6.8, and Lemma 6.9], which we slightly adapt. First we complement a linear set where the constant is the null-vector and the periods are linearly independent in Lemma 7. We continue by complementing a linear set with an arbitrary constant and linearly independent periods in Corollary 8. To complement a semilinear set where all the period sets are linearly independent in Theorem 9 we use DeMorgan's law: a semilinear set is a finite union of linear sets, so the complement is the intersection of the complements of the linear sets. For this intersection we use Theorem 6. Then we convert an arbitrary linear set to a semilinear set with linearly independent period sets in Lemma 10. Finally we insert the bounds from Lemma 10 into the bounds from Theorem 9 to complement an arbitrary semilinear set in Theorem 11.

Lemma 7 *Let $n,k \ge 1$, and $P \subseteq \mathbb{N}^k$ be linearly independent with $||P|| \le n$. Then there exists an index set I with $|I| \le 2^k+k-1$ such that, for each $i \in I$, there are subsets $P_i, C_i \subseteq \mathbb{N}^k$ with*
$$||P_i||, ||C_i|| \le (2k+1)k^{k/2}n^k$$
and $\mathbb{N}^k \setminus L(\vec{0},P) = \bigcup_{i\in I} L(C_i,P_i)$.

Proof: Let $P = \{\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_p\}$. Since the vectors in P are linearly independent, we conclude $p \le k$. For $i \in \{1,2,\ldots,k\}$ let $\vec{e}_i \in \mathbb{N}^k$ be the unit vector defined by $(\vec{e}_i)_i = 1$ and $(\vec{e}_i)_j = 0$ for $i \ne j$. By elementary vector space theory there exist $\vec{x}_{p+1},\vec{x}_{p+2},\ldots,\vec{x}_k \in \{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_k\}$ such that $\vec{x}_1,\vec{x}_2,\ldots,\vec{x}_k$ are linearly independent. Let Δ be the absolute value of the determinant of the matrix $(\vec{x}_1 \mid \vec{x}_2 \mid \cdots \mid \vec{x}_k)$. Moreover, let $\pi : \mathbb{N}^k \times \mathbb{N}^k \to \mathbb{N}^k$ be the projection on the first factor. For $J, K \subseteq \{1,2,\ldots,k\}$ we define
$$A_J = \{\, (\vec{y},\vec{a}) \in \mathbb{N}^k \times \mathbb{N}^k \mid a_j > 0 \text{ for all } j \in J \,\}$$
and
$$B_K = \left\{ (\vec{y},\vec{a}) \in \mathbb{N}^k \times \mathbb{N}^k \;\middle|\; \Delta\vec{y} + \sum_{i\in K} a_i\vec{x}_i = \sum_{i\in\{1,\ldots,k\}\setminus K} a_i\vec{x}_i \right\}.$$
Let Q_K and $D_{K,J}$ be the sets of minimal elements of $B_K \setminus \{\vec{0}\}$ and $B_K \cap A_J$. Looking at the proof of [5, Theorem 6.1] we see that $B_K \cap A_J = L(D_{K,J}, Q_K)$. The linearity of π implies
$$\pi(B_K \cap A_J) = L(\pi(D_{K,J}), \pi(Q_K)).$$
Next define
$$B'_K = \left\{ (\vec{y},\vec{a}) \in \mathbb{N}^k \times \mathbb{N}^k \;\middle|\; \vec{y} + \sum_{i\in K} a_i\vec{x}_i = \sum_{i\in\{1,\ldots,k\}\setminus K} a_i\vec{x}_i \text{ and } \vec{y} = \Delta\vec{z} \text{ for some } \vec{z} \in \mathbb{N}^k \right\}$$

and Q'_K and $D'_{K,J}$ to be the sets of minimal elements of $B'_K \setminus \{0\}$ and $B'_K \cap A_J$. Then the mapping $f : B'_K \to B_K$, defined *via* $(\vec{y},\vec{a}) \mapsto (\vec{y}/\Delta,\vec{a})$ is a bijection. The proof of [4] and Theorem 2 show that

$$||Q'_K||, ||D'_{K,J}|| \le (2k\Delta+1)k^{k/2}n^k \le \Delta \cdot (2k+1)k^{k/2}n^k.$$

With $Q_K = f(Q'_K)$ and $D_{K,J} = f(D'_{K,J})$ we get

$$||\pi(Q_K)||, ||\pi(D_{K,J})|| \le (2k+1)k^{k/2}n^k.$$

Set $G_1 = \bigcup_{\emptyset \neq K \subseteq \{1,\dots,k\}} \pi(B_K \cap A_K)$.

Because $\vec{x}_1,\vec{x}_2,\dots,\vec{x}_k$ are linearly independent every $\vec{y} \in \mathbb{N}^k$ can be written uniquely as $\vec{y} = \sum_{i=1}^{k} \lambda_{y,i}\vec{x}_i$ with $\lambda_{y,i} \in \mathbb{Q}$, for $i \in \{1,2,\dots,k\}$. Then $\vec{y} \in L(\vec{0},P)$ if and only if $\lambda_{y,i} \in \mathbb{N}$, for every $i \in \{1,2,\dots,p\}$ and $\lambda_{y,i} = 0$, for every $i \in \{p+1,p+2,\dots,k\}$. In the proof of [5, Lemma 6.7] it was shown that $\Delta\vec{y}$ can be written uniquely as $\Delta\vec{y} = \sum_{i=1}^{k} \mu_{y,i}\vec{x}_i$ with $\mu_{y,i} \in \mathbb{Z}$, for $i \in \{1,2,\dots,k\}$. Because of $\lambda_{y,i} = \mu_{y,i}/\Delta$ we get that $\vec{y} \in L(\vec{0},P)$ if and only if $\mu_{y,i}$ is a non-negative multiple of Δ, for every $i \in \{1,2,\dots,p\}$ and $\mu_{y,i} = 0$, for every $i \in \{p+1,p+2,\dots,k\}$. The set G_1 consists of all $\vec{y} \in \mathbb{N}^k$ such that at least one of the $\mu_{y,i}$ is negative. This implies $G_1 \subseteq \mathbb{N}^k \setminus L(\vec{0},P)$.

Now we set $G_2 = \bigcup_{i=p+1}^{k} \pi(B_\emptyset \cap A_{\{i\}})$. This set consists of all $\vec{y} \in \mathbb{N}^k$ such that all the $\mu_{y,i}$ are non-negative and there exists $i \in \{p+1,p+2,\dots,k\}$ such that $\mu_{y,i}$ is positive. This implies $G_2 \subseteq \mathbb{N}^k \setminus L(\vec{0},P)$.

For $i \in \{1,2,\dots,p\}$ and $r \in \{0,1,\dots,\Delta-1\}$ we set

$$E_{i,r} = \left\{ (\vec{y},\vec{a}) \in \mathbb{N}^k \times \mathbb{N}^p \;\middle|\; \Delta\vec{y} = \sum_{j=1}^{p} a_j\vec{x}_j \text{ and } a_i \bmod \Delta = r \right\}.$$

Let $R_{i,r}$ be the set of minimal elements of $E_{i,r} \setminus \{\vec{0}\}$. According to the proof of [5, Theorem 6.1] we get $E_{i,r} = L(R_{i,r}, R_{i,0})$, for $r > 0$. We set $\pi' : \mathbb{N}^k \times \mathbb{N}^p \to \mathbb{N}^k$ to be the projection on the first factor. Then $\pi'(E_{i,r}) = L(\pi'(R_{i,r}), \pi'(R_{i,0}))$, for $r > 0$, and

$$\bigcup_{r=1}^{\Delta-1} \pi'(E_{i,r}) = L(\bigcup_{r=1}^{\Delta-1} \pi'(R_{i,r}), \pi'(R_{i,0})).$$

Let $(\vec{y},\vec{a}) \in R_{i,r}$. Then we have $||\vec{a}|| \le \Delta$. This implies $||\vec{y}|| \le pn$. So we obtain $||\pi'(R_{i,r})|| \le pn$. Define $G_3 = \bigcup_{i=1}^{p} \bigcup_{r=1}^{\Delta-1} \pi'(E_{i,r})$. This is the set of all vectors $\vec{y} \in \mathbb{N}^k$ such that $\mu_{y,j} = 0$, for every $j \in \{p+1,p+2,\dots,k\}$, $\mu_{y,j} \ge 0$, for every $j \in \{1,2,\dots,p\}$, and $\mu_{y,j}$ is not divisible by Δ for at least one $j \in \{1,2,\dots,p\}$. Thus we have $G_1 \cup G_2 \cup G_3 = \mathbb{N}^k \setminus L(\vec{0},P)$. □

The next lemma gives a size estimation for the set $\mathbb{N}^k \setminus L(\vec{x}_0,P)$, for an arbitrary vector $\vec{x}_0$, instead of the null-vector, as in the previous theorem.

Lemma 8 *Let $k \ge 1$, subset $P \subseteq \mathbb{N}^k$ be linearly independent, and $\vec{x}_0 \in \mathbb{N}^k$. Then there exists an index set I with $|I| \le 2^k + 2k - 1$ such that, for each $i \in I$, there are $P_i, C_i \subseteq \mathbb{N}^k$ with*

$$\begin{aligned} ||P_i|| &\le (2k+1)k^{k/2}(||P||+1)^k, \\ ||C_i|| &\le (2k+1)k^{k/2}(||P||+1)^k + ||\vec{x}_0||, \\ |C_i| &\le \max(4^k k^{k^2/2+k}(||P||+1)^{k^2}, ||\vec{x}_0||), \end{aligned}$$

and $\mathbb{N}^k \setminus L(\vec{x}_0,P) = \bigcup_{i \in I} L(C_i,P_i)$.

Proof: For $j \in \{1,2,\ldots,k\}$ let

$$D_j = \{\,\vec{y} \in \mathbb{N}^k \mid y_\ell = 0 \text{ for } \ell \neq j \text{ and } y_j < (\vec{x}_0)_j\,\}$$

and $Q_j = \{\vec{e}_1,\ldots,\vec{e}_{j-1},\vec{e}_{j+1},\ldots,\vec{e}_k\}$, where the $\vec{e}_\ell$ are defined as in the proof of Lemma 7. Define the set $G = \bigcup_{j=1}^{k} L(D_j,Q_j)$. This is the set of all $\vec{y} \in \mathbb{N}^k$ such that $\vec{x}_0 \leq \vec{y}$ is false. So we have $G \subseteq \mathbb{N}^k \setminus L(\vec{x}_0,P)$.

Now let $Y = \{\vec{y} \in \mathbb{N}^k \mid \vec{x}_0 \leq \vec{y}\}$. Then $\mathbb{N}^k \setminus L(\vec{x}_0,P) = G \cup (Y \setminus L(\vec{x}_0,P))$. We have $Y \setminus L(\vec{x}_0,P) = (\mathbb{N}^k \setminus L(\vec{0},P)) + \vec{x}_0$. Due to Lemma 7 we have an index set J with $|J| \leq 2^k + k - 1$ and for each $j \in J$ subsets $C_j, P_j \subseteq \mathbb{N}^k$ with $||C_j||, ||P_j|| \leq (2k+1)k^{k/2}(||P||+1)^k$ such that $\mathbb{N}^k \setminus L(\vec{0},P) = \bigcup_{j\in J} L(C_j,P_j)$. This gives us $(\mathbb{N}^k \setminus L(\vec{0},P)) + \vec{x}_0 = \bigcup_{j\in J} L(C_j + \vec{x}_0, P_j)$. Because of $C_j \subseteq \mathbb{N}^k$ we obtain

$$|C_j + \vec{x}_0| = |C_j| \leq ((2k+1)k^{k/2}(||P||+1)^k+1)^k \leq (4k^{k/2+1}(||P||+1)^k)^k = 4^k k^{k^2/2+k}(||P||+1)^{k^2}.$$

This proves the stated claim. □

Now we are ready to deal with the complement of a semilinear set with linearly independent period sets.

Theorem 9 *Let $k \geq 1$ and $\bigcup_{i\in I} L(\vec{x}_i,P_i)$ be a semilinear subset of $\mathbb{N}^k$ with $I \neq \emptyset$ and linearly independent sets P_i. We set $n = \max_{i\in I} ||P_i||$ and $\ell = \max_{i\in I} ||\vec{x}_i||$. Define $q = \lceil \log_2 |I| \rceil$. Then there exists an index set J with*

$$|J| \leq (4k(n+1))^{5(k+2)(3k+2)^{q+1}}(\ell+1)^{k\cdot 2^{q+1}}$$

such that, for each $j \in J$, there are $P_j, C_j \subseteq \mathbb{N}^k$ with

$$||P_j|| \leq (4k(n+1))^{(k+2)(3k+1)^q},$$
$$||C_j|| \leq (4k(n+1))^{(k+2)(3k+2)^q+k}(\ell+1),$$

and $\mathbb{N}^k \setminus \bigcup_{i\in I} L(\vec{x}_i,P_i) = \bigcup_{j\in J} L(C_j,P_j)$.

Proof: Due to DeMorgan's law we have

$$\mathbb{N}^k \setminus \bigcup_{i\in I} L(\vec{x}_i,P_i) = \bigcap_{i\in I}\left(\mathbb{N}^k \setminus L(\vec{x}_i,P_i)\right).$$

Because of Lemma 8 for every $i \in I$ there exists an index set H_i with $|H_i| \leq 2^{k+1}$ such that, for each $h \in H_i$, there are $C_h, P_h \subseteq \mathbb{N}^k$ with

$$||P_h|| \leq (2k+1)k^{k/2}(n+1)^k \leq 3k^{k/2+1}(n+1)^k,$$
$$||C_h|| \leq (2k+1)k^{k/2}(n+1)^k + \ell \leq 3k^{k/2+1}(n+1)^k + \ell,$$

and $\mathbb{N}^k \setminus L(\vec{x}_i,P_i) = \bigcup_{h\in H_i} L(C_h,P_h)$. Theorem 6 gives us an index set J and for each $j \in J$ sets $C_j, P_j \subseteq \mathbb{N}^k$ with

$$\bigcup_{j\in J} L(C_j,P_j) = \bigcap_{i\in I}\left(\bigcup_{h\in H_i} L(C_h,P_h)\right) = \mathbb{N}^k \setminus \bigcup_{i\in I} L(\vec{x}_i,P_i)$$

and

$$|J| \leq (2^{k+1})^{2^q}(3k^{k/2+1}(n+1)^k+\ell+1)^{k\cdot 2^{q+1}}(4^{k+1}k^{k/2}\cdot 3k^{k/2+1}(n+1)^k+1)^{4(3k+2)^{q+1}}$$
$$\leq 2^{(k+1)\cdot 2^q}(4k^{k/2+1}(n+1)^k(\ell+1))^{k\cdot 2^{q+1}}(4^{k+2}k^{k+1}(n+1)^k)^{4(3k+2)^{q+1}}.$$

Now we order the factors and get that this is less than or equal to

$$2^{(k+1)\cdot 2^q+2k\cdot 2^{q+1}+8(k+2)(3k+2)^{q+1}}(k(n+1))^{(k+1)k\cdot 2^{q+1}+4(k+1)(3k+2)^{q+1}}(\ell+1)^{k\cdot 2^{q+1}}.$$

Because of $(5k+1)\le(k+2)(3k+2)$ we have

$$(k+1)\cdot 2^q+2k\cdot 2^{q+1}+8(k+2)(3k+2)^{q+1}=(5k+1)\cdot 2^q+8(k+2)(3k+2)^{q+1}\le 9(k+2)(3k+2)^{q+1}.$$

Furthermore $k\cdot 2^{q+1}\le(3k+2)^{q+1}$ gives us

$$(k+1)k\cdot 2^{q+1}+4(k+1)(3k+2)^{q+1}\le 5(k+1)(3k+2)^{q+1}.$$

So we get

$$|J|\le 2^{9(k+2)(3k+2)^{q+1}}(k(n+1))^{5(k+1)(3k+2)^{q+1}}(\ell+1)^{k\cdot 2^{q+1}}\le(4k(n+1))^{5(k+2)(3k+2)^{q+1}}(\ell+1)^{k\cdot 2^{q+1}}.$$

For each $j\in J$ we have

$$||P_j||\le(4^{k+1}k^{k/2}\cdot 3k^{k/2+1}(n+1)^k)^{(3k+1)^q}\le(4k(n+1))^{(k+2)(3k+1)^q}$$

and

$$\begin{aligned}||C_j||&\le(4^{k+1}k^{k/2}\cdot 3k^{k/2+1}(n+1)^k+1)^{(3k+2)^q}(3k^{k/2+1}(n+1)^k+\ell)\\&\le(4^{k+2}k^{k+1}(n+1)^k)^{(3k+2)^q}(4k^{k/2+1}(n+1)^k(\ell+1)).\end{aligned}$$

From $k^{(k+1)(3k+2)^q}k^{k/2+1}=k^{(k+1)(3k+2)^q+k/2+1}\le k^{(k+1)(3k+2)^q+k+(3k+2)^q}=k^{(k+2)(3k+2)^q+k}$ we finally deduce $||C_j||\le(4k(n+1))^{(k+2)(3k+2)^q+k}(\ell+1)$. □

Next we convert an arbitrary linear set to a semilinear set with linearly independent period sets. The idea is the following: If the periods are linearly dependent we can rewrite our linear set as a semilinear set, where in each period set one of the original periods is removed. By doing this inductively the period sets get smaller and smaller until they are finally linearly independent.

Lemma 10 *Let $L(\vec{x}_0,P)$ be a linear subset of $\mathbb{N}^k$ for some $k\ge 1$. We set $m=|P|$ and $n=||P||$. Then there exists an index set I with*

$$|I|\le(m+1)!\cdot m!/2^m\cdot(k^{k/2}n^k+1)^{m-1}$$

and, for each $i\in I$, a linearly independent subset $P_i\subseteq\mathbb{N}^k$ with $||P_i||\le n$ and a vector $\vec{x}_i\in\mathbb{N}^k$ with

$$||\vec{x}_i||\le||\vec{x}_0||+(m+1)(m+2)/2\cdot k^{k/2}n^{k+1}$$

such that $\bigcup_{i\in I}L(\vec{x}_i,P_i)=L(\vec{x}_0,P)$.

Proof: We prove this by induction on m. The statement of the lemma is clearly true for $m=0$ or $m=1$. So let $m\ge 2$ now. If P is linearly independent the statement of the lemma is trivial. Thus we assume P to be linearly dependent. Then there exists $p\in\{1,2,\ldots,\lfloor m/2\rfloor\}$ and pairwise different vectors $x_1,x_2,\ldots,x_p,y_1,y_2,\ldots,y_{m-p}\in P$ such that $X=\left\{\vec{a}\in\mathbb{N}^m\setminus\{\vec{0}\}\;\middle|\;H\cdot\vec{a}=\vec{0}\right\}$, where $H\in\mathbb{Z}^{k\times m}$ is the matrix $(x_1|x_2|\ldots|x_p|-y_1|-y_2|\ldots|-y_{m-p})$, is not empty. Let $\vec{a}$ be a minimal element of X. From Theorem 1 and Theorem 2 we deduce $||\vec{a}||\le(m+1)k^{k/2}n^k$. For $j\in\{1,2,\ldots,p\}$ let $C_j=$

$\{\vec{x}_0 + \lambda\vec{x}_j \mid \lambda \in \{0,1,\ldots,a_j-1\}\}$, if $a_j > 0$, and $C_j = \{\vec{x}_0\}$, otherwise. Furthermore let $Q_j = P \setminus \{\vec{x}_j\}$. In the proof of [5, Lemma 6.6] it was shown that $\bigcup_{j=1}^{p} L(C_j, Q_j) = L(\vec{x}_0, P)$. We can rewrite this set as $\bigcup_{j\in\{1,2,\ldots,p\},\vec{c}\in C_j} L(\vec{c}, Q_j)$. Here the size of the index set is smaller than or equal to $m/2 \cdot ||\vec{a}|| \leq (m+1)m/2 \cdot k^{k/2}n^k$ and for each such $\vec{c}$ we have $||\vec{c}|| \leq ||\vec{x}_0|| + ||\vec{a}|| \cdot n \leq ||\vec{x}_0|| + (m+1)k^{k/2}n^{k+1}$. Because of $|Q_j| = m-1$ for each $j \in \{1,2,\ldots,p\}$ and $\vec{c} \in C_j$, by induction, there exists an index set $I_{j,\vec{c}}$ with

$$|I_{j,\vec{c}}| \leq m! \cdot (m-1)!/2^{(m-1)} \cdot (k^{k/2}n^k + 1)^{m-2}$$

and, for each $i \in I_{j,\vec{c}}$, a linearly independent subset $R_i \subseteq \mathbb{N}^k$ with $||R_i|| \leq n$ and a vector $\vec{z}_i \in \mathbb{N}^k$ with

$$||\vec{z}_i|| \leq ||\vec{c}|| + m(m+1)/2 \cdot k^{k/2}n^{k+1}$$

such that $\bigcup_{i\in I_{j,\vec{c}}} L(\vec{z}_i, R_i) = L(\vec{c}, Q_j)$. This gives us

$$\bigcup_{j\in\{1,2,\ldots,p\},\vec{c}\in C_j, i\in I_{j,\vec{c}}} L(\vec{z}_i, R_i) = L(\vec{x}_0, P).$$

The size of this index set is smaller than or equal to

$$(m+1)m/2 \cdot k^{k/2}n^k \cdot m! \cdot (m-1)!/2^{(m-1)} \cdot (k^{k/2}n^k+1)^{m-2} \leq (m+1)! \cdot m!/2^m \cdot (k^{k/2}n^k+1)^{m-1}.$$

With

$$||\vec{z}_i|| \leq ||\vec{x}_0|| + (m+1)k^{k/2}n^{k+1} + m(m+1)/2 \cdot k^{k/2}n^{k+1} \leq ||\vec{x}_0|| + (m+1)(m+2)/2 \cdot k^{k/2}n^{k+1}$$

the lemma is proved. □

With Theorem 9 and Lemma 10 we are ready to complement an arbitrary semilinear set.

Theorem 11 *Let $k \geq 1$ and $\bigcup_{i\in I} L(\vec{x}_i, P_i)$ be a semilinear subset of $\mathbb{N}^k$ with $I \neq \emptyset$. We set $n = \max_{i\in I} ||P_i||$, $m = \max_{i\in I} |P_i|$, and $\ell = \max_{i\in I} ||\vec{x}_i||$. Define $b(k,n,m,I)$ as*

$$\left(\sqrt{k}(n+2)\right)^{k\cdot\log_2(3k+2)\cdot(3m+1)+3} \cdot (3k+2)^{-(2\log_2(e)+1)m+7} \cdot |I|^{\log_2(3k+2)}.$$

Then there exists an index set J with

$$|J| \leq 2^{b(k,n,m,I)} \cdot (\ell+2)^{\left(\sqrt{k}(n+2)\right)^{k\cdot(3m+1)+8} \cdot \left(2e^2\right)^{-m} \cdot |I|}$$

such that, for each $j \in J$, there are $P_j, C_j \subseteq \mathbb{N}^k$ with

$$||P_j|| \leq 2^{b(k,n,m,I)},$$
$$||C_j|| \leq 2^{b(k,n,m,I)} \cdot (\ell+1),$$

and $\mathbb{N}^k \setminus \bigcup_{i\in I} L(\vec{x}_i, P_i) = \bigcup_{j\in J} L(C_j, P_j)$.

Proof: Because of Lemma 10 there exists an index set $H \neq \emptyset$ with

$$|H| \leq (m+1)! \cdot m!/2^m \cdot (k^{k/2}n^k+1)^{m-1} \cdot |I|$$

and, for each $h \in H$, a linearly independent subset $P_h \subseteq \mathbb{N}^k$ with $||P_h|| \leq n$ and a vector $\vec{x}_h \in \mathbb{N}^k$ with

$$||\vec{x}_h|| \leq \ell + (m+1)(m+2)/2 \cdot k^{k/2}n^{k+1}$$

such that $\bigcup_{h\in H} L(\vec{x}_h, P_h) = \bigcup_{i\in I} L(\vec{x}_i, P_i)$. With Stirling's formula we get

$$(m+1)! \le (m+1)^{m+3/2}e^{-m} \qquad \text{and} \qquad m! \le (m+1)^{m+1/2}e^{-m+1}.$$

This gives us $(m+1)!\cdot m! \le (m+1)^{2m+2}e^{-2m+1} \le ((n+1)^k+1)^{2m+2}e^{-2m+1}$ and we get

$$\begin{aligned}|H| &\le (m+1)!\cdot m!/2^m \cdot (k^{k/2}n^k+1)^{m-1}\cdot |I| \\ &\le ((n+1)^k+1)^{2m+2}e^{-2m+1}/2^m \cdot (k^{k/2}n^k+1)^{m-1}\cdot |I| \\ &\le (k^{k/2}(n+2)^k)^{2m+2}e^{-2m+1}/2^m \cdot (k^{k/2}(n+2)^k)^{m-1}\cdot |I| \\ &= e\cdot\left(\sqrt{k}(n+2)\right)^{k\cdot(3m+1)}\cdot\left(2e^2\right)^{-m}\cdot |I|.\end{aligned}$$

We shall use Theorem 9 to get upper bounds for the complement. So we set $q = \lceil \log_2 |H| \rceil$. In all three bounds of Theorem 9 the exponent of $4k(n+1)$ is bounded from above by $(3k+2)^{q+3}$. We have

$$\begin{aligned}(3k+2)^{q+3} &\le (3k+2)^{\log_2|H|+4} \\ &= (3k+2)^4\cdot |H|^{\log_2(3k+2)} \\ &\le (3k+2)^4\cdot\left(e\cdot\left(\sqrt{k}(n+2)\right)^{k\cdot(3m+1)}\cdot\left(2e^2\right)^{-m}\cdot |I|\right)^{\log_2(3k+2)} \\ &= \left(\sqrt{k}(n+2)\right)^{k\cdot\log_2(3k+2)\cdot(3m+1)}\cdot(3k+2)^{-(2\log_2(e)+1)m+\log_2(e)+4}\cdot |I|^{\log_2(3k+2)} \\ &\le \left(\sqrt{k}(n+2)\right)^{-3}\cdot(3k+2)^{-1}\cdot b(k,n,m,I).\end{aligned}$$

Because of $\log_2(4k(n+1)) \le \left(\sqrt{k}(n+2)\right)^2$ we get

$$(4k(n+1))^{(3k+2)^{q+3}} \le 2^{\left(\sqrt{k}(n+2)\right)^{-1}\cdot(3k+2)^{-1}\cdot b(k,n,m,I)}. \tag{4}$$

For the sets P_j from Theorem 9 this implies $||P_j|| \le 2^{b(k,n,m,I)}$. For each $h\in H$ we have

$$\begin{aligned}||\vec{x}_h||+1 &\le \ell + (m+1)(m+2)/2\cdot k^{k/2}n^{k+1}+1 \\ &\le \ell + k^{k/2}(n+2)^k(n+3)^k n^{k+1}/2+1 \\ &\le \ell + \left(\sqrt{k}n(n+2)(n+3)\right)^{k+1}/2+1 \\ &\le \ell + \left(\sqrt{k}(n+2)^3\right)^{k+1} \\ &= \ell + 2^{(k+1)\cdot\log_2\left(\sqrt{k}(n+2)^3\right)} \\ &\le 2^{(k+1)\cdot\log_2\left(\sqrt{k}(n+2)^3\right)}\cdot(\ell+1).\end{aligned}$$

From $(3k+2)^{q+3} \le \left(\sqrt{k}(n+2)\right)^{-3}\cdot(3k+2)^{-1}\cdot b(k,n,m,I)$ we get

$$(k+1)\cdot\log_2\left(\sqrt{k}(n+2)^3\right) \le \left(\sqrt{k}(n+2)\right)^3 \le (3k+2)^{-1}\cdot b(k,n,m,I).$$

This leads to $||\vec{x}_h||+1 \le 2^{(3k+2)^{-1}\cdot b(k,n,m,I)}\cdot(\ell+1)$. Together with Inequality (4) this implies

$$||C_j|| \le 2^{b(k,n,m,I)}\cdot(\ell+1)$$

for the sets C_j from Theorem 9, because of $3k+2 \geq 2$. For each $h \in H$ we have

$$\begin{aligned}(||\vec{x}_h||+1)^{k\cdot 2^{q+1}} &\leq \left(\ell + 2^{(k+1)\cdot\log_2\left(\sqrt{k}(n+2)^3\right)}\right)^{4k\cdot|H|} \\ &\leq (\ell+2)^{(k+1)\cdot\log_2\left(\sqrt{k}(n+2)^3\right)\cdot 4k\cdot e\cdot\left(\sqrt{k}(n+2)\right)^{k\cdot(3m+1)}\cdot\left(2e^2\right)^{-m}\cdot|I|} \\ &\leq (\ell+2)^{\left(\sqrt{k}(n+2)\right)^{k\cdot(3m+1)+8}\cdot\left(2e^2\right)^{-m}\cdot|I|}\end{aligned}$$

because

$$4e\cdot k(k+1)\cdot\log_2\left(\sqrt{k}(n+2)^3\right) \leq 12\cdot 2k^2\cdot 3\cdot\log_2\left(\sqrt{k}(n+2)\right) \leq 72k^2\cdot\sqrt{k}\cdot(n+2) \leq 2^7\cdot\left(\sqrt{k}\right)^5(n+2).$$

With Inequality (4) we get $|J| \leq 2^{b(k,n,m,I)}\cdot(\ell+2)^{\left(\sqrt{k}(n+2)\right)^{k\cdot(3m+1)+8}\cdot\left(2e^2\right)^{-m}\cdot|I|}$ for the set J from Theorem 9. This proves our theorem. □

The size of the resulting index set and the norms for the resulting periods and constants are bounded from above by $2^{(\nu+2)^{O(m)}\cdot|I|^{\log(3k+2)}}$, if k is constant and, as before, ν is the maximum of n and ℓ. So we observe that the size increase is exponential in ν and $|I|$ and double exponential in m.

3.4 Inverse homomorphism on semilinear sets

Finally, we consider the descriptional complexity of the inverse homomorphism. We follow the lines of the proof on the inverse homomorphism closure given in [5]. Since inverse homomorphism commutes with union, we only need to look at linear sets. The vectors in the pre-image of a linear set, with respect to a homomorphism, can be described by a system of linear equations. Now we use the same techniques as in the proof of Theorem 4: out of the minimal solutions of the system of equations we can build periods and constants of a semilinear description of the pre-image. With Theorems 1 and 2 we estimate the size of the minimal solutions to get upper bounds for the norms of the resulting periods and constants.

Theorem 12 *Let $k_1, k_2 \geq 1$ and $\bigcup_{i\in I} L(\vec{c}_i, P_i)$ be a semilinear subset of $\mathbb{N}^{k_2}$. We set $n = \max_{i\in I}||P_i||$, $m = \max_{i\in I}|P_i|$, and $\ell = \max_{i\in I}||\vec{c}_i||$. Moreover let $H \in \mathbb{N}^{k_2\times k_1}$ be a matrix and $h: \mathbb{N}^{k_1} \to \mathbb{N}^{k_2}$ be the corresponding linear function $\vec{x} \mapsto H\vec{x}$. Then for every $i \in I$ there exist $Q_i, C_i \subseteq \mathbb{N}^{k_1}$ with*

$$\begin{aligned}||Q_i|| &\leq (k_1+m+1)k_2^{\min(k_1+m,k_2)/2}\cdot(||H||+1)^{\min(k_1,k_2)}(n+1)^{\min(m,k_2)}, \\ ||C_i|| &\leq (k_1+m+1)k_2^{\min(k_1+m,k_2)/2}\cdot(||H||+1)^{\min(k_1,k_2)}(n+1)^{\min(m,k_2)}\ell,\end{aligned}$$

and $h^{-1}\left(\bigcup_{i\in I} L(\vec{c}_i, P_i)\right) = \bigcup_{i\in I} L(C_i, Q_i)$.

Proof: Let $i \in I$ be fixed and define P_i to be $\{\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_p\}$. Then the set of vectors

$$\left\{\vec{x} \in \mathbb{N}^{k_1} \;\middle|\; H\vec{x} \in L(\vec{c}_i, P_i)\right\}$$

is equal to $\left\{\vec{x} \in \mathbb{N}^{k_1} \;\middle|\; \exists \lambda_1, \lambda_2, \ldots, \lambda_p \in \mathbb{N} : H\vec{x} = \vec{c}_i + \lambda_1\vec{y}_1 + \lambda_2\vec{y}_2 + \cdots + \lambda_p\vec{y}_p\right\}$.

Now let $\tau: \mathbb{N}^{k_1}\times\mathbb{N}^p \to \mathbb{N}^{k_1}$ be the projection on the first component and let $J \in \mathbb{Z}^{k_2\times(k_1+p)}$ be the matrix $J = (H \mid -\vec{y}_1 \mid -\vec{y}_2 \mid \cdots \mid -\vec{y}_p)$. We obtain

$$\left\{\vec{x} \in \mathbb{N}^{k_1} \;\middle|\; H\vec{x} \in L(\vec{c}_i, P_i)\right\} = \tau\left(\left\{\vec{x} \in \mathbb{N}^{k_1+p} \;\middle|\; J\vec{x} = \vec{c}_i\right\}\right).$$

Let $C \subseteq \mathbb{N}^{k_1+p}$ be the set of minimal elements of $\left\{\vec{x} \in \mathbb{N}^{k_1+p} \;\middle|\; J\vec{x} = \vec{c}_i\right\}$ and $Q \subseteq \mathbb{N}^{k_1+p}$ be the set of minimal elements of $\left\{\vec{x} \in \mathbb{N}^{k_1+p} \setminus \{\vec{0}\} \;\middle|\; J\vec{x} = \vec{0}\right\}$. In the proof of [5, Theorem 6.1] it is shown that $L(C,Q) = \left\{\vec{x} \in \mathbb{N}^{k_1+p} \;\middle|\; J\vec{x} = \vec{c}_i\right\}$. With $p \leq m$ and $||P_i|| \leq n$, we derive from Theorems 1 and 2 that

$$||Q|| \leq (k_1+m+1)k_2^{\min(k_1+m,k_2)/2} \cdot (||H||+1)^{\min(k_1,k_2)}(n+1)^{\min(m,k_2)}$$

With $||\vec{c}_i|| \leq \ell$ we get

$$||C|| \leq (k_1+m+1)k_2^{\min(k_1+m,k_2)/2} \cdot (||H||+1)^{\min(k_1,k_2)}(n+1)^{\min(m,k_2)}\ell.$$

Since τ is linear we have $L(\tau(C),\tau(Q)) = \left\{\vec{x} \in \mathbb{N}^{k_1} \;\middle|\; H\vec{x} \in L(\vec{c}_i,P_i)\right\}$. Moreover, we have the inequalities $||\tau(Q)|| \leq ||Q||$ and $||\tau(C)|| \leq ||C||$. Because of $h^{-1}\left(\bigcup_{i\in I} L(\vec{c}_i,P_i)\right) = \bigcup_{i\in I} h^{-1}\left(L(\vec{c}_i,P_i)\right)$ our theorem is proved. □

We see that the index set of the semilinear set is not changed under inverse homomorphism. If k_1 and k_2 are constant, then the norms of the periods and constants of the resulting semilinear set are in $O\left((||H||+1)^{\min(k_1,k_2)}(m+1)(\nu+1)^{k_2+1}\right)$. Again ν is the maximum of n and ℓ. Thus, the size increase for inverse homomorphism is polynomial with respect to all parameters.

References

[1] S. Beier, M. Holzer & M. Kutrib (2017): *Operational State Complexity and Decidability of Jumping Finite Automata*. In: Proceedings DLT, LNCS, 10396, Springer, pp. 96–108, doi:10.1007/978-3-319-62809-7_6.

[2] D. Chistikov & Ch. Haase (2016): *The Taming of the Semi-Linear Set*. In: Proceedings ICALP, LIPIcs 55, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, pp. 128:1–128:13, doi:10.4230/LIPIcs.ICALP.2016.128.

[3] L. E. Dickson (1913): *Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors*. American J. Math. 35(4), pp. 413–422, doi:10.2307/2370405 .

[4] J. von zur Gathen & M. Sieveking (1978): *A bound on solutions of linear integer equalities and inequalities*. Proc. Am. Math. Soc. 72, pp. 155–158, doi:10.1090/S0002-9939-1978-0500555-0.

[5] S. Ginsburg & E. H. Spanier (1964): *Bounded ALGOL-like languages*. Trans. AMS 113, pp. 333–368, doi:10.2307/1994067.

[6] S. Ginsburg & E. H. Spanier (1966): *Finite-Turn Pushdown Automata*. SIAM J. Comput. 4, pp. 429–453, doi:10.1137/0304034.

[7] T.-D. Huynh (1980): *The complexity of semilinear sets*. In: Proceedings ICALP, LNCS 85, Springer, pp. 324–337, doi:10.1007/3-540-10003-2_81.

[8] G. J. Lavado, G. Pighizzini & S. Seki (2013): *Converting nondeterministic automata and context-free grammars into Parikh equivalent one-way and two-way deterministic automata*. Inform. Comput. 228–229, pp. 1–15, doi:10.1016/j.ic.2013.06.003.

[9] G. J. Lavado, G. Pighizzini & S. Seki (2014): *Operational State Complexity under Parikh Equivalence*. In: Proceedings DCFS, LNCS 8614, Springer, pp. 294–305, doi: 10.1007/978-3-319-09704-6_26.

[10] A. Meduna & P. Zemek (2012): *Jumping Finite Automata*. Internat. J. Found. Comput. Sci. 23, pp. 1555–1578, doi:10.1142/S0129054112500244.

[11] R. J. Parikh (1966): *On Context-Free Languages*. J. ACM 13, pp. 570–581, doi:10.1145/321356.321364.