Logical theory of the additive monoid of subsets of natural integers

Christian Choffrut* Serge Grigorieff*

May 5, 2014

Abstract

We consider the logical theory of the monoid of subsets of $\mathbb{N}$ endowed solely with addition lifted to sets: no other set theoretical predicate or function, no constant (contrarily to previous work by Jeż and Okhotin cited below). We prove that the class of true Σ_5 formulas is undecidable and that the whole theory is recursively isomorphic to second-order arithmetic. Also, each ultimately periodic set A (viewed as a predicate $X = A$) is Π_4 definable and their collection is Σ_6. Though these undecidability results are not surprising, they involve technical difficulties witnessed by the following facts: 1) no elementary predicate or operation on sets (inclusion, union, intersection, complementation, adjunction of 0) is definable, 2) The class of subsemigroups is not definable though that of submonoids is easily definable. To get our results, we code integers by a Π_3 definable class of submonoids and arithmetic operations on $\mathbb{N}$ by Δ_5 operations on this class.

Contents

1 Introduction

2 Preliminaries

2.1 Submonoids

2.2 Maximal submonoids

2.3 Basic predicates

2.3.1 Removing definable constants

2.3.2 Basic constants

2.3.3 Some classes of subsets

2.3.4 Least element in a set

2.3.5 Singleton sets

*LIAFA, CNRS and Université Paris 7 Denis Diderot, France.

3 **Using submonoids to approximate and emulate**
3.1 Approximating inclusion and membership
3.1.1 Approximating inclusion
3.1.2 Approximating membership
3.2 Definability issues of special submonoids
3.2.1 Definability of each special submonoid
3.2.2 Definability of the class of special submonoids
3.3 Addition and multiplication on special submonoids
3.3.1 Insight into the proof
3.3.2 Addition on special submonoids
3.3.3 Multiplication on special submonoids
4 Complexity of the theory
4.1 Emulating second-order arithmetic
4.2 The theory of addition on sets is undecidable
5 Non definable predicates
5.1 What is so special about the predicate $0 \in X$?
5.2 Other non definable predicates
6 Structural definability in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$
6.1 Families of sets all containing 0
6.2 Families of sets invariant by translation
6.3 Definable sets of integers
6.4 Definability with an extra predicate
7 Remarkable definable subsets and classes of subsets
7.1 Operations on sets with close minimum elements
7.2 Fixed submonoids
7.3 Regular subsets of $\mathbb{N}$
7.3.1 Fixed regular subsets of $\mathbb{N}$
7.3.2 Finite and cofinite subsets of $\mathbb{N}$
7.3.3 The class of regular subsets of $\mathbb{N}$
8 Conclusion

1 Introduction

{s:introduction}

The object of this paper could hardly be more elementary since we are concerned with the class of subsets of nonnegative integers equipped with addition as unique operation. Presburger studied in 1929 the first-order logic of integers with addition and showed that this logic admits quantifier elimination on the language enriched with the order relation and all arithmetic congruences. Consequently, the theory is decidable and it was proved in 1974 [3] by Michael J. Fischer and Michael O. Rabin that its time complexity is upper bounded by a double exponential. In the sixties, the class of relations defined by the logic received a simple algebraic characterization by Seymour Ginsburg as the semilinear subsets of integers, [5]. It can thus be reasonably said that this logic is well-understood.

Our purpose is still a first-order theory but the domain is the power set of $\mathbb{N}$ with set addition and equality, formally $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$. At the beginning of our investigation we came up every day with different properties. Some could be considered as the source of inspiration for exercises in an introductory course in logic or as entertaining mathematical recreation. Others played a more important role and are kept in this paper, but all had a low complexity in the arithmetical hierarchy, i.e., they were expressible with very few quantifier alternations. However, we could not build on them to get a consistent and general view of the problem. E.g., we were not even able to answer the question whether or not the property for a subset to be recognizable by a finite automaton is expressible. The final picture to the contrary is that the expressiveness of the theory is extremely powerful but, at least the way we did it, this was obtained by working out predicates of higher complexity.

It should not be surprising that the submonoids of $\mathbb{N}$ play a crucial role since a nonempty subset X is a submonoid if and only if it satisfies the condition $X + X = X$. Submonoids of $\mathbb{N}$ have a deceiving simplicity. Contrarily to submonoids of nonunary free monoids, they are finitely generated. They are related to an intriguing and well-celebrated problem attributed to Frobenius which asks the following. Say a submonoid is *numerical* if it is generated by a finite subset of integers with greatest common divisor equal to 1. These submonoids are known to be cofinite in $\mathbb{N}$ but what precisely is the largest integer not in the submonoid? There is a rich literature on the topic and many conferences are dedicated to the classification of these monoids [4, 9, 2, 11], nonetheless the problem seems to be far from solved and we could not find any result that would help us in our investigation.

We now turn to a quick presentation of our work. Section 2 gathers all basic material of algebraic or logical type used in the sequel and is essentially meant for the reader unfamiliar with the domain. Section 3 establishes the main properties of our paper which can be summarized as saying that under certain restrictions which cannot be relaxed, membership of an element to a subset and subset inclusion can be expressed via special (and simple) submonoids. Based on these results, Section 4 shows that the theory is highly undecidable by interpreting the second-order theory of arithmetic in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$. Actually, the Σ_5 fragment is already undecidable, see Theorem 4.2. At this point of the article the only useful subsets all contain the integer 0. Section 5 investigates to what extent other classes of subsets are expressible. We show in particular that we cannot express the fact that a subset is obtained from another by just adding 0. The same is true of the simplest set theoretical predicates (inclusion, union, intersection, complementation) and of the class of subsemigroups of $\mathbb{N}$. Nevertheless, this leaves the problem of trying to extend the classes of subsets expressible in the logic which is done in section 6. In particular, we show that a singleton class $\{A\}$ is definable with set addition if and only if it is definable in second-order arithmetic. An inventory of typical predicates expressible in the theory to be found in Section 7 serves an illustrative purpose among which the class of regular subsets of $\mathbb{N}$ which is Σ_7 definable.

It is worthwhile mentioning the works of J̇ez and Okhotin since they can be interpreted as studying the Diophantine theory of the current structure enriched with all ultimately periodic subsets of $\mathbb{N}$ (as set constants). In [6] they show that there exists an encoding of the subsets of $\mathbb{N}$ under which each recursive subset of $\mathbb{N}$ is the encoding of the unique solution of some system of equations involving the operation of sum of subsets and the regular subsets as unique constants. Furthermore they prove the satisfiability of this theory to be Π^0_1-complete. Because ultimately periodic constants are Π_4 definable, their undecidable result is in accordance with ours and leaves the open question of finding the minimum undecidable fragment.

2 Preliminaries

{s:preliminaries}

In this section we recall classical and introduce elementary properties of two types: algebraic and logic.

2.1 Submonoids

{ss:basic-on-submonoi

Given a non negative integer n and two subsets $X, Y \subseteq \mathbb{N}$ we define

$$nX = \{nx \mid x \in X\} \tag{1}$$ {eq:nX}

$$X + Y = \{x + y \mid x \in X, y \in Y\} \tag{2}$$ {eq:XplusY}

Observe that $2X \neq X + X$.

{def:star}

Definition 2.1. A subset $X \subseteq \mathbb{N}$ is a *subsemigroup* if it is closed under addition, i.e., $X + X \subseteq X$. A subsemigroup is a *submonoid* if it is nonempty and contains 0. Equivalently a submonoid is a nonempty subset satisfying the condition $X + X = X$.

The *submonoid generated* by Y, denoted Y^*, is the minimum submonoid containing Y, i.e. containing every finite sum of elements of Y

$$Y^* = \{0\} \cup \bigcup_{n \geq 1} \overbrace{Y + \cdots + Y}^{n \text{ times}}$$

The subset Y is a *generating subset* of Y^*.

We are concerned with the first-order theory of $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$. It is not surprising that the submonoids of $\mathbb{N}$ play a central role since a set X containing 0 is a submonoid if and only if $X = X + X$ holds. Most of the following is folklore and does not contain anything new. For the sake of completeness we recall it with some detail. We start with two trivial observations the proofs of which are omitted.

{p:semigroups-submono

Proposition 2.2. *A set $X \subseteq \mathbb{N}$ such that $0 \in X$ is a monoid if and only if $X \setminus \{0\}$ is a subsemigroup.*

{p:sum-of-submonoids}

Proposition 2.3. *If X and Y are two submonoids, so is $X + Y$.*

A remarkable property of $\mathbb{N}$ is that its submonoids are finitely generated. This can be stated more precisely.

{p:N-submonoids}

Proposition 2.4. *1. Every submonoid X of $\mathbb{N}$ is finitely generated and has a minimum generating set $G(X)$ which is equal to*

$$G(X) = S \setminus (S + S) \quad \textit{where } S = X \setminus \{0\} \tag{3}$$ {eq:minimum-generatin

The set $G(X)$ is called the minimum generator *or the* minimum generating set *of X and its elements are called the* generators *of X.*

2. A submonoid X is of the form $\{0\}$ or of the form

$$X = b\big(F \cup (a + \mathbb{N})\big) \tag{4}$$ {eq:general-form-subm

where $b \geq 1$ and $0 \in F \subseteq \{0, \ldots, a-1\}$.

Observe that the numeric submonoids (i.e. those generated by a finite subset of $\mathbb{N}$ with greatest common divisor equal to 1, cf. [4, 11]) correspond to $b = 1$. They are exactly the submonoids which are cofinite.

Proof. 1. Let b be the greatest common divisor of the elements in X. Then $X \subseteq b\mathbb{N}$. Let $a_1, \ldots, a_n \in X$ such that g.c.d.$(a_1, \ldots, a_n) = b$. By Bézout there exist $x_1, \ldots, x_\ell, y_1, \ldots, y_{n-\ell} \in \mathbb{N}$ such that

$$x_1a_1 + \cdots + x_\ell a_\ell - y_1a_{\ell+1} - \cdots - y_{n-\ell}a_n = b$$

(up to a permutation of $a_1, \ldots, a_n$). Let $L = x_1a_1 + \cdots + x_\ell a_\ell$ and $R = y_1a_{\ell+1} + \ldots + y_{n-\ell}a_n$ and observe that $L, R \subseteq b\mathbb{N}$. For all integers k, set $k = qa_1 + r$ where $0 \leq r < a_1$ and $q \in \mathbb{N}$. Then

$$(a_1 - 1)R + kb \;=\; qba_1 + (a_1 - 1 - r)R + rL \;\in\; a_1\mathbb{N} + \cdots + a_n\mathbb{N}\,.$$

which shows that all multiples of b greater than $(a_1 - 1)R$ are generated by $a_1, \ldots, a_n$ hence that X is generated by the finite class consisting of all its elements less than or equal to

$$\max\{a_1, \ldots, a_n, (a_1 - 1)R\} \tag{5}$$ {eq:threshold}

It is clear that $G(X)$ as in equation (3) generates X. Assume by contradiction that there exists a generating set H not containing $G(X)$ and let α be an element in $G(X) \setminus H$. We assume without loss of generality that $0 \notin H$.

Since $X = H^*$ we have $\alpha = \beta + \gamma$ where $\beta \in H$ and $\gamma \in H^* \setminus \{0\}$. Then $\alpha \in S + S$ which contradicts the definition of $G(X)$.

2. Observe that the integer in (5) is a multiple of b, say ab. Then in order to obtain (4) it suffices to consider the subset F satisfying $bF = X \cap b\{0, \ldots, a-1\}$ □

The specific form of nonzero submonoids leads us to the following general notion.

{de:ultimate period}

Definition 2.5. A set $X \subseteq \mathbb{N}$ has *ultimate period* b if there exists an integer K such that for all $n \geq K$ we have

$$n \in X \iff n + b \in X$$

If this holds for some $b \geq 1$, X is *ultimately periodic.*

The following result is classical.

{p:regular}

Proposition 2.6. *Let* $X \subseteq \mathbb{N}$.
1. The following conditions are equivalent:
- (i) X *is ultimately periodic,*
- (ii) X *is a regular subset of* $\mathbb{N}$.
- (iii) $X = A \cup (B + b\mathbb{N})$ *where* $a, b \in \mathbb{N}$, $\emptyset \neq A \subseteq [0, a[$ *and* $B \subseteq [a, a+p[$.

The set X *is finite if and only if* $B = \emptyset$ *in* (iii).

2. If X *has ultimate period* b *then all multiples of* b *are ultimate periods.*

3. When X *is a submonoid, the integer* b *of equation (4) is its minimum ultimate period.*

{rk:algo GX}

Remark 2.7. Equation (3) of Proposition 2.4 gives a simple algorithm to compute $G(X) = \{g_0, \ldots, g_n\}$ provided its ultimate period b is known. Suppose the submonoid X has minimum nonzero element m. Let $g_0 = m$ and let g_{i+1} be the minimum element of X not in $\sum_{j=0}^{j=i} g_j\mathbb{N}$. Halt when $\{g_0, \ldots, g_n\}$ generate m/b successive elements of the periodic tail of X.

The above representation of submonoids allows us to express the inclusion and intersection of submonoids simply.

{p:pair submonoids}

Proposition 2.8. *Any pair of nonzero submonoids* X, Y *is of the form*

$$\left\{ \begin{array}{rcl} X & = & bF \cup (ad + b\mathbb{N}) \\ Y & = & cG \cup (ad + c\mathbb{N}) \end{array} \right. \quad \textit{with} \quad \left\{ \begin{array}{l} b \geq 1,\ c \geq 1,\ d = \text{l.c.m.}(b, c) \\ 0 \in F \cap G, \quad F, G \textit{ finite} \\ bF \cup cG \subseteq \{0, \ldots, ad - 1\} \end{array} \right.$$

The intersection monoid is $X \cap Y = (bF \cap cG) \cup (ad + d\mathbb{N})$.
In particular, $X \subseteq Y$ *if and only if* $bF \subseteq cG$ *and* c *divides* b.

2.2 Maximal submonoids

{sect:maximal}{not:maximalsubmono}

Notation 2.9. We write $X \lhd Y$ whenever X is an inclusion-maximal proper submonoid of the submonoid Y.

{p:maximal}

Proposition 2.10. *Let X be a submonoid of $\mathbb{N}$ with $G(X)$ as minimum generating set.*

1. The proper maximal submonoids of X are the sets $X \setminus \{g\}$ where $g \in G(X)$.

2. Every generator of X distinct from g is a generator of $X \setminus \{g\}$ (but there may be other ones, cf. Example 2.11). In other words,

$$G(X) \setminus \{g\} \ \subseteq \ G(X \setminus \{g\}) \tag{6}$$ {eq:GX-g}

Proof. Let $g \in G(X)$. All elements in $X \setminus \{g\}$ are of the form

$$\sum_{h \in G} x_h h \quad \text{with} \quad x_g = 0 \quad \text{or} \quad \sum_{h \in G} x_h \geq 2$$

These elements clearly define a proper submonoid of X and this monoid is maximal. Conversely, consider a proper submonoid X' of X. There exists some $g \in G(X) \setminus X'$ hence $X' \subseteq X \setminus \{g\}$ and equality holds in case X' is maximal.

Finally, for $g \in G(X)$, the following inclusion is straightforward:

$$\big(S \setminus (S+S)\big) \setminus \{g\} \ \subseteq \ (S \setminus \{g\}) \setminus \big((S \setminus \{g\}) + (S \setminus \{g\})\big) \ .$$

Thus, every X-generator distinct from g is an $(X \setminus \{g\})$-generator. $\square$

{ex:maximal}

Example 2.11. *The subset $X = \{0\} \cup \{3,5,6\} \cup 8 + \mathbb{N}$ is the submonoid with minimum generating set $\{3,5\}$ (use Remark 2.7). It thus has two proper maximal submonoids $X_1 \lhd X$ and $X_2 \lhd X$. The minimum generating set of $X_1 = X \setminus \{3\}$ is $\{5,6,8,9\}$ and the minimum generating set of $X_2 = X \setminus \{5\}$ is $\{3,8,10\}$. Thus X_1 has 4 proper maximal submonoids and X_2 has 3 proper maximal submonoids.*

Consequently, every submonoid which is not reduced to 0 has finitely many proper maximal submonoids but some monoids fail to have minimal proper supermonoids. The following result characterizes them.

{p:no minimal super}

Proposition 2.12. *The submonoids which have no minimal supermonoid are $\{0\}$ and the sets $b\mathbb{N}$, $b \geq 0$.*

Proof. Suppose Y is a minimal nonzero supermonoid of X and let $G(Y)$ be its minimum generating set. Then $X = Y \setminus \{g\}$ for some $g \in G(Y)$ and $G(Y) \setminus \{g\} \subseteq G(X)$. First, using Proposition 2.4, we show that $\{0\}$

and the sets $b\mathbb{N}$, $b \geq 1$ have no minimal supermonoid. We argue by way of contradiction.

Case $X = \{0\}$. Then $G(X) = \emptyset$ hence $G(Y) = \{g\}$ so that $Y = g\mathbb{N}$. Now, $2g\mathbb{N}$ is a submonoid strictly between $\{0\}$ and Y, contradicting the minimality of Y over X.

Case $G(X) = \{b\}$, *i.e.* $X = b\mathbb{N}$. For $b = 1$ it is trivial so we assume $b > 1$. Then $G(Y)$ has at most two generators. It cannot have only one generator because $G(X)$ would be empty. Thus, $G(Y) = \{b, c\}$ for some $c \notin b\mathbb{N}$. Then by Propostion 2.3, $b\mathbb{N} + (b+1)c\mathbb{N}$ is a submonoid strictly between $X = b\mathbb{N}$ and $Y = b\mathbb{N} + c\mathbb{N}$, contradicting the minimality of Y over X.

We now show that every submonoid X distinct from $\{0\}$ and the sets $b\mathbb{N}$, $b \geq 1$, has a minimal supermonoid. This condition on X, together with Equation (4) supra, insure that $X = b(F \cup (a+\mathbb{N}))$ with $b \geq 1$, $0 \in F$, $a \geq 2$ and $a - 1 \notin F$. The set $Y = \{0\} \cup b((a-1)+\mathbb{N})$ is clearly a submonoid. By Proposition 2.3, $X+Y$ is a supermonoid of X. A simple computation shows that $(X+Y) \setminus X = \{b(a-1)\}$, i.e., that $X+Y$ is a minimal supermonoid of X. □

2.3 Basic predicates

{ss:basic-on-logic}

In further sections we try to evaluate the complexity of the predicates which are expressible in the logic. Here we content ourselves with gathering the most elementary predicates.

We recall that a predicate is Σ_n (resp. Π_n) if it is defined by a formula that begins with some existential (resp. universal) quantifiers and alternates $n-1$ times between series of existential and universal quantifiers. It is Δ_n if it is both Σ_n and Π_n. It is $\Sigma_n \wedge \Pi_n$ if it is equivalent to a conjunction of a Σ_n and a Π_n formulas. We assume the reader has some familiarity with computing the logical complexity. As an example of the type of computation consider the $\Sigma_1 \wedge \Pi_1$ formula

$$\theta(x,y,z,t,u) \quad \equiv \quad \exists t \ \phi(x,y,t) \ \wedge \ \forall u \ \psi(x,z,u)$$

Assume that x is the sole common free variable of ϕ and ψ. Then

$$\begin{aligned} \theta(x,y,z,t,u) \quad &\Longleftrightarrow \quad \exists t \ \forall u \ \big(\phi(x,y,t) \wedge \psi(x,z,u)\big) \\ &\Longleftrightarrow \quad \forall u \ \exists t \ \big(\phi(x,y,t) \wedge \psi(x,z,u)\big) \\ \exists x \ \theta(x,y,z,t,u) \quad &\Longleftrightarrow \quad \exists x \ \exists t \ \forall u \ \big(\phi(x,y,t) \wedge \psi(x,z,u)\big) \end{aligned}$$

showing both that the predicate associated to θ is Δ_2 and that $\exists x\, \theta$ is Σ_2. Such a type of computation will not be explicitly carried out in the sequel.

2.3.1 Removing definable constants

{sss:removing definab

Since we deal with definability in a particular structure, the following classical result in logic will be heavily used.

{p:remove constants}

Proposition 2.13. *Let $\mathcal{M}$ be any logical structure and a an element of $\mathcal{M}$. Let $n, p \in \mathbb{N}$. Suppose a is Δ_n definable in $\mathcal{M}$, i.e. $x = a$ is equivalent to a Σ_n formula and to a Π_n formula. Then, for every Σ_p (resp. Π_p) formula $\phi(x, \vec{y})$ (with free variables x and possibly some other ones), there exists a $\Sigma_{\max(n,p)}$ (resp. $\Pi_{\max(n,p)}$) formula $\psi(\vec{y})$ such that $\phi(a, \vec{y})$ is equivalent to $\psi(\vec{y})$.*

Proof. Suppose $x = a$ is equivalent to formulas

$$\exists \vec{v_1}\ \forall \vec{v_2} \ldots Q_n \vec{v_n}\ F(\vec{v_1}, \ldots, \vec{v_n}, x) \quad , \quad \forall \vec{v_1}\ \exists \vec{v_2} \ldots R_n \vec{v_n}\ G(\vec{v_1}, \ldots, \vec{v_n}, x)$$

where Q_n (resp. R_n) is $\forall$ if n is even (resp. odd) and is $\exists$ otherwise.
Letting $s = \max(n, p)$, if ϕ is $\exists \vec{z_1}\ \forall \vec{z_2} \ldots Q_p \vec{z_p}\ A(\vec{z_1}, \ldots, \vec{z_p}, x, \vec{y})$ then

$$\phi(a, \vec{y}) \iff \exists \vec{z_1}\ \exists \vec{v_1}\ \forall \vec{z_2}\ \forall \vec{v_2} \ldots Q_s \vec{z_s}\ Q_s \vec{v_s} \\ \big(F(\vec{v_1}, \ldots, \vec{v_n}, x)\ \wedge\ A(\vec{z_1}, \ldots, \vec{z_p}, x, \vec{y})\big)$$

and if ϕ is $\forall \vec{z_1}\ \exists \vec{z_2} \ldots R_p \vec{z_p}\ B(\vec{z_1}, \ldots, \vec{z_p}, x, \vec{y})$ then

$$\phi(a, \vec{y}) \iff \forall \vec{z_1}\ \forall \vec{v_1}\ \exists \vec{z_2}\ \exists \vec{v_2} \ldots R_s \vec{z_s}\ R_s \vec{v_s} \\ \big(G(\vec{v_1}, \ldots, \vec{v_n}, x)\ \Rightarrow\ B(\vec{z_1}, \ldots, \vec{z_p}, x, \vec{y})\big)$$

□

2.3.2 Basic constants

{sss:basic constants}{p:0-N}

Proposition 2.14. *1. The predicate $X = \emptyset$ is Π_1.*
2. The predicate $X = \{0\}$ is Π_1.
3. The predicate $0 \in X$ is Σ_1.
4. The predicate $X = \mathbb{N}$ is $\Sigma_1 \wedge \Pi_1$.

Proof. 1. $X = \emptyset$ if and only if $\forall Y\ X + Y = X$.
2. $\{0\}$ is the neutral element of $\mathcal{P}(\mathbb{N})$ hence is the unique set satisfying $\forall Y\ X + Y = Y$.
3. $0 \in X$ if and only if $\exists Y\ (Y \neq \emptyset \wedge X + Y = Y)$.
4. $X = \mathbb{N}$ if and only if $0 \in X\ \wedge\ \forall Y\ (0 \in Y \Rightarrow X + Y = X)$. □

2.3.3 Some classes of subsets

{sss:class-submonoids

The following two classes of subsets are easily definable.

{p:submonoid Sigma1}

Proposition 2.15. *1. The class of submonoids of $\mathbb{N}$ is Σ_1 definable.*
2. The class of final segments $\{n + \mathbb{N} \mid n \in \mathbb{N}\}$ is $\Sigma_1 \wedge \Pi_1$ definable.

Proof. 1. X is a submonoid if and only if it is nonempty and $X + X = X$.
2. $X \in \{n + \mathbb{N} \mid n \in \mathbb{N}\}$ if and only if $X \neq \emptyset \wedge \forall Y\ (0 \in Y \Rightarrow X + Y = X)$. □

2.3.4 Minimum element in a set

{sss:minimum}

The minimum element of a nonempty set X is denoted by $\min X$.

{p:min}

Proposition 2.16. *The following predicates are definable by formulas of the stated complexity:*

$$\begin{array}{lllllll}
1. & \text{(a)} & \min X & \le & k & \textit{is} & \Pi_2 \quad \textit{for } k \ge 1 \\
 & \text{(b)} & \min X & = & 0 & \textit{is} & \Sigma_1 \\
 & \text{(c)} & \min X & = & 1 & \textit{is} & \Pi_2 \\
 & \text{(d)} & \min X & = & k & \textit{is} & \Sigma_2 \wedge \Pi_2 \quad \textit{for } k \ge 2 \\
2. & \text{(a)} & \min X & \le & \min Y & \textit{is} & \Sigma_1 \\
 & \text{(b)} & \min X & = & \min Y & \textit{is} & \Sigma_1 \\
 & \text{(c)} & \min X & \le & \min Y + k & \textit{is} & \Pi_2 \quad \textit{for } k \ge 1 \\
 & \text{(d)} & \min X & = & \min Y + 1 & \textit{is} & \Pi_2 \\
 & \text{(e)} & \min X & = & \min Y + k & \textit{is} & \Sigma_2 \wedge \Pi_2 \quad \textit{for } k \ge 2 \\
3. & \text{(a)} & \min X + \min Y & \le & \min Z & \textit{is} & \Sigma_1 \\
 & \text{(b)} & \min X + \min Y & = & \min Z & \textit{is} & \Sigma_1
\end{array}$$

Proof. 1a. $\min X \le k$ if and only if $X \ne \emptyset$ and X is not the sum of $k+1$ sets which do not contain 0 :

$$X \ne \emptyset \ \wedge \ \forall X_1, \dots, X_{k+1} \ (X = X_1 + \dots + X_{k+1} \Rightarrow \bigvee_{i=1}^{i=k+1} 0 \in X_i)$$

Claim 3 of Proposition 2.14 yields the stated logical complexity.
1b. Again use claim 3 of Proposition 2.14.
1c. Express that $\min X \le 1$ and $\min X \ne 0$.
1d. Express that $\min X \le k$ and $\min X \not\le k-1$.
2a. $\min X \le \min Y$ if and only if

$$\exists A, B, R, S \ \ (0 \in R \ \wedge \ 0 \in S \ \wedge \ X = A + R \ \wedge \ Y = A + B + S)$$

Only if. Set $A = \{\min X\}$, $B = \{\min Y - \min X\}$, and $R = X - \min X$ and $S = Y - \min Y$.
If. $\min X = \min A \le \min A + \min B = \min Y$.
2b. Express that $\min X \le \min Y$ and $\min Y \le \min X$.
2c. $\min X \le \min Y + k$ if and only if

$$\forall A, R, B_1, \dots, B_{k+1}$$
$$((Y = A + R \ \wedge \ 0 \in R \ \wedge \ X = A + B_1 + \dots + B_{k+1}) \Rightarrow \bigvee_{i=1}^{i=k+1} 0 \in B_i)$$

2d & 2e. Express that $\min X \le \min Y + k$ and $\min X \not\le \min Y + k - 1$.
3a. $\min X + \min Y \le \min Z$ if and only if

$$\exists A, B, R, S, T \ \ (0 \in R \ \wedge \ 0 \in S$$
$$\wedge \ X = A + R \ \wedge \ Y = B + S \ \wedge \ Z = A + B + T)$$

3b. $\min X + \min Y = \min Z$ if and only if

$$\exists A, B, R, S, T \ \ (0 \in R \ \wedge \ 0 \in S \ \wedge \ 0 \in T \\ \wedge \ X = A + R \ \wedge \ Y = B + S \ \wedge \ Z = A + B + T)$$

□

2.3.5 Singleton sets

{pssingleton}

Proposition 2.17. *1. The predicate "*X is a singleton*" is* Π_2.
2. The predicate $X = \{0\}$ *is* Π_1.
3. The predicate $X = \{k\}$ *is* $\Sigma_2 \wedge \Pi_2$ *for* $k \geq 1$.
4. The predicate $Y \neq \emptyset \wedge X = \{\min Y\}$ *is* Π_2

Proof. 1. X is a singleton if and only if

$X \neq \emptyset$ and $\forall Y \ ((Y \neq \emptyset \wedge \min Y = \min X) \Rightarrow \exists Z \ Y = X + Z)$

Only if. Let $X = \{\ell\}$. Since $\min Y = \ell$ we have $(Y - \ell) + \{\ell\} = Y$ hence we can let $Z = Y - \ell$.
If. Let $Y = \{\min X\}$. Equality $\{\min X\} = X + Z$ implies X and Z are singleton sets and $X = \{\min X\}$ and $Z = \{0\}$.
Complexity: to remove the constant $\emptyset$, apply Proposition 2.13.
2. Already done in claim 2 of Proposition 2.14.
3. $X = \{k\}$ if and only if X is a singleton and $\min X = k$. We conclude with claim 3 of Proposition 2.16.
4. We express that X is a singleton and $\min X = \min Y$. □

3 Using submonoids to approximate and emulate

{s:using-submonoids}

3.1 Special cases of inclusion and membership

{ss:approx inclusion

The importance of the submonoids of $\mathbb{N}$ relies on the fact that they provide some approximation of two important relations, namely subset inclusion $Y \subseteq X$ and membership $x \in X$.

3.1.1 Special cases of inclusion

{sss:approx inclusion

We will prove in Theorem 5.7 that the inclusion predicate $Y \subseteq X$ is not expressible in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$. However, when X is a submonoid and Y contains 0 the inclusion is expressible.

{p:inclusion-in-submo

Proposition 3.1. *Let* $0 \in Y$ *and let* X *be a submonoid. Then*

$$Y \subseteq X \iff Y + X = X$$

Proof. Indeed, from right to left we have $X = Y + X \supseteq Y + \{0\} = Y$. Conversely, since $0 \in Y$ we have $X \subseteq Y + X$ and since $Y \subseteq X$ and X is a submonoid we have $Y + X \subseteq X + X = X$. □

The following result helps us to tightly evaluate syntactical complexity.

{p:maximal proper is

Proposition 3.2. *The predicate* $\lhd$ *is* $\Sigma_1 \wedge \Pi_1$. *Also, it is* Π_1 *on submonoids: there exists a* Π_1 *formula* $F(X,Y)$ *such that*

$$X, Y \text{ are submonoids} \Rightarrow (F(X,Y) \Leftrightarrow X \lhd Y)$$

Proof. Indeed, $Y \lhd X$ if and only if

$$\begin{aligned} &X, Y \text{ are submonoids } \wedge\ Y \subseteq X\ \wedge\ Y \neq X \\ &\quad \wedge\ \forall Z\ \big((Z \text{ is a submonoid } \wedge\ Y \subseteq Z \subseteq X) \Rightarrow (Z = Y\ \vee\ Z = X)\big) \end{aligned}$$

To conclude we use Propositions 2.15 and 3.1. □

3.1.2 Special cases of membership

{sss:approx membershi

Concerning the approximation of membership we use a special (whence the notation) type of submonoids which is appropriate to evaluate the complexity of the logic.

{def:Sn}

Definition 3.3. For each integer $n \geq 1$ we let

$$S_n = \{0\} \cup (n + \mathbb{N})$$

These submonoids are called *special*. The class of special submonoids is denoted by $\boldsymbol{\mathit{Special}}$.

E.g., $S_1 = \mathbb{N}$ and $S_2 = \{0\} \cup (2 + \mathbb{N}) = \mathbb{N} \setminus \{1\}$ is the largest proper submonoid of $\mathbb{N}$ since the minimum generating subset of $\mathbb{N}$ is $\{1\}$, cf. Proposition 2.10 Claim 1.

The following shows that the submonoids S_n, $n \geq 1$, can be used to test membership of a fixed n in X provided n is strictly greater than $\min(X)$. In particular, if X contains 0 then the property "$n > 0$ and belongs to X" for a fixed n is expressible since the restriction n greater than 0 is no longer necessary, see Proposition 2.14 Claim 3. Definability of the subset S_n is done in Theorem 3.6.

{l:Sn-minus-Sn+1}

Lemma 3.4. *Let* $m \in \mathbb{N}$ *and* $n \geq 1$. *If* $X \neq \emptyset$ *and* $m = \min X$ *then* $m + n \in X$ *if and only if* $X + S_n = X + S_{n+1}$.

Proof. Observe that

$$\begin{aligned} X + S_n &= X + (\{0\} \cup (n + \mathbb{N})) &= (X + \{0\}) \cup (X + (n + \mathbb{N})) \\ &= X \cup ((m + n) + \mathbb{N}) \\ X + S_{n+1} &= X \cup ((m + n + 1) + \mathbb{N})\ . \end{aligned}$$

Thus, $X + S_{n+1} \subseteq X + S_n$ and $(X + S_n) \setminus (X + S_{n+1}) = \{m + n\} \setminus X$. □

3.2 Definability issues of special submonoids

{ss:Sn}

We first show that each S_n can be defined. This is done by carefully investigating their proper maximal submonoids. In a second step we show that the fact of being an S_n, i.e., the class $\{S_n \mid n \geq 1\}$ is definable. This also relies on properties of the containment relation between submonoids.

3.2.1 Definability of each special submonoid

{sss:each-Sn}

Recall the notation $G(X)$ for the minimum generating set of X, Proposition 2.4.

{l:maxi submonoids Sn

Lemma 3.5. *Assume $n \geq 1$.*
1. $G(S_n) = \{n, \ldots, 2n-1\}$. Thus (cf. Proposition 2.10), S_n is a monoid with exactly n maximal proper submonoids.
2. $G(S_n \setminus \{n\}) = G(S_{n+1}) = \{n+1, \ldots, 2n+1\}$. Thus, $S_n \setminus \{n\}$ is a monoid with exactly $n+1$ maximal proper submonoids.
3. $G(S_n \setminus \{n+1\}) = \{n\} \cup \{n+2, \ldots, 2n-1\} \cup \{2n+1\}$. Thus, $S_n \setminus \{n+1\}$ is a monoid with exactly n maximal proper submonoids.
4. If $n+2 \leq i \leq 2n-1$ then $G(S_n \setminus \{i\}) = \{n, n+1, \ldots, i-1, i+1, \ldots, 2n-1\}$. Thus, $S_n \setminus \{i\}$ is a monoid with exactly $n-1$ maximal proper submonoids.

Proof. The right to left inclusions of the form $G(\ldots) \supseteq \ldots$ of the four claims are straightforward. We check the left to right inclusions.

1. For all integers $k \geq 0$ we have $2n+k = n+(n+k) \in (X \setminus \{0\}) + (X \setminus \{0\})$.
2. We apply Claim 1 with $n+1$ in place of n.
3. For all $k \geq 2$ we have $2n+k = n+(n+k) \in (X \setminus \{0\}) + (X \setminus \{0\})$.
4. For all $0 \leq k \neq i$ we have $2n+k = n+(n+k) \in (X \setminus \{0\}) + (X \setminus \{0\})$ and for $k = i$ we have $2n+i = (n+1)+(n+i-1)$. $\square$

We now convert the previous result formally in our logic.

{thm:Sn}

Theorem 3.6. *1. The predicate $X = S_1$ is $\Sigma_1 \wedge \Pi_1$.*
2. For each $n \geq 2$, the predicate $X = S_n$ is Δ_2.

Proof. 1. Since $S_1 = \mathbb{N}$ this is Proposition 2.14 Claim 4.
2. In view of applying Lemma 3.5, we introduce variables $X_1, \ldots, X_n$ to represent $S_1, \ldots, S_n$ and consider some formulas involving these variables.

(1) Let A be the formula expressing $X_1 = \mathbb{N}$.
(2) Let B be the formula expressing $X_n \lhd X_{n-1} \lhd \cdots \lhd X_2 \lhd X_1$.
(3) Let C be the formula which expresses that, for $m = 1, \ldots, n$, there exist $m+1$ pairwise distinct maximal proper submonoids of X_m.

Lemma 3.5 insures that every maximal submonoid of S_m has at most m maximal submonoids, except $S_m \setminus \{m\} = S_{m+1}$ which has $m+1$ maximal submonoids. Thus, a straightforward induction on $m = 1, \ldots, n$ shows that the formula $A \wedge B \wedge C$ implies that $X_m = S_m$. As a consequence, both formulas

$$\begin{aligned} & \exists X_1 \ldots X_n \; (A \wedge B \wedge C \wedge \; X_n = X) \\ \text{and} \quad & \forall X_1 \ldots X_n \; (A \wedge B \wedge C \Rightarrow \; X_n = X) \end{aligned}$$

express that $X = S_n$. By Proposition 3.2, formulas A, B are $\Sigma_1 \wedge \Pi_1$ and formula C is Σ_2. This shows that the predicate $X = S_n$ is Σ_2 and is Π_2. $\square$

3.2.2 Definability of the class of special submonoids

{sss:class-Sn}

We now look for a formula defining the class of special submonoids S_n, $n \geq 1$.

{def:partial M}

Definition 3.7. If M is a submonoid of $\mathbb{N}$ with m as minimum nonzero element, we denote by ∂M the submonoid $M \setminus \{m\}$ of M.

The following technical result based on Proposition 2.10 is crucial for a definition of the submonoids S_n. It is general and relates the generators of a submonoid with its maximal submonoids.

{p:creative}

Proposition 3.8. *Let M be a submonoid of $\mathbb{N}$, K a maximal submonoid of M and g a generator of M such that $K = M \setminus \{g\}$. For $k \in \mathbb{N}$, the following conditions are equivalent:*

(1) *There are exactly k generators of $M \setminus \{g\}$ which are not in $G(M)$, i.e. $G(M \setminus \{g\})$ is equal to $G(M) \setminus \{g\}$ augmented with k elements.*
(2) *There are exactly k sets in the class $\mathcal{Z}$ of maximal submonoids L of K such that K is the unique submonoid Z satisfying $L \lhd Z \lhd M$.*
(3) *There are exactly k sets in the class $\mathcal{Y}$ of maximal submonoids L of K such that K is the unique submonoid Z satisfying $L \subsetneq Z \subsetneq M$.*

Proof. We use Proposition 2.4. Any maximal submonoid L of K is of the form $L = K \setminus \{h\} = M \setminus \{g, h\}$ for some $h \in G(K)$. There are obviously only two sets Z such that $M \setminus \{g, h\} \subsetneq Z \subsetneq M$, namely $M \setminus \{g\} = K$ and $M \setminus \{h\}$. Thus, $K \in \mathcal{Y}$ if and only if $M \setminus \{h\}$ is not a submonoid. Observe that the sole possible reason for a failure of $L \lhd M \setminus \{h\} \lhd M$ is that $M \setminus \{h\}$ is not a submonoid. Thus, $K \in \mathcal{Z}$ if and only if $M \setminus \{h\}$ is not a submonoid. To conclude, observe that $M \setminus \{h\}$ is not a submonoid if and only if $h \notin G(M)$. $\square$

{def:creative}

Definition 3.9. Let M be a submonoid of $\mathbb{N}$ and $k \in \mathbb{N}$. A generator g of M is *k-creative* if condition (1) of Proposition 3.8 holds. A maximal submonoid K of M is k-creative if condition (2) of Proposition 3.8 holds, i.e. if $K = M \setminus \{g\}$ where g is a k-creative generator of M.

We shall write $(\geq \ell)$-creative to mean k-creative for some $k \geq \ell$.

Proposition 2.4 yields the following result.

{l:M setminus m}

Lemma 3.10. *If M is a submonoid different from $\{0\}$ then $m = \min(M \setminus \{0\})$ is a (≥ 2)-creative generator of M. Thus, ∂M is a (≥ 2)-creative maximal submonoid of M.*

Proof. Since m cannot be a sum of two nonzero elements of M we see that $m \in G(M)$. Also, $G(M \setminus \{m\}) \supseteq \{2m, m+p\}$ hence g is (≥ 2)-creative. $\square$

{def:good}

Definition 3.11. A submonoid M of $\mathbb{N}$ is *good* if ∂M is its unique maximal submonoid which is is (≥ 2)-creative, i.e. $\min M$ is the sole (≥ 2)-creative generator of M.

{l:Sn good}

Lemma 3.12. *The submonoids S_n, $n \geq 2$, are good.*

Proof. We use Lemma 3.5. The generators of S_n are $n, \ldots, 2n-1$.

Case of $S_n \setminus \{n\}$. The generators of $S_n \setminus \{n\} = S_{n+1}$ are $n+1, \ldots, 2n+1$. Only two of them (namely, $2n, 2n+1$) are not in $G(S_n)$. Thus, n is a 2-creative generator of S_n.

Case of $S_n \setminus \{n+1\}$. The generators of $S_n \setminus \{n+1\}$ are the elements in $\{n\} \cup \{n+2, \ldots, 2n-1\} \cup \{2n+1\}$. Only one of them (namely, $2n+1$) is not in $G(S_n)$. Thus, $n+1$ is a 1-creative generator of S_n.

Case of $S_n \setminus \{i\}$ with $n+2 \leq i \leq 2n-1$. The generators of $S_n \setminus \{i\}$ are the elements in $\{n, \ldots, 2n-1\} \setminus \{i\}$. All of them are in $G(S_n)$. Thus, i is a 0-creative generator of S_n.

This shows that $n = \min S_n$ is the sole (≥ 2)-creative generator of S_n. $\square$

The submonoids S_n are not the sole examples of good submonoids.

Example 3.13. *The monoid $M = \{0, 6, 7, 8, 9\} \cup (11 + \mathbb{N})$ with minimum generating set $\{6, 7, 8, 9, 11\}$ is good. Indeed, using Remark 2.7, we get the facts shown in the following table:*

g	$M \setminus \{g\}$	$G(M \setminus \{g\})$	
6	$\{0,7,8,9\} \cup (11+\mathbb{N})$	$\{7,8,9,11,\mathbf{12},\mathbf{13}\}$	6 *is* 2-*creative*
7	$\{0,6,8,9\} \cup (11+\mathbb{N})$	$\{6,8,9,11,\mathbf{13}\}$	7 *is* 1-*creative*
8	$\{0,6,7,9\} \cup (11+\mathbb{N})$	$\{6,7,9,11\}$	8 *is* 0-*creative*
9	$\{0,6,7,8\} \cup (11+\mathbb{N})$	$\{6,7,8,11\}$	9 *is* 0-*creative*
11	$\{0,6,7,8,9\} \cup (12+\mathbb{N})$	$\{6,7,8,9\}$	11 *is* 0-*creative*

Remark 3.14. Not every submonoid is good. For instance, the submonoid $X = \{0, 3, 5, 6\} \cup (8 + \mathbb{N})$ in Example 2.11 has two (≥ 2)-creative generators. Indeed, $G(X) = \{3, 5\}$ and $G(X \setminus \{3\}) = \{5, 6, 8, 9\}$ and $G(X \setminus \{5\}) = \{3, 8, 10\}$ hence 3 is 3-creative and 5 is 2-creative.

{l:good Sigma4}

Lemma 3.15. *1. The following predicate is Σ_2 :*

$$K \text{ is a } (\geq 2)\text{-creative maximal submonoid of the submonoid } M$$

2. The class of good submonoids is Π_2.

Proof. 1. Using condition (2) of Proposition 3.8, let $A(M,K)$ be the formula

$$\begin{aligned} K \lhd M \ \wedge \ \exists L_1, L_2 \ \big(L_1 \neq L_2 \ \wedge \ L_1 \lhd K \ \wedge \ L_2 \lhd K \\ \wedge \ \forall L \ (L \text{ is a submonoid } \Rightarrow \\ \big(L = K \Leftrightarrow (L_1 \subsetneq L \subsetneq M)\big) \ \wedge \ (L = K \Leftrightarrow L_2 \subsetneq L \subsetneq M)\big)\big) \end{aligned}$$

Since the class of submonoids is Σ_1 and the predicate $\lhd$ is $\Sigma_1 \wedge \Pi_1$ and inclusion of submonoids is Σ_0 (cf. Propositions 3.1, 2.15, 3.2), the above formula $A(M,K)$ is Σ_2.

2. By definition M is good if and only if M is a submonoid and there exists a unique K such that $A(M,K)$. Using Lemma 3.10, we see that it suffices to say that there exists at most one K such that $A(M,K)$, i.e.

$$M \text{ is a submonoid } \wedge \ \forall K' \forall K'' \ ((A(M,K') \wedge A(M,K'')) \Rightarrow K' = K'')$$

Since A is Σ_2, this formula is Π_2. □

{l:partial}

Lemma 3.16. *The following predicate is $\Sigma_2 \wedge \Pi_2$*

$$\{(L,M) \mid M \textit{ is a good submonoid and } L = \partial M\}$$

Proof. By Lemma 3.10, when M is good, ∂M is the unique K such that $A(M,K)$. Thus, the formula M *is good* $\wedge$ $A(M,L)$ defines the considered predicate. It is Π_2 by Lemma 3.15. □

We can now get a useful extension of Lemma 3.4.

{l:m in L}

Lemma 3.17. *Let L, M be submonoids and m be the minimum nonzero element of M. Then $m \in L$ if and only if $L + \partial M = L + M$.*

Proof. Trivially, if $m \in L$ then $L + \partial M = L + M$. Suppose now $L + \partial M = L + M$. Since $m \in L + M$ we have $m \in L + \partial M$ hence $m = x + y$ with $x \in L$ and $y \in \partial M$. Since all nonzero elements of ∂M are strictly greater than m we have $y = 0$ hence $m = x \in L$. □

The following proves the definability of the class of special submonoids S_n.

{thm:iss}

Theorem 3.18. *The class* `Special` $= \{S_n \mid n \geq 1\}$ *is* Π_3.

Proof. Consider the following formula $\texttt{Special}(X)$ which (by Lemma 3.17 and Proposition 3.1) is Π_4 and expresses that X is a submonoid and, for any good M with m as minimum nonzero element, if $m \in X$ then $M \subseteq X$:

$$\begin{aligned} &X \text{ is a submonoid and } \forall M\ \forall L \\ &\qquad \big(M \text{ is a good submonoid and } L = \partial M \text{ and } X + L = X + M \\ &\qquad\qquad\qquad\qquad\qquad\qquad \Rightarrow X + M = X\big) \end{aligned}$$

The submonoids S_n clearly satisfy this property. Conversely, if X satisfies this property and n is the minimum nonzero element of X then, applying the property with the good submonoid $M = S_n$, we get $S_n \subseteq X$ hence $S_n = X$. □

{p:succ Sn}

Proposition 3.19. *The following predicates are* Π_3 *:*

$$\begin{array}{rcl} \texttt{Succ}_1(X,Y) & \equiv & (X,Y) \in \{(S_n, S_{n+1}) \mid n \geq 1\} \\ \texttt{Succ}_k(X,Y) & \equiv & (X,Y) \in \{(S_n, S_{n+k}) \mid n \geq 1\} \\ \texttt{Succ}_*(X,Y) & \equiv & (X,Y) \in \{(S_n, S_{n+k}) \mid n, k \geq 1\} \end{array}$$

For $k = 1$ *we simply write* $\texttt{Succ}$ *in place of* $\texttt{Succ}_1$.

Proof. Observe that

$$\begin{array}{rcl} \texttt{Succ}_1(X,Y) & \Longleftrightarrow & X \text{ is special and } Y = \partial X \\ \texttt{Succ}_k(X_0,Y) & \Longleftrightarrow & \texttt{Special}(X_0) \wedge \forall X_1 \ldots \forall X_k \\ & & \left(\left(\bigwedge_0^{k-1} X_i \text{ is good and } X_{i+1} = \partial X_i\right) \Rightarrow Y = X_k\right) \\ \texttt{Succ}_*(X,Y) & \Longleftrightarrow & X, Y \text{ are special and } Y \subseteq X \end{array}$$

then apply Theorem 3.18, Lemma 3.16 and Proposition 3.1. □

3.3 Addition and multiplication on special submonoids

{ss:operations-on-Sn}

Here we show that the set of submonoids of the form S_n, with $n \geq 1$, can be equipped with two definable operations $\oplus$ and $\otimes$ which make it isomorphic to $\langle \mathbb{N} \setminus \{0\}; +, \times, = \rangle$.

3.3.1 Insight into the proof

{sss:sketch}

This paragraph is meant to give some intuition behind the formal proofs of the next two ones. The idea is to define two operations on the family of special submonoids, namely an addition $(S_n, S_p) \to S_{n+p}$ and a multiplication $(S_n, S_p) \to S_{n \times p}$.

The addition is defined via the finite initial segments by observing that $\{0,\ldots,n+p\} = \{0,\ldots,n\}+\{0,\ldots,p\}$ holds and by using the correspondence $S_n \mapsto \{0,\ldots,n\}$ (which is definable, cf. Proposition 3.25).

Based on a number theoretic result which we recall below, multiplication can be expressed by using addition and divisibility. Divisibility is defined via the sets of the form $n\mathbb{N}$ by observing that n divides p if and only if $p\mathbb{N} \subseteq n\mathbb{N}$ and by using the correspondence $S_n \mapsto n\mathbb{N}$ (which is definable, cf. Proposition 3.28).

{l:mult from add divi

Lemma 3.20. *Multiplication on $\mathbb{N}$ is definable from addition and divisibility. More precisely, it is Δ_1 relative to addition and some predicates which are themselves Π_1 relative to divisibility.*

Proof. Schnirelman's famous result (1931) insures the existence of a constant K such that every integer ≥ 2 is the sum of at most K primes. Olivier Ramaré, [8], showed that $K \leq 7$. If $x = \sum_{i=1}^{a} p_i$ and $y = \sum_{j=1}^{b} q_j$ with $a, b \leq 7$ and the numbers p_i, q_j are primes then $x \times y = \sum_{i=1}^{a} \sum_{j=1}^{b} p_i \times q_j$. Now, the product of two primes p, q (distinct or not) is the unique number with p, q as sole proper divisors. Let $s \,|\, t$ mean that s is a divisor of t, let $P(x)$ mean that x is prime and let $A(x,p,q)$ mean that p, q are prime and $x = p \times q$. Then

$$\begin{array}{rcl} P(p) & \equiv & p \neq 1 \wedge \forall s\ (s \,|\, p \iff s = 1 \vee s = p) \\ A(x,p,q) & \equiv & P(p) \wedge P(q) \wedge x \neq p, q \\ & & \wedge\ \forall s\ (s \,|\, x \iff s = 1 \vee s = x \vee s = p \vee s = q) \end{array}$$

are Π_1 relative to divisibility. Also, the predicate $z = x \times y$ is expressed as the conjunction of the formulas $(x = 0 \vee y = 0) \Rightarrow z = 0$, $x = 1 \Rightarrow z = y$ and $y = 1 \Rightarrow z = x$ and any one of the following formulas:

$$E(x,y,z) \equiv x, y \geq 2 \Rightarrow \bigvee_{a,b \in \{1,\ldots,7\}} \exists (x_i, y_j, z_{i,j})_{1 \leq i \leq a}^{1 \leq j \leq b} \left(\varphi \ \wedge\ z = \sum_{i,j} z_{i,j}\right)$$

$$F(x,y,z) \equiv x, y \geq 2 \Rightarrow \bigvee_{a,b \in \{1,\ldots,7\}} \forall (x_i, y_j, z_{i,j})_{1 \leq i \leq a}^{1 \leq j \leq b} \left(\varphi \Rightarrow z = \sum_{i,j} z_{i,j}\right)$$

where φ is $x = \sum_i x_i \ \wedge\ y = \sum_j y_i \ \wedge \bigwedge_{i,j} P(x_i) \wedge P(y_j) \wedge A(z_{i,j}, x_i, y_j)$. □

3.3.2 Addition on special submonoids

{def:addition}

Definition 3.21. We denote by `Initial` the class of all initial segments $\{0,\ldots,n\}$ for some $n \in \mathbb{N}$.

{p:initial segment}

Proposition 3.22. *The class* `Initial` *is* Π_4

Proof. Observe that $X \in$ `Initial` if and only if $0 \in X$ and $X \neq \mathbb{N}$ and, for all $x \geq 2$, if $x \in X$ then $x - 1 \in X$. This is expressible as follows:

$$\begin{array}{l} 0 \in X \ \wedge\ X \neq \mathbb{N} \ \wedge \forall Z, T, U \\ \quad \big((\mathtt{Succ}(Z,T) \wedge \mathtt{Succ}(T,U) \wedge X + T = X + U) \Rightarrow X + Z = X + T\big) \end{array}$$

using the predicate `Succ` defined in Proposition 3.19 and Lemma 3.4. □

As explained in paragraph 3.3.1 we view an integer as the maximal element of an initial segment which allows us to indirectly express that it belongs to some subset containing 0.

{p:max X in Y}

Proposition 3.23. *The following two predicates are* Π_4

$$X \in \texttt{Initial} \wedge 0 \in Y \wedge \max X \in Y \quad , \quad X \in \texttt{Initial} \wedge 0 \in Y \wedge \max X \notin Y$$

Proof. Observe that the maximum element of a finite initial segment X non reduced to $\{0\}$ is the integer n such that $n \in X$ and $n+1 \notin X$. Thus, $\max X \in Y$ is expressed by the formula

$$\begin{aligned} &\texttt{Initial}(X) \ \wedge\ 0 \in Y \ \wedge \big(X \neq \{0\} \Rightarrow \forall Z\ \forall T\ \forall U \\ &\quad \big(\texttt{Succ}(Z,T) \wedge \texttt{Succ}(T,U) \wedge (X+Z = X+T) \wedge (X+T \neq X+U) \\ &\qquad\qquad\qquad\qquad\qquad \Rightarrow Y+Z = Y+T\big)\end{aligned}$$

Idem for the second predicate with $\max X \notin Y$: just replace the last equality $Y+Z = Y+T$ by an inequality. The stated complexity comes from Propositions 3.22, 3.19 and 2.14. □

{p:max-X+max-Y}

Proposition 3.24. *The following predicate is* Π_4

$$X, Y, Z \in \texttt{Initial} \ \wedge\ \max X + \max Y = \max Z$$

Proof. Observe that equality $\max X + \max Y = \max Z$ is equivalent to $X + Y = Z$ when X, Y, Z are finite initial segments. □

{p:SmaxX}

Proposition 3.25. *The following predicate is* Π_4 *:*

$$X \in \texttt{Initial} \ \wedge\ X \neq \{0\} \ \wedge\ Y = S_{\max X}$$

Proof. Observe that $\max X$ is the largest n such that $\max X \in S_n$. Thus, the predicate can be expressed as follows:

$$\begin{aligned} &X \in \texttt{Initial} \ \wedge\ X \neq \{0\} \ \wedge\ \max X \in Y \wedge\ \texttt{Special}(Y) \\ &\qquad\qquad\qquad\qquad \wedge\ \forall Z\ (\texttt{Succ}(Y,Z) \Rightarrow \max X \notin Z)\end{aligned}$$

using Proposition 3.23 and Lemma 3.4. □

{thm:Sn+p}

Theorem 3.26. *The relation* $\{(S_n, S_p, S_{n+p}) \mid n, p \geq 1\}$ *is* Δ_5.
We write $T \oplus U = V$ *if* (T, U, V) *is in this relation.*

Proof. Using Propositions 3.25 and 3.24, $T \oplus U = V$ holds if and only if it any one of the following formulas holds

$$\varphi \ \wedge\ \exists I\ \exists J\ \exists K\ (\psi\ \wedge I + J = K) \quad , \quad \varphi\ \wedge\ \forall I\ \forall J\ \forall K\ (\psi\ \Rightarrow I + J = K)$$

$$\text{where} \left\{ \begin{array}{lll} \varphi & \equiv & \texttt{Special}(T) \ \wedge\ \texttt{Special}(U) \ \wedge\ \texttt{Special}(V) \\ \psi & \equiv & I, J, K \in \texttt{Initial} \ \wedge\ I, J, K \neq \{0\} \\ & & \wedge\ T = S_{\max I} \ \wedge\ U = S_{\max J} \ \wedge\ V = S_{\max K} \end{array} \right. .$$ □

3.3.3 Multiplication on special submonoids

{sss:multiplication}

We introduce two useful predicates.

{p:the-bNs}

Proposition 3.27. *The predicate* $\texttt{Periodic} = \{n\mathbb{N} \mid n \geq 1\}$ *is* Π_2.

Proof. By Proposition 2.12, the sets $n\mathbb{N}$, $n \geq 1$, are the nonzero submonoids with no minimal supermonoid:

$$\texttt{Periodic}(X) \iff (X \neq \{0\} \text{ is a submonoid and } \forall Y\ \neg(X \lhd Y))$$

Conclude with Proposition 3.2. □

{p:Sn-divides-p}

Proposition 3.28. *1. The predicate* $\mathcal{B} = \{(S_n, n\mathbb{N}) \mid n \geq 1\}$ *is* Π_3.
2. The relation $\{(S_n, S_p) \mid n \geq 1$ *and* n *divides* $p\}$ *is* Δ_4.

Proof. 1. Recall that $\texttt{Periodic}(X)$ is the Π_2 predicate of Proposition 3.27. Observing that $n\mathbb{N}$ is the unique submonoid which is included in S_n and not in $\partial(S_n) = S_{n+1}$, the predicate $(X, Y) \in \mathcal{B}$ is expressible as

$$\texttt{Special}(X)\ \wedge\ \texttt{Periodic}(Y)\ \wedge\ Y \subseteq X\ \wedge\ \forall Z\ (Z = \partial(X) \Rightarrow Y \not\subseteq Z)$$

Theorem 3.18, Proposition 3.1 and Lemma 3.16 give the complexity.
2. Recall that n divides p if and only if $p\mathbb{N} \subseteq n\mathbb{N}$. Thus, the formulas

$$\begin{cases} \exists Y\ \exists Y'\ (\mathcal{B}(X, Y)\ \wedge \mathcal{B}(X', Y')\ \wedge\ Y' \subseteq Y) \\ \forall Y\ \forall Y'\ ((\mathcal{B}(X, Y)\ \wedge \mathcal{B}(X', Y')) \Rightarrow Y' \subseteq Y) \end{cases}$$

(which are Σ_4 and Π_4) define the divisibility predicate on the submonoids S_n. □

{thm:Snxp}

Theorem 3.29. *The relation* $\{(S_n, S_p, S_{n\times p}) \mid n, p \geq 1\}$ *is* Δ_5. *We write* $T \otimes U = V$ *if* (T, U, V) *is in this relation.*

Proof. Do the following in the formulas $E(x, y, z)$ and $F(x, y, z)$ given in the proof of Lemma 3.20 (and involving the predicates P and A):
- replace the variables $x, y, z, x_i, y_j, z_{i,j}$ by $X, Y, Z, X_i, Y_j, Z_{i,j}$,
- replace addition by the Δ_5 predicate $\oplus$ (cf. Theorem 3.26),
- using claim 2 of Proposition 3.28, replace the predicates P and A which are Π_1 relative to divisibility, by Π_4 predicates in $X, Y, Z, X_i, Y_j, Z_{i,j}$. □

We now extend the two elementary operations of addition and multiplication to an arbitrary polynomial.

{cor:Snpoly}

Corollary 3.30. *Let* $T(x_1, \ldots, x_k)$ *be a polynomial with non zero coefficients in* $\mathbb{N}$ *and variables in* $\{x_1, \ldots, x_n\}$ *with* $n \geq 1$. *The following relation is* Δ_5 *:*

$$\{(S_{n_1}, \ldots, S_{n_k}, S_p) \mid n_1, \ldots, n_k \geq 1 \text{ and } p = T(n_1, \ldots, n_k)\}$$

Proof. There are polynomials $T_0, \ldots, T_s$ such that $T = T_s$ and, for $0 \le i \le s$, T_i is the constant 1 or a variable or $T_j + T_\ell$ or $T_j \times T_\ell$ with $j, \ell < i$. Let I be the set of numbers i such that $T_i = 1$, let J be the set of pairs (i, m) such that $T_i = x_m$, A (resp. M) be the set of triples (i, j, ℓ) such that $T_i = T_j + T_\ell$ (resp. $T_i = T_j \times T_\ell$). The relation is expressed by either of the following formulas with free variables $X_1, \ldots, X_k, X$:

$$\exists Z_1 \ldots \exists Z_s \ (\varphi \ \wedge \ X = Z_s) \quad , \quad \forall Z_1 \ldots \forall Z_s \ (\varphi \ \Rightarrow \ X = Z_s)$$

where φ is $\bigwedge_{i \in I} Z_i = S_1 \ \wedge \ \bigwedge_{(i,m) \in J} Z_i = X_m$
$$\wedge \ \bigwedge_{(i,j,\ell) \in A} Z_i = Z_j \oplus Z_\ell \ \wedge \ \bigwedge_{(i,j,\ell) \in M} Z_i = Z_j \otimes Z_\ell$$

By Theorems 3.6, 3.26, 3.29, these formulas are respectively Σ_5 and Π_5. ☐

4 Complexity of the theory

{s:theory-complexity}

4.1 Emulating second-order arithmetic

{ss:2d arithm}

Here, we show that we can interpret the second-order theory of arithmetic in the theory of $\langle \mathcal{P}(\mathbb{N}); =, + \rangle$.

{thm:2d arithm}

Theorem 4.1. *1. To each second-order arithmetical formula φ one can computably associate a formula* $\texttt{Trad}(\varphi)$ *so that if φ has m free first-order variables and n free second-order variables then* $\texttt{Trad}(\varphi)$ *has $m + n$ free variables and, for all $a_1, \ldots, a_m \in \mathbb{N}$ and $A_1, \ldots, A_n \subseteq \mathbb{N}$,*

$$\langle \mathbb{N}, \mathcal{P}(\mathbb{N}); =, \in, 1, +, \times \rangle \models \varphi(a_1, \ldots, a_m, A_1, \ldots, A_n) \iff$$
$$\langle \mathcal{P}(\mathbb{N}); =, + \rangle \models \texttt{Trad}(\varphi)(S_{1+a_1}, \ldots, S_{1+a_m}, \{0\} \cup (1+A_1), \ldots, \{0\} \cup (1+A_1))$$

2. If φ is quantifier-free then $\texttt{Trad}(\varphi)$ *can be taken either Σ_5 or Π_5. If φ is in prenex form with a nonempty quantifier prefix of the form $Q_1\xi_1 \ldots Q_k\xi_k$ where the variables ξ_i are first or second order variables and there are ℓ alternating blocks of quantifiers $\exists, \forall$ in $Q_1 \ldots Q_k$ then* $\texttt{Trad}(\varphi)$ *can be taken $\Sigma_{\ell+4}$ if $Q_1 = \exists$ and $\Pi_{\ell+4}$ if $Q_1 = \forall$.*

Proof. 1. The transformation $\texttt{Trad}$ is defined as the composition $\Omega \circ \Theta$ of two reductions. The first reduction Θ allows to go from the second-order arithmetical structure of $\mathbb{N}$ to that of $\mathbb{N} \setminus \{0\}$. The second reduction Ω allows to go to the structure $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$.

The reduction Θ maps a second-order arithmetical formula φ to another such formula $\Theta(\varphi)$ with the same free variables so that, for all $a_1, \ldots, a_m \in \mathbb{N}$ and $A_1, \ldots, A_n \subseteq \mathbb{N}$,

$$\langle \mathbb{N}, \mathcal{P}(\mathbb{N}); =, \in, 1, +, \times \rangle \models \varphi(a_1, \ldots, a_m, A_1, \ldots, A_n) \iff$$
$$\langle \mathbb{N} \setminus \{0\}, \mathcal{P}(\mathbb{N} \setminus \{0\}); =, \in, 1, +, \times \rangle \models \Theta(\varphi)(1+a_1, \ldots, 1+a_m, 1+A_1, \ldots, 1+A_n)$$

Thus, in $\Theta(\varphi)$, integer quantifications are over $\mathbb{N}\setminus\{0\}$ and set quantifications are over $\mathcal{P}(\mathbb{N}\setminus\{0\})$. This is simply done by replacing in φ any equation over integers $Q(x_1,\ldots,x_k) = R(x_1,\ldots,x_k)$ by the equation obtained from $Q(x_1-1,\ldots,x_k-1) = R(x_1-1,\ldots,x_k-1)$ by developing and moving any monomial with negative coefficient from one side to the other side of the equation. For instance, $\Theta(x+y=z)$ is obtained by the above process from $(x-1)+(y-1) = z-1$ hence $\Theta(x+y=z)$ is $x+y=z+1$; similarly $\Theta(x\times y = z)$ is obtained from $(x-1)\times(y-1) = z-1$ and is therefore $x\times y+2 = x+y+z$. As for subformulas $x \in X$, they are left unchanged.

We now define Ω which maps a second-order arithmetical formula ψ to a formula $\Omega(\psi)$ with the same number of free variables so that, for all $b_1,\ldots,b_m \in \mathbb{N}\setminus\{0\}$ and $B_1,\ldots,B_n \subseteq \mathbb{N}\setminus\{0\}$,

$$\begin{aligned}&\langle \mathbb{N}\setminus\{0\}, \mathcal{P}(\mathbb{N}\setminus\{0\}); =, \in, 1, +, \times\rangle \models \psi(b_1,\ldots,b_m,B_1,\ldots,B_n)\\ &\qquad\iff \langle \mathcal{P}(\mathbb{N}); =, +, \rangle \models \Omega(\psi)(S_{b_1},\ldots,S_{b_m},\{0\}\cup B_1,\ldots,\{0\}\cup B_n)\end{aligned}$$

First, we distinguish two disjoint infinite families of set variables $(U_i)_{i\in\mathbb{N}}$ and $(V_i)_{i\in\mathbb{N}}$. The variables U_i in $\Omega(\psi)$ are to vary over the class of special submonoids (i.e. the sets S_n), they correspond to first-order variables in ψ: if x_i takes value $n \in \mathbb{N}\setminus\{0\}$ then U_i is to take value $S_n \in \mathcal{P}(\mathbb{N})$. The variables V_i in $\Omega(\psi)$ correspond to the second-order variables in ψ : if X_i takes value $B \in \mathcal{P}(\mathbb{N}\setminus\{0\})$ then V_i is to take value $\{0\}\cup B \in \mathcal{P}(\mathbb{N})$.

In view of Claim 2, we inductively define two variants of Ω, namely $\Omega^\exists$ and $\Omega^\forall$.

(1) If ψ is an atomic formula $Q(x_1,\ldots,x_k) = R(x_1,\ldots,x_k)$ then $\Omega^\exists(\psi)$ and $\Omega^\forall(\psi)$ are the Σ_5 and Π_5 formulas

$$\begin{aligned}\Omega^\exists(\psi) &\equiv \exists U\ \big(A(U_1,\ldots,U_k,U)\ \wedge\ B(U_1,\ldots,U_k,U)\big)\\ \Omega^\forall(\psi) &\equiv \texttt{Special}(U_k)\ \wedge\ldots\wedge\ \texttt{Special}(U_k)\ \wedge\ \forall U\ \forall U'\\ &\qquad ((A(U_1,\ldots,U_k,U)\ \wedge\ B(U_1,\ldots,U_k,U')) \Rightarrow U = U')\end{aligned}$$

where A, B are Σ_6 formulas associated to Q and R by Corollary 3.30 (thus, the i-th first-order variable x_i is replaced by the variable U_i). Note: the subformulas $\texttt{Special}(U_i)$ are omittted in $\Omega^\exists(\psi)$ since they are implied by $A(U_1,\ldots,U_k,U)$.

(2) If ψ is $x_i \in X_m$ then, relying on Lemma 3.4, $\Omega^\exists(\psi)$ and $\Omega^\forall(\psi)$ are the Σ_4 and Π_4 formulas (cf. Proposition 3.19)

$$\begin{aligned}\Omega^\exists(\psi) &\equiv 0 \in V_m\ \wedge\ \exists U\ (\texttt{Succ}(U_i,U)\ \wedge\ V_m + U_i = V_m + U)\\ \Omega^\forall(\psi) &\equiv \texttt{Special}(U_i)\ \wedge\ 0 \in V_m\\ &\qquad \wedge\ \forall U\ (\texttt{Succ}(U_i,U)\ \Rightarrow\ V_m + U_i = V_m + U)\end{aligned}$$

(3) $\Omega^\exists$ and $\Omega^\forall$ commute with conjunction and disjunction.

(4) $\Omega^\exists(\neg\varphi) = \neg\Omega^\forall(\varphi)$ and $\Omega^\forall(\neg\varphi) = \neg\Omega^\exists(\varphi)$.

(5) $\Omega^\exists(\exists x_i\ \psi) = \Omega^\forall(\exists x_i\ \psi) = \exists U_i\ (\texttt{Special}(U_i) \wedge \Omega^\exists(\psi))$
$\Omega^\exists(\forall x_i\ \psi) = \Omega^\forall(\forall x_i\ \psi) = \forall U_i\ (\texttt{Special}(U_i) \Rightarrow \Omega^\forall(\psi))$

(6) $\Omega^\exists(\exists X_m\ \psi) = \Omega^\forall(\exists X_m\ \psi) = \exists V_m\ (0 \in V_m \wedge \Omega^\exists(\psi))$
$\Omega^\exists(\forall X_m\ \psi) = \Omega^\forall(\forall X_m\ \psi) = \forall V_m\ (0 \in V_m \Rightarrow \Omega^\forall(\psi))$

Letting Ω be either $\Omega^\exists$ or $\Omega^\forall$, Corollary 3.30 and Lemma 3.4, show that the above clauses insure the wanted property of Ω hence also those of $\texttt{Trad} : \varphi \mapsto \Omega(\Theta(\varphi))$.
2. The assertion about quantifier-free formulas φ is clear from clauses (1) and (2). An easy induction on the complexity of φ shows that if φ has ℓ alternating blocks of quantifiers and the last one is a Q-block then $\mathrm{Trad}(\Omega^Q(\Theta(\varphi)))$ is $\Sigma_{\ell+4}$ if the first block is a $\exists$-block and is $\Pi_{\ell+4}$ if the first block is a $\forall$-block. □

4.2 The theory of addition on sets is undecidable

{ss:the-theory-is-und

The complexity results concerning the theory $\langle \mathcal{P}(\mathbb{N}); +, =\rangle$ are direct consequences of the results in the previous sections.

{thm:theory undecidab

Theorem 4.2. *The class of Σ_5 sentences true in $\langle \mathcal{P}(\mathbb{N}); +, =\rangle$ is undecidable.*

Proof. Use the undecidability of the Diophantine theory of $\langle \mathbb{N}; =, 1, +, \times\rangle$ (Matiyasevich's celebrated result) and Theorem 4.1. □

The theory of addition on sets is, in fact, *highly* undecidable.

{thm:theory highly un

Theorem 4.3. *The class $\mathcal{T}$ of sentences true in $\langle \mathcal{P}(\mathbb{N}); =, +\rangle$ is recursively isomorphic to the second order theory $\mathcal{A}$ of $\langle \mathbb{N}; =, +, \times\rangle$, i.e. there exists a computable bijection θ between the set of first-order formulas in the language $\{=, +\}$ and the set of second-order formulas in the language $\{=, +, \times\}$ such that $\mathcal{T} = \theta^{-1}(\mathcal{A})$.*

{rk:beyond}

Remark 4.4. The class of sentences with quantifications over $\mathbb{N}$ only which are true in $\langle \mathbb{N}; =, 1, +, \times\rangle$ (i.e. the first order theory of arithmetic) is Δ^1_1 and not Σ^0_n for any $n \in \mathbb{N}$. As for the class of sentences with quantifications over $\mathbb{N}$ and over $\mathcal{P}(\mathbb{N})$ which are true in $\langle \mathbb{N}, \mathcal{P}(\mathbb{N}); =, \in, 1, +, \times\rangle$ (i.e. the second order theory of arithmetic), it is Δ^2_1 and not Σ^1_n for any $n \in \mathbb{N}$. Thus, the Turing degree of the second order theory of arithmetic is an order of magnitude higher than that of the first order theory.

Proof of Theorem 4.3. Recall Myhill's isomorphism theorem which is the computable analog of Cantor-Bernstein's theorem in set theory (cf. [10] Theorem VI page 85, or [7] Theorem III.7.13 page 325): if $X, Y \subseteq \mathbb{N}$ and there exists computable injective maps $\varphi : \mathbb{N} \to \mathbb{N}$ and $\psi : \mathbb{N} \to \mathbb{N}$ such that $X = \varphi^{-1}(Y)$ and $Y = \psi^{-1}(X)$ then there exists a computable bijective map $\theta : \mathbb{N} \to \mathbb{N}$ such that $X = \theta^{-1}(Y)$. Thus, to prove the theorem it suffices to get injective computable reductions of $\mathcal{T}$ to $\mathcal{A}$ and of $\mathcal{A}$ to $\mathcal{T}$.

Recursive reduction of $\mathcal{T}$ to $\mathcal{A}$. To each sentence about $\langle \mathcal{P}(\mathbb{N}); =, + \rangle$ associate the second order arithmetical formula obtained by replacing any subformula $X + Y = Z$ by

$$\forall r \ (r \in Z \iff \exists p, q \ (n = p + q \wedge p \in X \wedge q \in Y) \ .$$

Recursive reduction of $\mathcal{A}$ to $\mathcal{T}$. Use Theorem 4.1. □

5 Non definable predicates

{s:nondefinable}

5.1 What is so special about the predicate $0 \in X$?

{ss:zero}

As implicitly used in numerous instances in the previous sections, the existence of 0 in a subset seems to be the crux for proving remarkable properties such as those in paragraph 3.1. If the logic could allow us to add 0 to an arbitrary subset then we could extend these properties to all subsets. However this is not possible. We show that the following predicate is not definable:

$$Y = X \cup \{0\} \tag{7}$$ {eq:hypothetical-pred

The proof uses two ingredients. The first one, cf. Lemma 5.2 below, insures that equations and inequations between set variables can be split into conditions on elements of $\mathbb{N}$ and conditions on subsets which are either empty or contain 0. This transformation is lifted to formulas in Lemma 5.3 below. The second ingredient, cf. Lemma 5.4 below, is a general result about formulas consisting of combinations of claims on disjoint sets of variables.

{not:M}

Notation 5.1. Let $\mathcal{P}_0(\mathbb{N})$ be the class of sets which contain 0. Let $\mathcal{P}_{0,\emptyset}(\mathbb{N}) = \mathcal{P}_0(\mathbb{N}) \cup \{\emptyset\}$ be the class of sets which contain 0 or are empty. Let

(1) $\mathcal{E}$ be the subset $\{\emptyset\}$ of $\mathcal{P}_{0,\emptyset}$

(2) $+_{\mathbb{N}}$ and $=_{\mathbb{N}}$ be addition and equality on $\mathbb{N}$,

(3) $+_{\mathcal{P}_{0,\emptyset}(\mathbb{N})}$ and $=_{\mathcal{P}_{0,\emptyset}(\mathbb{N})}$ be addition and equality of sets in $\mathcal{P}_{0,\emptyset}(\mathbb{N})$.

We consider the following two sort structure:

$$\mathcal{M} \ = \ \langle \mathbb{N}, \mathcal{P}_{0,\emptyset}(\mathbb{N}); +_{\mathbb{N}}, +_{\mathcal{P}_{0,\emptyset}}, \mathcal{E}, =_{\mathbb{N}}, =_{\mathcal{P}_{0,\emptyset}} \rangle$$

{l:split}

Lemma 5.2. *Let I, J be disjoint subsets such that $I \cup J = \{1, \ldots, n\}$. Then, for all integers $a_1, \ldots, a_n \in \mathbb{N}$ and sets $A_1, \ldots, A_n \in \mathcal{P}_{0,\emptyset}$,*

$$\langle \mathcal{P}(\mathbb{N}); +, = \rangle \models \sum_{i \in I} a_i + A_i = \sum_{j \in J} a_j + A_j \iff \mathcal{M} \models \psi(a_1, \ldots, a_n, A_1, \ldots, A_n)$$

where $\psi(x_1, \ldots, x_n, X_1, \ldots, X_n)$ is a Boolean combination of formulas $\mathcal{E}(X_\ell)$, for $\ell = 1, \ldots, n$, and equalities $\sum_{i \in I} x_i = \sum_{j \in J} x_j$ and $\sum_{i \in I} X_i = \sum_{j \in J} X_j$.

Proof. Consider the class $\mathcal{S}$ of solutions of equation $\sum_{i \in I} X_i = \sum_{j \in J} X_j$ in $\mathcal{P}(\mathbb{N})$. Then $\mathcal{S} = \mathcal{Z} \cup (\mathcal{S} \setminus \mathcal{Z})$ where

(i) $\mathcal{Z}$ is the class of n-tuples of sets satisfying $X_i = X_j = \emptyset$ for some $i \in I$ and $j \in J$.

(ii) $\mathcal{S} \setminus \mathcal{Z}$ is the class of n-tuples in $\mathcal{S}$ consisting of nonempty sets.

Since $a + \emptyset = \emptyset$ for all $a \in \mathbb{N}$, we have

$$\begin{aligned} \mathcal{Z} \;=\; \{(a_1 + A_1, \ldots, a_n + A_n) \mid a_1, \ldots, a_n \in \mathbb{N},\ A_1, \ldots, A_n \in \mathcal{P}_{0,\emptyset}(\mathbb{N}), \\ A_i = A_j = \emptyset \text{ for some } i \in I \text{ and } j \in J\}\ . \end{aligned} \quad (8) \quad \texttt{\{eq:Z\}}$$

Let $\mathcal{S}_0 = \mathcal{S} \cap (\mathcal{P}_0(\mathbb{N}))^n$ and $\mathcal{R} = \{(a_1, \ldots, a_n) \in \mathbb{N}^n \mid \sum_{i \in I} a_i = \sum_{j \in J} a_j\}$. Observe that if $B_1, \ldots, B_n \in \mathcal{P}(\mathbb{N})$ are all nonempty and $a_i = \min B_i$, $A_i = B_i - a_i$ (i.e. $B_i = a_i + A_i$ with $A_i \in \mathcal{P}_0(\mathbb{N})$) for $i = 1, \ldots, n$, then equality $\sum_{i \in I} B_i = \sum_{j \in J} B_j$ holds if and only if both equalities $\sum_{i \in I} b_i = \sum_{j \in J} b_j$ and $\sum_{i \in I} A_i = \sum_{j \in J} A_j$ hold. Thus,

$$\mathcal{S} \setminus \mathcal{Z} \;=\; \{(a_1 + A_1, \ldots, a_n + A_n) \mid \overrightarrow{a} \in \mathcal{R},\ \overrightarrow{A} \in \mathcal{S}_0\}\ . \quad (9) \quad \texttt{\{eq:S-Z\}}$$

Consider the following formulas (recall $\mathcal{E}(X)$ expresses $X = \emptyset$):

$$\begin{aligned} \psi_{\mathcal{Z}}(x_1, \ldots, x_n, X_1, \ldots, X_n) &\equiv \bigvee_{i \in I, j \in J} \mathcal{E}(X_i) \wedge \mathcal{E}(X_j) \\ \psi_{\mathcal{S} \setminus \mathcal{Z}}(x_1, \ldots, x_n, X_1, \ldots, X_n) &\equiv \neg\mathcal{E}(X_1) \wedge \ldots \wedge \neg\mathcal{E}(X_n) \\ &\quad \wedge \sum_{i \in I} x_i = \sum_{j \in J} x_j \wedge \sum_{i \in I} X_i = \sum_{j \in J} X_j \end{aligned}$$

Equalities (8) and (9) show that, for $a_1, \ldots, a_n \in \mathbb{N}$ and $A_1, \ldots, A_n \in \mathcal{P}_{0,\emptyset}(\mathbb{N})$,

$$\begin{aligned} &\langle \mathcal{P}(\mathbb{N}); +, = \rangle \models \sum_{i \in I} a_i + A_i = \sum_{j \in J} a_j + A_j \\ &\iff \mathcal{M} \models \psi_{\mathcal{Z}}(a_1, \ldots, a_n, A_1, \ldots, A_n) \vee \psi_{\mathcal{S} \setminus \mathcal{Z}}(a_1, \ldots, a_n, A_1, \ldots, A_n) \end{aligned}$$

□

We now lift Lemma 5.2 to formulas with quantifications over $\mathcal{P}(\mathbb{N})$.

{l:lift split}

Lemma 5.3. *For any Boolean combination $F(\overrightarrow{Y}, \overrightarrow{X})$ of equalities between sums of the set variables $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_k$, there exists a Boolean combination $\mathcal{T}(F)$ of*

1. *equalities between sums of the set variables $U_1, \ldots, U_n, V_1, \ldots, V_k$,*
2. *equalities between sums of the integer variables*
3. *formulas $\mathcal{E}(X_1), \ldots, \mathcal{E}(X_n), \mathcal{E}(Y_1), \ldots, \mathcal{E}(Y_k)$,*

such that, for all integers $a_1, \ldots, a_n \in \mathbb{N}$, all sets $A_1, \ldots, A_n$ in $\mathcal{P}_{0,\emptyset}(\mathbb{N})$ (i.e. each A_i is either empty or contains 0), any sequence $Q_1, \ldots, Q_k$ of quantifiers $\exists$ or $\forall$,

$$\begin{aligned} &\langle \mathcal{P}(\mathbb{N}); +, = \rangle \models Q_1 Y_1 \cdots Q_k Y_k \ F(\overrightarrow{Y}, a_1 + A_1, \ldots, a_n + A_n) \\ &\qquad \iff \mathcal{M} \models Q_1 v_1 \ Q_1 V_1 \cdots Q_k v_k \ Q_k V_k \ \mathcal{T}(F)(\overrightarrow{v}, \overrightarrow{a}, \overrightarrow{V}, \overrightarrow{A}) \end{aligned} \tag{10}$$ {eq:TF}

Proof. If E is an equation then Lemma 5.2 gives a formula ψ_E which is a convenient $\mathcal{T}(E)$. For a Boolean combination F of equations $E_1, \ldots, E_p$, let $T(F)$ be the same Boolean combination with $\psi_{E_1}, \ldots, \psi_{E_p}$.

Having defined $\mathcal{T}(F)$, we now prove the Lemma by induction on k. The case $k = 0$ (i.e. no prefix of quantifications) is clear from Lemma 5.2. Suppose (10) holds for the quantification prefix $Q_1 \ldots Q_k$ and any Boolean combination F. Then, for $a_1, \ldots, a_n \in \mathbb{N}$, $A_1, \ldots, A_n$ in $\mathcal{P}_{0,\emptyset}(\mathbb{N})$ and $Q \in \{\exists, \forall\}$,

$$\begin{aligned} &\langle \mathcal{P}(\mathbb{N}); +, = \rangle \models QZ \ Q_1 Y_1 \cdots Q_k Y_k \ F(\overrightarrow{Y}, Z, a_1 + A_1, \ldots, a_n + A_n) \\ &\iff Qb \in \mathbb{N} \ QB \in \mathcal{P}_{0,\emptyset}(\mathbb{N}) \\ &\qquad \langle \mathcal{P}(\mathbb{N}); +, = \rangle \models \overrightarrow{QY} \ F(\overrightarrow{Y}, b + B, a_1 + A_1, \ldots, a_n + A_n) \\ (\dagger) \quad &\iff Qb \in \mathbb{N} \ QB \in \mathcal{P}_{0,\emptyset}(\mathbb{N}) \\ &\qquad \mathcal{M} \models Q_1 v_1 \ Q_1 V_1 \cdots Q_k v_k \ Q_k V_k \ \mathcal{T}(F)(\overrightarrow{v}, b, \overrightarrow{a}, \overrightarrow{V}, B, \overrightarrow{A}) \\ &\iff \mathcal{M} \models Qw \ QW \ Q_1 v_1 \ Q_1 V_1 \cdots Q_k v_k \ Q_k V_k \\ &\qquad\qquad \mathcal{T}(F)(\overrightarrow{v}, w, \overrightarrow{a}, \overrightarrow{V}, W, \overrightarrow{A}) \end{aligned}$$

where line ($\dagger$) is obtained using the induction hypothesis. □

{l:formula-splitting}

Lemma 5.4 (Splitting lemma)**.** *Consider two disjoint sets of variables $z_1, \ldots, z_k, Z_1, \ldots, Z_k$ and $t_1, \ldots, t_n, T_1, \ldots, T_n$ and let $\Phi(\overrightarrow{t}, \overrightarrow{T})$ be a formula with $2n$ free variables of the form*

$$\Phi(\overrightarrow{t}, \overrightarrow{T}) \quad \equiv \quad Q_k z_k Q_k Z_k \cdots \ Q_1 z_1 Q_1 Z_1 \ \mathcal{B}(\overrightarrow{z}, \overrightarrow{Z}, \overrightarrow{t}, \overrightarrow{T})$$

where the Q_i's are quantifiers in $\{\exists, \forall\}$ and where $\mathcal{B}(\overrightarrow{z}, \overrightarrow{Z}, \overrightarrow{t}, \overrightarrow{T})$ is a Boolean combination of atomic formulas depending on the variables $\overrightarrow{z}, \overrightarrow{t}$ only and atomic formulas depending on the variables $\overrightarrow{Z}, \overrightarrow{T}$ only. Then there *exists $r \geq 1$ and finitely many formulas ϕ_ℓ, ψ_ℓ, for $\ell = 1, \ldots, r$ each having n variables, such that $\Phi(\overrightarrow{t}, \overrightarrow{T})$ is logically equivalent to*

$$\bigvee_{\ell=1,\ldots,r} \phi_\ell(\overrightarrow{t}) \wedge \psi_\ell(\overrightarrow{T})$$

Proof. We argue by induction on $k \in \mathbb{N}$. The initial case $k = 0$ is an instance of the classical disjunctive normal form. We now show the induction step. Suppose the Lemma is true for $k-1$ with $k \geq 1$. Then there exists $r \geq 1$ and $\phi_{k-1,\ell}(z_k, \overrightarrow{t})$, $\psi_{k-1,\ell}(Z_k, \overrightarrow{T})$, for $\ell = 1, \ldots, r$, such that

$$\begin{aligned} & Q_{k-1}z_{k-1}Q_{k-1}Z_{k-1}\cdots Q_1z_1Q_1Z_1\ \mathcal{B}(z_1, \ldots, z_{k-1}, z_kZ_1, \ldots, Z_{k-1}, Z_k, \overrightarrow{t}, \overrightarrow{T}) \\ & \qquad\qquad \iff \bigvee_{\ell=1,\ldots,r} \phi_{k-1,\ell}(z_k, \overrightarrow{t}) \wedge \psi_{k-1,\ell}(Z_k, \overrightarrow{T}) \end{aligned}$$

Then, in case Q_k is $\exists$,

$$\begin{aligned} & \exists z_k \exists Z_k\ Q_{k-1}z_{k-1}Q_{k-1}Z_{k-1}\cdots Q_1z_1Q_1Z_1\ \mathcal{B}(\overrightarrow{z}, \overrightarrow{Z}, \overrightarrow{t}, \overrightarrow{T}) \\ & \quad \iff \exists z_k \exists Z_k\ \textstyle\bigvee_{\ell=1,\ldots,r} \phi_{k-1,\ell}(z_k, \overrightarrow{t}) \wedge \psi_{k-1,\ell}(Z_k, \overrightarrow{T}) \\ & \quad \iff \textstyle\bigvee_{\ell=1,\ldots,r} \big(\exists z_k \phi_{k-1,\ell}(z_k, \overrightarrow{t})\big) \wedge \big(\exists Z_k \psi_{k-1,\ell}(Z_k, \overrightarrow{T})\big) \\ & \quad \iff \textstyle\bigvee_{\ell=1,\ldots,r} \phi_{k,\ell}(\overrightarrow{t}) \wedge \psi_{k,\ell}(\overrightarrow{T}) \end{aligned}$$

where $\phi_{k,\ell}(\overrightarrow{t})$ is $\exists z_k \phi_{k-1,\ell}(\overrightarrow{t})$ and the same with $\psi_{k,\ell}(\overrightarrow{T})$. Finally, the case Q_k is $\forall$ is treated similarly by first converting from disjunctive to conjunctive form and, after distributing the $\forall$ quantifiers, converting back from conjunctive to disjunctive form. $\square$

We finally come to the wanted nondefinability result.

{thm:plus-0-not-expre

Theorem 5.5. *The predicate $Y = \{0\} \cup X$ is not definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$.*

Proof. By way of contradiction, assume that the predicate $Y = \{0\} \cup X$ can be defined in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$. There exists a Boolean combination $F(\overrightarrow{Z}, X, Y)$ of equalities of sums of the sets Z_i and X, Y such that

$$\langle \mathcal{P}(\mathbb{N}); +, = \rangle \models\ Y = \{0\} \cup X \iff Q_1Z_1 \cdots Q_kZ_k\ F(\overrightarrow{Z}, X, Y) \qquad (11)$$ {eq:0 cup X}

By Lemma 5.3, for all $a, b \in \mathbb{N}$ and sets $A, B \in \mathcal{P}_{0,\emptyset}(\mathbb{N})$,

$$\begin{aligned} & \langle \mathcal{P}(\mathbb{N}); +, = \rangle \models\ A = \mathbb{N}\ \wedge\ b + B = \{0\} \cup (a + A) \\ & \qquad \iff \mathcal{M} \models\ Q_1v_1\ Q_1V_1 \cdots Q_kv_k\ Q_kV_k\ \mathcal{T}(F)(\overrightarrow{v}, \overrightarrow{V}, a, b, A, B) \qquad (12) \end{aligned}$$ {eq:TF in proof}

where $\mathcal{T}(F)(\overrightarrow{v}, \overrightarrow{V}, u, v, U, V)$ is a Boolean combination of formulas with free variables among u, v and formulas with free variables among U, V. By

Lemma 5.4, we get formulas $\varphi_\ell(u,v)$, $\psi_\ell(U,V)$, $\ell = 1,\ldots,L$, such that

$$\langle \mathcal{P}(\mathbb{N}); +, = \rangle \models b + B = \{0\} \cup (a + A) \iff \mathcal{M} \models \bigwedge_{\ell=1,\ldots,L} \varphi_\ell(a,b) \wedge \psi_\ell(A,B) \quad (13)$$

{eq:conjunct of varph

Now, for each $n \geq 1$, we have $0 + S_n = \{0\} \cup (n + \mathbb{N})$ hence there exists ℓ such that $\varphi_\ell(n,0)$ and $\psi_\ell(\mathbb{N}, S_n)$ are true. Since there are finitely many such indices ℓ, there exists two values of n, say $p,q \geq 1$, $p \neq q$, with the same associated ℓ. In particular, $\varphi_\ell(p,0) \wedge \psi_\ell(\mathbb{N}, S_q)$ is true yielding equality $p + \mathbb{N} = 0 + S_q$, contradicting the condition $p \neq q$. □

5.2 Other nondefinable predicates

{ss:nondefinable}

Theorem 5.5 implies the nondefinability of many other predicates. We select some of them in this section. In particular, we compare four ways to code integers by sets in definable classes:

`Final`	=	$\{n + \mathbb{N} \mid n \geq 1\}$	(see Proposition 2.14)
`Single`	=	$\{\{n\} \mid n \geq 1\}$	(see Proposition 2.17)
`Special`	=	$\{\{0\} \cup n + \mathbb{N} \mid n \geq 1\}$	(see Theorem 3.18)
`Periodic`	=	$\{n\mathbb{N} \mid n \geq 1\}$	(see Proposition 3.27)

{def:leadsto}

Definition 5.6. 1. Given predicates $A_1,\ldots,A_k$ and B over $\mathcal{P}(\mathbb{N})$, we say that B is *definable from* $A_1,\ldots,A_k$ in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$ if B is first-order definable in the structure $\langle \mathcal{P}(\mathbb{N}); +, =, A_1,\ldots,A_k \rangle$. We then write $(A_1,\ldots,A_k) \leadsto B$.
2. A and B are *definable from each other* in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$ if $A \leadsto B$ and $B \leadsto A$.

{thm:inclusion-not-ex

Theorem 5.7. *No predicate in Table 1 is definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$. Moreover, any two of them are definable from each other in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$.*

{open:equidefinabilit

Open problem 5.8. *Is there a non definable predicate which is not definable from each other with the predicates in Table 1?*

{table:non-def}

$E(X,Y)$	Inclusion	$X \subseteq Y$
$F(X,Y)$	Adjoin 0	$Y = X \cup \{0\}$
$F_1(X,Y,Z)$	Union	$X \cup Y = Z$
$F_2(X,Y,Z)$	Intersection	$X \cap Y = Z$
$F_3(X,Y)$	Complement	$X = \mathbb{N} \setminus Y$
$F_4(X,Y)$	Star	$Y = X^*$
$G(X,Y)$	Coding	$(X,Y) \in \{(n+\mathbb{N}, S_n) \mid n \geq 1\}$
$G_1(X,Y)$	interchange	$(X,Y) \in \{(\{n\}, S_n) \mid n \geq 1\}$
$G_2(X,Y)$		$(X,Y) \in \{(n+\mathbb{N}, n\mathbb{N}) \mid n \geq 1\}$
$G_3(X,Y)$		$(X,Y) \in \{(\{n\}, n\mathbb{N}) \mid n \geq 1\}$
$H(X,Y)$	Membership	$(X,Y) \in \{(n+\mathbb{N}, B) \mid n \in B\}$
$H_1(X,Y)$		$(X,Y) \in \{(\{n\}, B) \mid n \in B\}$
$H_2(X,Y)$		$(X,Y) \in \{(n\mathbb{N}, B) \mid n \in B\}$
$H_3(X,Y)$		$(X,Y) \in \{(S_n, B) \mid n \in B\}$
$H_4(X,Y)$		$(X,Y) \in \{(A, B) \mid \min A \in B\}$
$H_5(X,Y)$		$(X,Y) \in \{(A, B) \mid \text{the } 2^{\text{d}} \text{ elem. of } A \text{ is in } B\}$
$J(X)$	Semigroup	X is a semigroup (i.e. $X + X \subseteq X$)
$J_1(X)$		$X \in \{an + n\mathbb{N} \mid n \geq 1, a \in \mathbb{N}\}$
$J_2(X)$		$X \in \{n + n\mathbb{N} \mid n \geq 1\}$

Table 1: Some predicates not definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$

Proof of Theorem 5.7. Together with the predicates `Final`, `Single`, `Special` and `Periodic` recalled supra, we also freely use some definable predicates such as $\min X \leq \min Y$ (cf. § 2.3.4) and

$$\begin{aligned} \theta(X,Y) &= \{(A, S_n) \mid (\min A) + n \in A\} && \text{(cf. Lemma 3.4)} \\ \sigma(X,Y) &\equiv X \text{ is a submonoid containing } Y \ \wedge\ 0 \in Y \\ & && \text{(cf. Propositions 2.14, 2.15, 3.1)} \end{aligned}$$

We also freely use definable constants and functions (cf. §2.3.1): $\emptyset$, $\{0\}$, $\mathbb{N}$, $X \mapsto \{\min X\}$ (defined for $X \neq \emptyset$) and

$$\begin{array}{llll} \texttt{Succc} = S_n \mapsto S_{n+1} & \text{for } n \geq 1 & & \text{(cf. Proposition 3.19)} \\ S = n\mathbb{N} \mapsto S_n & \text{for } n \geq 1 & & \text{(cf. Proposition 3.28)} \\ \pi = S_n \mapsto n\mathbb{N} & \text{for } n \geq 1 & & \text{(cf. Proposition 3.28)} \\ S_{\max} = A \mapsto S_{\max A} & \text{for initial segments } A \neq \{0\} & & \text{(cf. Proposition 3.25)} \end{array}$$

The structure of the proof is as follows:

$$E \rightsquigarrow F \rightsquigarrow \begin{cases} F_4 \rightsquigarrow G \rightsquigarrow G_2 \rightsquigarrow G_3 \rightsquigarrow G_1 \rightsquigarrow G & \\ F_3 \rightsquigarrow G & G_1 \rightsquigarrow E \rightsquigarrow H \rightsquigarrow H_1 \rightsquigarrow G_1 \\ F_2 \rightsquigarrow E & (E, G_3) \rightsquigarrow H_2 \rightsquigarrow H_3 \rightsquigarrow G_1 \\ F_1 \rightsquigarrow E \rightsquigarrow J \rightsquigarrow J_1 \rightsquigarrow J_2 \rightsquigarrow G_3 & H_1 \rightsquigarrow H_4 \rightsquigarrow H_5 \rightsquigarrow H_3 \end{cases}$$

$E \rightsquigarrow F$. Use E to express that Y is the smallest set containing X and $\{0\}$.

$F \rightsquigarrow F_i$ for $i = 1, 2, 3$. For $i = 3$, we have to relate, for any n, condition $n \in X$ with condition $n \in Y$. The case $n = 0$ is treated apart whereas the case $n \geq 1$ can be treated with θ relatively to $X \cup \{0\}$ (which is obtained using F). The cases $i = 1, 2$ are similar. Thus,

$$\begin{array}{lcl}
F_3(X,Y) & \Leftrightarrow & (0 \in X \Leftrightarrow 0 \notin Y) \wedge \forall X_0\ \forall Y_0\ \big(F(X,X_0) \wedge F(Y,Y_0) \\
& & \qquad\qquad \Rightarrow \forall T\ (\texttt{Special}(T) \Rightarrow (\theta(X_0,T) \Leftrightarrow \neg\theta(Y_0,T)))\big) \\
F_1(X,Y,Z) & \Leftrightarrow & (0 \in Z \Leftrightarrow (0 \in X \vee 0 \in Y)) \\
& & \wedge \forall X_0\ \forall Y_0\ \forall Z_0\ \big(F(X,X_0) \wedge F(Y,Y_0) \\
& & \Rightarrow \forall T\ (\texttt{Special}(T) \Rightarrow (\theta(Z_0,T) \Leftrightarrow (\theta(X_0,T) \vee \theta(Y_0,T)))\big) \\
F_2(X,Y,Z) & \Leftrightarrow & (0 \in Z \Leftrightarrow (0 \in X \wedge 0 \in Y)) \\
& & \wedge \forall X_0\ \forall Y_0\ \forall Z_0\ \big(F(X,X_0) \wedge F(Y,Y_0) \\
& & \Rightarrow \forall T\ (\texttt{Special}(T) \Rightarrow (\theta(Z_0,T) \Leftrightarrow (\theta(X_0,T) \wedge \theta(Y_0,T)))\big)
\end{array}$$

$F \rightsquigarrow F_4$. Express that X^* is the smallest submonoid containing $X \cup \{0\}$.

$$F_4(X,Y) \quad \Leftrightarrow \quad \exists X_0\ (F(X,X_0) \wedge \sigma(Y,X_0) \wedge \forall T\ (\sigma(T,X_0) \Leftrightarrow \sigma(T,Y)))$$

$F_1 \rightsquigarrow E$, $F_2 \rightsquigarrow E$. $E(X,Y) \Leftrightarrow F_1(X,Y,Y) \Leftrightarrow F_2(X,Y,X)$.

$F_3 \rightsquigarrow G$. Observe that if X is a final segment then its complement is an initial segment and apply Proposition 3.25:

$$\begin{array}{lcl}
G(X,Y) & \Leftrightarrow & \texttt{Final}(X) \wedge (X = 1 + \mathbb{N} \Rightarrow Y = \mathbb{N}) \\
& & \qquad \wedge\ (X \neq 1 + \mathbb{N} \Rightarrow \exists Z\ (F_3(X,Z) \wedge Y = \texttt{Succ}(S_{\max}(Z))))
\end{array}$$

$F_4 \rightsquigarrow G$. $G(X,Y) \Leftrightarrow \texttt{Final}(X) \wedge F_4(X,Y)$.

$G \rightsquigarrow G_2$. $G_2(X,Y) \Leftrightarrow \texttt{Periodic}(Y) \wedge G(X, S(Y))$.

$G_2 \rightsquigarrow G_3$. $G_3(X,Y) \Leftrightarrow \texttt{Single}(X) \wedge G_2(X + \mathbb{N}, Y)$.

$G_3 \rightsquigarrow G_1$. $G_1(X,Y) \Leftrightarrow \texttt{Periodic}(Y) \wedge \texttt{Single}(X) \wedge G_3(X, S(Y))$

$G_1 \rightsquigarrow G$. $G(X,Y) \Leftrightarrow \texttt{Final}(X) \wedge G_1(\min(X), Y)$.

$G_1 \rightsquigarrow E$. Observe that $X \subseteq Y$ if and only if three conditions hold: 1) $\min X \geq \min Y$ 2) $\min X \in Y$ and 3) for all $p, q \geq 1$, if $(\min X) + p = (\min Y) + q$ then $(\min X) + p \in X$ implies $(\min Y) + q \in Y$. Now, $\min X \in Y$ if and only if $\min X = \min Y$ or $(\min Y) + a \in Y$ with $a = \min X - \min Y \geq 1$. In the next formula $\widetilde{X}, \widetilde{Y}$ denote $S_{\min X}, S_{\min Y}$ and P denotes S_p ($p > 0$) and Q denotes S_q (with $q = p + a$ and $p \geq 0$). Also line 3 expresses $\min X \in Y$ (in the sole necessary case where $\min X > \min Y$) and line 4 expresses that if $p,q \geq 1$ and $\min X + p = \min Y + q$ holds then $\min X + p \in X$ implies $\min Y + q \in Y$.

$$
\begin{aligned}
E(X,Y) \quad \Leftrightarrow \quad & \min X \geq \min Y \wedge\ \exists \widetilde{X}\ \exists \widetilde{Y}\ \forall P\ \forall Q \\
& \big(G_1(\{\min X\}, \widetilde{X}) \wedge G_1(\{\min Y\}, \widetilde{Y}) \wedge\ \texttt{Special}(P) \wedge\ \texttt{Special}(Q)\big) \\
& \Rightarrow \big(((\min X > \min Y \wedge \widetilde{X} = \widetilde{Y} \oplus Q) \Rightarrow \theta(Y,Q))) \\
& \quad \wedge \widetilde{X} \oplus P = \widetilde{Y} \oplus Q \Rightarrow (\theta(X,P) \Rightarrow \theta(Y,Q)\big)
\end{aligned}
$$

$E \rightsquigarrow H$. $H(X,Y) \Leftrightarrow \texttt{Final}(X) \wedge E(\{\min X\}, Y)$.

$H \rightsquigarrow H_1$. $H_1(X,Y) \Leftrightarrow \texttt{Single}(X) \wedge H(X + \mathbb{N}, Y)$.

$H_1 \rightsquigarrow G_1$. $G_1(X,Y) \Leftrightarrow \texttt{Special}(Y) \wedge H_1(X,Y) \wedge \neg H_1(X, \texttt{Succ}(Y))$.

$(E, G_3) \rightsquigarrow H_2$. $H_2(X,Y) \Leftrightarrow \texttt{Periodic}(X) \wedge \exists Z\ (G_3(Z,X) \wedge E(Z,Y))$.

$H_2 \rightsquigarrow H_3$. $H_3(X,Y) \Leftrightarrow \texttt{Special}(X) \wedge H_2(\pi(X), Y)$.

$H_3 \rightsquigarrow G_1$. $G_1(X,Y) \Leftrightarrow \texttt{Single}(X) \wedge \texttt{Special}(Y) \wedge H_3(Y,X) \wedge \neg H_3(\texttt{Succ}(Y), X)$.

$H_1 \rightsquigarrow H_4$. $H_4(X,Y) \Leftrightarrow H_1(X + \mathbb{N}, Y)$.

$$
\begin{aligned}
H_4 \rightsquigarrow H_5.\ H_5(X,Y) \Leftrightarrow\ & X \neq \emptyset \wedge \neg\texttt{Single}(X) \\
& \wedge \exists Z\ \big(\texttt{Single}(Z) \wedge (\min Z > \min X) \wedge H_4(Z,X) \wedge H_4(Z,Y) \\
& \wedge \forall T\ ((\texttt{Single}(T) \wedge\ \min X < \min T < \min Z) \Rightarrow \neg H_4(T,X))\big).
\end{aligned}
$$

$H_5 \rightsquigarrow H_3$. $H_3(X,Y) \Leftrightarrow \texttt{Special}(X) \wedge H_5(X,Y)$.

$E \rightsquigarrow J$. Observe that X is a semigroup if and only if $X + X \subseteq X$.

$J \rightsquigarrow J_1$. Let $\mathcal{T} = \{a + n\mathbb{N} \mid a \in \mathbb{N}, n \geq 1\}$. Then $\mathcal{T}(X)$ if and only if $X = Y + Z$ for some Y, Z satisfying $\texttt{Single}(Y)$ and $\texttt{Periodic}(Z)$. Now, $J_1(X)$ if and only if $\mathcal{T}(X)$ and X is a semigroup: implication $\Rightarrow$ is obvious, conversely, if $X = a + n\mathbb{N}$ is a semigroup then its ultimate period is n. It divides all elements of X and in particular a and $J_1(X)$ holds.

$J_1 \rightsquigarrow J_2$. Observe that $J_2(X)$ if and only if $J_1(X)$ and $\min X \neq 0$ and $\{\min X\}$ is the unique singleton set such that $X = T + Y$ for some T, Y such that $T \neq \{0\}$, $\texttt{Single}(T)$ and $J_1(Y)$.

$J_2 \rightsquigarrow G_3$. $G_3(X,Y) \Leftrightarrow \exists Z\ (J_2(Z) \wedge X = \{\min Z\} \wedge Z = X + Y)$. $\square$

Out of the six pairs of the four codings of $\mathbb{N} \setminus \{0\}$ by `Final`, `Single`, `Special` and `Periodic`, Table 5.2 tells that four conversions are incomparable with respect to $\rightsquigarrow$ (cf. predicates G and G_i, $i = 1, 2, 3$). In contrast the two remaining ones are definable as shown in the next result. {p:contrast definable

Proposition 5.9. *The following predicates are respectively* Π_2 *and* Π_3.

$$(X,Y) \in \{(n + \mathbb{N}, \{n\}) \mid n \geq 1\} \quad , \quad (X,Y) \in \{(n\mathbb{N}, S_n) \mid n \geq 1\}$$

Proof. Observe that $(X,Y) \in \{(n+\mathbb{N}, \{n\}) \mid n \geq 1\}$ if and only if $0 \notin Y$ and Y is a singleton set and $Y + \mathbb{N} = X$. Expressing the last equality as $\forall Z\ (Z = \mathbb{N} \Rightarrow Y + Z = X)$ Propositions 2.14 and 2.17 give the Π_2 complexity.

For the second predicate see Proposition 3.28. $\square$

6 Logical definability in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$

{s:definable-subsets-

6.1 Families of sets all containing 0

{ss:sets-containing-0

A simple application of Theorem 4.1 proves the following result.

{thm:sets-containing-

Theorem 6.1. *Suppose $\mathcal{F} \subseteq \mathcal{P}(\mathbb{N})$ is a class of sets all containing 0. Then $\mathcal{F}$ is definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$ if and only if it is definable in second-order arithmetic.*

Proof. Observe that $\mathcal{F}$ is definable in second-order arithmetic if and only if so is $\{A \mid \{0\} \cup (1 + A) \in \mathcal{F}\}$ and apply Theorem 4.1. □

{cor:sets-containing-

Corollary 6.2. *Suppose $\mathcal{F} \subseteq \mathcal{P}(\mathbb{N})$ is a class of sets all containing an integer in $\{0, \ldots, n\}$. Then $\mathcal{F}$ is definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$ if and only if it is definable in second-order arithmetic.*

Proof. The $\Rightarrow$ implication is trivial. Conversely, suppose $\mathcal{F}$ is definable in second-order arithmetic. For $i = 1, \ldots, n$, let $\mathcal{F}_i = \mathcal{F} \cap \{X \mid \min X = i\}$ and $\mathcal{G}_i = \{X - \min X \mid X \in \mathcal{F}_i\}$. Then the formulas $\mathcal{G}_i$ are also definable in second-order arithmetic. Since all sets in the formulas $\mathcal{G}_i$ contain 0, these formulas $\mathcal{G}_i$ are definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$ (use Theorem 6.1). Then so are the sets $\{i + X \mid X \in \mathcal{G}_i\} = \mathcal{F}_i$ hence also their union which is $\mathcal{F}$. □

6.2 Families of sets invariant by translation

{ss:sets-invariant-tr

There is yet another application of Theorem 4.1 which extends the class of subsets definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$.

{thm:translation-inva

Theorem 6.3. *Suppose $\mathcal{F} \subseteq \mathcal{P}(\mathbb{N})$ is a class of subsets such that for all subsets $A \subseteq \mathbb{N}$ and all integers $a \in \mathbb{N}$ it holds*

$$A \in \mathcal{F} \Leftrightarrow A + a \in \mathcal{F}$$

Then $\mathcal{F}$ is definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$ if and only if it is definable in second-order arithmetic.

Proof. Let $\mathcal{F}_0$ be the subclass of subsets in $\mathcal{F}$ containing 0. For all $B \in \mathcal{F}$ the subset $B - \min B$ is in $\mathcal{F}_0$. Clearly $\mathcal{F}_0$ is definable in second-order arithmetic, thus in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$ by a formula $\phi(X)$. Then $\mathcal{F}$ is definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$ by the formula

$$\exists X \ \exists Y \ (\phi(X) \wedge \mathrm{Sing}(Y) \wedge Z = Y + X)$$

□

{cor:sets-invariant-b

Corollary 6.4. *The following classes of subsets are definable*

(i) $\{A \subseteq \mathbb{N} \mid A$ *is finite*$\}$,

(ii) $\{A \subseteq \mathbb{N} \mid A$ *is cofinite*$\}$,

(iii) $\{A \subseteq \mathbb{N} \mid A$ *is regular by a finite automaton*$\}$*).*

Proof. Assertions (i) and (ii) are clear. Concerning assertion (iii) recall (cf. Proposition 2.6) that a subset $A \subseteq \mathbb{N}$ is recognizable by a finite automaton if and only if it is a finite union of subsets of the form $a + b\mathbb{N}$ with $a, b \in \mathbb{N}$. Thus, this class satisfies the conditions of Theorem 6.3. $\square$

6.3 Definable sets of integers

{thm:definable-sets}

Theorem 6.5. *Let $A \in \mathcal{P}(\mathbb{N})$. The following conditions are equivalent:*

(1) *As a set of integers, A is definable in second-order arithmetic,*

(2) *As a class of sets, $\{A\}$ is definable in second-order arithmetic,*

(3) *$\{A\}$ is definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$.*

Proof. (1) $\Leftrightarrow$ (2) is straightforward. (3) $\Rightarrow$ (2) is obvious.
(2) $\Rightarrow$ (3). Let $a = \min A$. If $\{A\}$ is definable in second-order arithmetic then so is $\{A - a\}$. By Theorem 6.1 $\{A - a\}$ is definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$ (since $0 \in A - a$) hence so is $\{A\} = \{X + a \mid X \in \{A - a\}\}$. $\square$

6.4 Definability with an extra predicate

{thm:extra}

Theorem 6.6. *Let $\mathcal{F} \subseteq \mathcal{P}(\mathbb{N})$ and A be any predicate in Table 1. Then $\mathcal{F}$ is definable in second-order arithmetic if and only if $\mathcal{F}$ is definable in the structure $\langle \mathcal{P}(\mathbb{N}); +, A, = \rangle$.*

Proof. Implication $\Leftarrow$ is obvious since all predicates in Table 1 are definable in second-order arithmetic. Conversely, suppose $\mathcal{F}$ is definable in second-order arithmetic. Then so are the classes

$$\mathcal{F}^+ = \mathcal{F} \cap \{A \mid 0 \in A\}\ ,\ \mathcal{F}^- = \mathcal{F} \cap \{A \mid 0 \notin A\}\ ,\ \mathcal{H} = \{\{0\} \cup A \mid A \in \mathcal{F}^-\}$$

By Theorem 6.1 $\mathcal{F}^+$ and $\mathcal{H}$ are definable in $\langle \mathcal{P}(\mathbb{N}); +, = \rangle$. Using the predicate $F(X, Y)$ (which insures $Y = \{0\} \cup X$) one can then define $\mathcal{F}^-$ from $\mathcal{H}$. Finally, $\mathcal{F} = \mathcal{F}^+ \cup \mathcal{F}^-$ is definable in $\langle \mathcal{P}(\mathbb{N}); +, F, = \rangle$. Since all predicates in Table 1 are definable from each other, $\mathcal{F} = \mathcal{F}^+ \cup \mathcal{F}^-$ is also definable in $\langle \mathcal{P}(\mathbb{N}); +, A, = \rangle$ for any A in Table 1. $\square$

7 Remarkable definable sets and classes

{s:with-Sn}

7.1 Operations on sets with close minimum elements

{ss:operations-on-sub

General set-theoretical operations are not definable. Here we show sufficient conditions for some of these operations to be definable.

{p:same-min}

Proposition 7.1. *The predicate* $\min X = \min Y \ \wedge\ \phi(X,Y,Z)$ *is* Π_4 *when* $\phi(X,Y,Z)$ *is either* $X \cup Y = Z$ *or* $X \cap Y = Z$ *or* $X \subseteq Y$*.*

Proof. We argue for union. The $\min X = \min Y$ condition is Σ_1 (cf. Proposition 2.16) and, by Lemma 3.4, $X \cup Y = Z$ if and only if, for all $n \geq 1$,

$$Z + S_n = Z + S_{n+1} \iff ((X + S_n = X + S_{n+1}) \vee (Y + S_n = Y + S_{n+1}))$$

Theorem 3.18 and Proposition 3.19 yield the stated logical complexity. □

{p:almost-same-min}

Proposition 7.2. *1. The predicate* $|\min X - \min Y| \leq k \ \wedge\ \phi(X,Y,Z)$ *is* Π_4 *when the integer* $k \geq 1$ *is fixed and* $\phi(X,Y,Z)$ *is either* $X \cup Y = Z$ *or* $X = \mathbb{N} \setminus Y$ *or* $X \subseteq Y$*.*
2. The following predicate is Π_4 *when the integer* $k \in \mathbb{N}$ *is fixed:*

$$\max(|\min(X) - \min(Y)|, |\min X - \min Z|) \leq k \ \wedge\ X \cap Y = Z$$

Proof. Point 1. Since the condition $|\min X - \min Y| \leq k$ is a disjunction of conditions $\min X = \min Y + \ell$ with $|\ell| \leq k$, it suffices to prove that the predicate $\min X = \min Y + k \ \wedge\ \phi(X,Y,Z)$ is Π_5. We argue for union. The $\min(X) = \min(Y) + k$ condition is $\Sigma_2 \wedge \Pi_2$ (cf. Proposition 2.16). Assuming $\min(X) = \min(Y) + k$, Lemma 3.4 insures that $X \cup Y = Z$ if and only if

1) $\min Z = \min Y$,
2) $Z + S_n = Z + S_{n+1} \iff Y + S_n = Y + S_{n+1}$ for all $n \in \{1, \ldots, k-1\}$,
3) $Z + S_k = Z + S_{k+1}$,
4) for all $n \geq 1$,
$$Z + S_{k+n} = Z + S_{k+n+1} \iff (X + S_n = Z + S_{n+1} \ \vee\ Y + S_{k+n} = Y + S_{k+n+1}).$$

Theorems 3.6, 3.18 and Proposition 3.19 yield the stated logical complexity. The proof of Point 2 is similar. □

Remark 7.3. Observe that the intersection is not definable even when the minima of the two subsets differ by 1 : we have to also bound the difference between $\min(X \cap Y)$ and $\min X$. For instance, $\{(n + \mathbb{N}, S_n) \mid n \geq 2\}$ which is $G(X,Y)$ of Theorem 5.7 is expressed as follows

$$\exists Z\ (\texttt{Special}(Z) \wedge Y = \partial Z \wedge (X = Y \cap (Z + 1)))$$

Indeed, if $Z = S_n$ then Y and $1 + Z$ have close minimum elements since $\min(1 + Z) = 1$ and $\min Y = 0$ but $\min(Y \cap (1 + Z)) = n + 1$ is not close to $\min Y$.

7.2 Fixed submonoids

{ss:fixed submonoids}

Proposition 2.16 can be generalized as follows.

{l:n-first}

Lemma 7.4. *Let X be a subset with minimum element equal to m and let $a_1 < \ldots < a_n$ be elements of $\mathbb{N}$. The following predicates $\phi_m(X)$ and $\psi_m(X)$ are Δ_2 :*

$$\begin{array}{ll} \phi_m(X) & : \min X = m \textit{ and } a_1, \ldots, a_n > m \textit{ belong to } X \\ \psi_m(X) & : \min X = m \textit{ and } a_1, \ldots, a_n > m \textit{ do not belong to } X \end{array}$$

Proof. Using Lemma 3.4, we have:

$$\phi_m(X) \equiv \min X = m \wedge \bigwedge_{1 \le i \le n} X + S_{a_i - m} = X + S_{a_i - m + 1}$$
$$\psi_m(X) \equiv \min X = m \wedge \bigwedge_{1 \le i \le n} X + S_{a_i - m} \neq X + S_{a_i - m + 1}$$

The Δ_2 complexity is a corollary of Theorem 3.6 and Proposition 2.13. □

{thm:each submonoid}

Theorem 7.5. *Let M be a submonoid. The predicate $X = M$ is definable and its complexity is as follows*
1. Case $M = \{0\}$. The predicate $X = M$ is Π_1.
2. Case $M = \mathbb{N}$. The predicate $X = M$ is $\Sigma_1 \wedge \Pi_1$.
3. Case $M = \{0\} \cup (a + \mathbb{N}) = S_a$ with $a \geq 2$. The predicate $X = M$ is Δ_2.
4. For the general case the predicate $X = M$ is Π_2.

Proof. Claims 1, 2: cf. Proposition 2.14. Claim 3: cf. Theorem 3.6.
4. Let $G = \{g_1, \ldots, g_n\}$ be the minimum generating set for M which we assume ordered. Then M is the smallest submonoid containing G. If $\phi_0(X)$ is the formula of Lemma 7.4 with $g_1, \ldots, g_n$ in place of $a_1, \ldots, a_n$ the predicate $X = M$ is equivalent to

$$\phi_0(X) \wedge \ X + X = X \wedge \forall Y \ \big((\phi_0(Y) \wedge (Y + Y = Y)) \Rightarrow X + Y = Y\big)$$

expressing that X is the smallest submonoid containing G. The Π_2 complexity comes from Lemma 7.4. □

7.3 Regular subsets of $\mathbb{N}$

{ss:regular-subsets-i

Here we give a precise estimate of the structural complexity of some subsets and classes of subsets of particular importance, the definability of which is a consequence of Theorem 6.3.

7.3.1 Fixed regular subsets of $\mathbb{N}$

{thm:fixedregular}

Theorem 7.6. *If $R \subseteq \mathbb{N}$ is regular then the predicate $X = R$ is Π_4. In case $R = \emptyset$ or $R = \{0\}$ it is Π_1. In case R is a singleton different from 0 it is $\Sigma_2 \wedge \Pi_2$.*

Proof. By Proposition 2.6, $R = A \cup (B + p\mathbb{N})$ with $a, p \in \mathbb{N}$ and $\emptyset \neq A \subseteq [0, a[$ and $B \subseteq [a, a+p[$. First, we introduce some formulas. We set $m = \min A$ and $b = a + p$ with the convention $b = a$ when B is empty. The following Σ_4 and Π_4 predicates tell which elements in the initial interval $[0, b[$ belong to the set and which do not.

$$
\begin{aligned}
F^{\exists}(X) \quad & \equiv m = \min X \ \wedge \ \exists Y_1, \ldots, Y_{b-m-1} Z_1, \ldots, Z_{b-m-1} \\
& \bigwedge_{i \in A \cup B \setminus \{m\}} (Y_i = S_{i-m} \ \wedge \ \mathtt{Succ}(Y_i, Z_i) \ \wedge \ X + Y_i = X + Z_i) \\
& \wedge \bigwedge_{i \in \{m+1, \ldots, b-1\} \setminus (A \cup B)} (Y_i = S_{i-m} \ \wedge \ \mathtt{Succ}(Y_i, Z_i) \ \wedge \ X + Y_i \neq X + Z_i)
\end{aligned}
$$

$$
\begin{aligned}
F^{\forall}(X) \quad & \equiv m = \min X \ \wedge \ \forall Y, Z \\
& \bigwedge_{i \in A \cup B \setminus \{m\}} ((Y = S_{i-m} \ \wedge \ \mathtt{Succ}(Y, Z)) \Rightarrow X + Y = X + Z) \\
& \wedge \bigwedge_{i \in \{m+1, \ldots, b-1\} \setminus (A \cup B)} ((Y = S_{i-m} \ \wedge \ \mathtt{Succ}(Y, Z)) \Rightarrow X + Y \neq X + Z)
\end{aligned}
$$

If the subset is finite, it suffices to express the fact that it does not contain any integer greater than or equal to a. This leads to the Π_4 formula

$$
\begin{aligned}
G(X) \quad & \equiv \forall Y, Z, T \ \big((T = S_{a-m} \wedge \ \mathtt{Succ}(Y, Z) \ \wedge \ Y + T = T) \\
& \Rightarrow X + Y \neq X + Z\big)
\end{aligned}
$$

Thus when R is finite it is expressed by the predicate $F^{\forall}(X) \wedge G(X)$.

When the subset is infinite, i.e., when $B \neq \emptyset$ we must say that the subset of R consisting of all elements greater than or equal to a is periodic of period p, equivalently for all $x \geq a$ we have $x \in R \Leftrightarrow x + p \in R$ which is expressed by the Π_4 formula

$$
\begin{aligned}
H(X) \quad & \equiv \ \forall T, Y, Y', Z, Z' : \\
& T = S_{a-m} \wedge \mathtt{Succ}(Y, Y') \wedge \mathtt{Succ}(Z, Z') \wedge \mathtt{Succ}_p(Y, Z) \\
& \Rightarrow (X + Y = X + Y' \Leftrightarrow X + Z = X + Z')
\end{aligned}
$$

Thus, when R is infinite it is expressed by the formula $F^{\forall}(X) \wedge H(X)$.

The remaining cases are a consequence of Proposition 2.17.

$\square$

7.3.2 Finite and cofinite subsets of $\mathbb{N}$

{ss:finitecofinite}

Proposition 7.7. *The predicates "X is finite" and X is cofinite" are* Σ_5.

Proof. For the finiteness predicate, use Lemma 3.4 to express that $\min(X)+n \notin X$ for all large enough n :

$$\exists Y\ \big(\texttt{Special}(Y) \wedge \forall Z, W\ ((Y+Z=Y \wedge \texttt{Succ}(Z,W)) \Rightarrow X+Z \neq X+W\big)$$

For the cofiniteness predicate, express that $\min(X)+n \in X$ for all large enough n. The logical complexity is given by Theorem 3.18 and Proposition 3.19. □

7.3.3 The class of regular subsets of $\mathbb{N}$

{ss:regularclass}

Theorem 7.8. *The predicate "X is regular" is* Σ_6.

Proof. Observe that X is regular if and only if X is finite or is periodic with period $p \geq 1$. This latter assertion means that there exist two integers n and p such that all for all integers $x \geq n$ we have $x \in X$ if and only if $x+p \in X$. Using Theorems 3.26 and Proposition 7.7 this can be expressed as follows (where the variables N, P encode the above integers n and p and where the pairs of variables (V, V') and (W, W') respectively encode x and $x+p$):

$$\begin{aligned} &X \text{ is finite } \vee\ \exists N\ \exists P\ \forall V, W, V', W' \\ &\quad \big(\texttt{Special}(N) \ \wedge\ \texttt{Special}(P) \ \wedge\ N+V=N \ \wedge\ V \oplus P = W \\ &\qquad\qquad \wedge\ \texttt{Succ}(V,V') \ \wedge\ \texttt{Succ}(W,W') \\ &\qquad\qquad\qquad \Longrightarrow (X+V=X+V' \Leftrightarrow X+W=X+W')\big) \end{aligned}$$

□

8 Conclusion

{s:conclusion}

This paper proves the undecidability of the Σ_5 theory of additive monoid of subsets of $\mathbb{N}$. The decision problem for positive Σ_1 formulas (i.e. no negation) is trivially decidable since every equation is satisfied when all the variables are equal to the emptyset. What about the full Σ_1 theory? Care: we are looking at formulas in which the atomic subformulas are equations between variables (such as $XYXZ = ZXX$): no parameter is allowed. When regular sets are allowed as parameters then the decision problem for systems of equations becomes undecidable, cf. [6].

The question "What is definable and what is not definable in the additive monoid of subsets of $\mathbb{N}$?" is largely answered in this paper. Some definability results involve logically complex definitions: up to Σ_6 for the class of regular sets. Are such complex definitions optimal?

The additive monoid of subsets of $\mathbb{N}$ can be seen as the monoid of tally languages. What about the monoid of languages over an alphabet with at least two letters. This question is investigated in a forthcoming paper [1].

References

[1] Christian Choffrut and Serge Grigorieff. *Logical theory of the monoid of languages over a non tally alphabet.* In preparation.

[2] Leslie Cohn. On the submonoids of the additive group of integers, http://www.macs.citadel.edu/cohnl/submonoids2002.pdf.

[3] Michael Jo Fischer and Michael O. Rabin. Super-exponential complexity of presburger arithmetic. In *SIAM-AMS Symposium in Applied Mathematics*, volume 7, page 2741, 1974.

[4] Pedro A. García-Sánchez. Numerical semigroups minicourse, http://www.ugr.es/ pedro/minicurso-porto.pdf.

[5] Seymour Ginsburg and Edwin H. Spanier. Semigroups, Presburger formulas, and languages. *Pacific J. Math.*, 16:285–296, 1966.

[6] Artur Jeż and Alexander Okhotin. Equations over sets of natural numbers with addition only. In *STACS*, pages 577–588, 2009.

[7] Piergiorgio Odifreddi. *Classical recursion theory. Vol. 1: The theory of functions and sets of natural integers.* North Holland, 1989.

[8] Olivier Ramaré. On Shnirelman's constant. 22(4):645–706, 1995.

[9] Jorge Luis Ramírez-Alfonsín. *The Diophantine Frobenius Problem.* Oxford University Press, 2005.

[10] Hartley Rogers. *Theory of recursive functions and effective computability.* McGraw Hill, 1967.

[11] José Carlos Rosales and Pedro A. García-Sánchez. *Numerical semigroups.* Number 20 in Developments in Mathematics. Springer, 2009.