On deciding linear arithmetic constraints over p-adic integers for all primes

Christoph Haase
Department of Computer Science, University of Oxford, Oxford, UK

Alessio Mansutti
Department of Computer Science, University of Oxford, Oxford, UK

Abstract

Given an existential formula Φ of linear arithmetic over p-adic integers together with valuation constraints, we study the p-universality problem which consists of deciding whether Φ is satisfiable for all primes p, and the analogous problem for the closely related existential theory of Büchi arithmetic. Our main result is a coNEXP upper bound for both problems, together with a matching lower bound for existential Büchi arithmetic. On a technical level, our results are obtained from analysing properties of a certain class of p-automata, finite-state automata whose languages encode sets of tuples of natural numbers.

2012 ACM Subject Classification Theory of computation → Logic

Keywords and phrases linear arithmetic, Büchi arithmetic, p-adic numbers, automatic structures

Funding This work is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 852769, ARiAT).

1 Introduction

In the light of the undecidability of Hilbert's tenth problem, the decidability of the Diophantine problem for addition and divisibility established by Lipshitz [20] is a non-trivial and interesting result. The latter problem consists of deciding whether a system of divisibility constraints of the form $p(\boldsymbol{x}) \mid q(\boldsymbol{x})$, with p and q being linear polynomials, has a solution over the integers. Lipshitz' proof of decidability relies on a local-to-global principle. He showed that every such system can be transformed into an equi-satisfiable one that has a solution if and only if an associated restricted system of linear equations with simple p-adic valuation constraints is satisfiable over the p-adic integers for every prime p. We call the latter problem the *p-universality problem*. To decide p-universality for the restricted class he considered, Lipshitz showed that it suffices to only check satisfiability for all primes p up to a certain threshold that can be computed from the input. One main result of this paper is to show that the latter result can be generalised: p-universality is decidable in coNEXP for *arbitrary* systems of linear equations over p-adic integers together with *general* linear p-adic valuation constraints. For linear equations with first-order variables ranging over the whole field of the p-adic numbers and restricted valuation constraints that allow to impose a partial order on the p-adic valuations of the first-order variables, a quantifier-elimination procedure was given by Dolzmann and Sturm from which it is possible to derive a coNEXP upper bound for p-universality in this setting [8]. Their result also shows that, in their setting, the set of those primes for which a solution exists is either finite or co-finite.

Linear arithmetic over p-adic integers with valuation constraints is closely related to *Büchi arithmetic*. Büchi arithmetic of base $p \geq 2$, p not necessarily prime, is the first-order theory of the structure $(\mathbb{N}, +, =, V_p)$, an extension of Presburger arithmetic with a unary V_p function such that $V_p(a) = b$ if and only if b is the largest power of p dividing a without remainder, i.e., there is some $k \in \mathbb{N}$ such that $b = p^k$, $b \mid a$ and $p \cdot b \nmid a$. Büchi showed that this theory is decidable using an automata-based approach, and conversely that Büchi arithmetic of base p defines the sets of numbers recognisable by *p-automata*, finite-state automata defining tuples of natural numbers encoded as words of tuples over the alphabet $\{0, \dots, p-1\}$ [5], though the latter result was incorrectly stated by Büchi and later correctly stated and proved by Bruyère [3]. One central line of research in Büchi arithmetic has been to understand the properties of this theory when the base p is variable. For instance, the celebrated Cobham-Semënov theorem states that if a set $M \subseteq \mathbb{N}^d$ is separately definable in Büchi arithmetic of multiplicatively independent bases p and q, then M is definable in Presburger arithmetic [7, 26]. Another main result of this paper is to show coNEXP-completeness of the analogue of p-universality for existential Büchi arithmetic: given an existential formula Ψ of Büchi arithmetic, decide whether Ψ is satisfiable in all bases $p \geq 2$. Note that p-universality does not imply definability in Presburger arithmetic as, for instance, the formula $V_p(x) = y$ is not definable in Presburger arithmetic, but it is p-universal.

Both coNEXP upper bounds are obtained by establishing doubly-exponential upper bounds on the smallest p for which a given formula becomes unsatisfiable. As a structural result, we obtain that for linear equations over p-adic integers with valuation constraints, the set of those primes p for which a given instance is satisfiable is precisely contained in an ultimately periodic set. On a technical level, our results are obtained by analysing properties of p-automata. While the latter have been studied for decades, only recently have they been instrumental in obtaining tight complexity bounds for long-standing open problems about the complexity of the satisfiability problem of the existential theories of the two arithmetic theories we consider in this paper [11]. A key observation we exploit for our approach is that the set of states of a p-automaton accepting the solutions of a system of linear Diophantine equations does *not* depend on p. Note that the quantifier-elimination approach employed by Dolzmann and Sturm [8] does not seem applicable in our setting as it works over the whole p-adic numbers and relies on them being a field. Moreover, Büchi arithmetic does not have a quantifier-elimination procedure, even when extended with additional predicates definable in existential Büchi arithmetic [13].

2 Preliminaries and main results

The symbols $\mathbb{Z}$, $\mathbb{N}$ and $\mathbb{Q}$ denote the set of integers, natural and rational numbers, respectively. We write $\mathbb{P}$ for the set of prime numbers, and $\overline{\mathbb{Z}}$ to denote the set of integers extended with the symbol ∞ such that $n \leq \infty$ for all $n \in \mathbb{Z}$. All numbers are assumed to be encoded in binary, unless otherwise stated. For any object, we denote by $\langle \cdot \rangle$ the size of its encoding.

Linear arithmetic constraints over p-adic integers. Let $p \geq 2$ be a fixed prime number. Given a non-zero rational number $q \in \mathbb{Q}$, the p-adic valuation $v_p(q)$ is defined as the unique integer $k \in \mathbb{Z}$ such that $q = p^k \cdot \frac{a}{b}$ for $a, b \in \mathbb{Z}$ not divisible by p, and $v_p(0) = \infty$. The valuation v_p induces the p-adic absolute value $|\cdot|_p$ defined as $|q|_p = p^{-v_p(q)}$. The field of p-adic numbers $\mathbb{Q}_p$ is obtained as the Cauchy completion of the field of the rational numbers under $|\cdot|_p$. Any p-adic number different from 0 has a unique p-adic expansion as an infinite power series $\sum_{i=k}^{\infty} a_i p^i$ for some $k \in \mathbb{Z}$, $a_k \neq 0$ and $a_i \in [0, p-1]$ for all $i \geq k$. The ring $\mathbb{Z}_p$ of p-adic integers consists of all p-adic numbers for which this k is non-negative. By *linear arithmetic constraints over p-adic integers*, we refer to the first-order theory of the two-sorted structure $(\{\mathbb{Z}_p, \overline{\mathbb{Z}}\}, 0, 1, +, =, <, v_p)$. All constants, relational and functional symbols have their natural semantics, and v_p is the p-adic valuation mapping p-adic integers to the valuation ring $\overline{\mathbb{Z}}$. For simplicity, we view the constants 0 and 1 as well as binary addition $+$ as being defined for both sorts. However, addition is restricted between elements of the *same* sort. The equality relation $=$ is defined on both $\overline{\mathbb{Z}}$ and $\mathbb{Z}_p$, whereas the less-than relation $<$ is restricted to the valuation ring $\overline{\mathbb{Z}}$. Usually, the letters u, v refer to first-order variables interpreted over $\mathbb{Z}_p$, and x, y, z refer to variables over $\overline{\mathbb{Z}}$. We rely on the axiom system for integer arithmetic enriched with infinity presented in [17] to treat linear terms over the valuation ring containing the symbol ∞.

Note that we allow arbitrary Boolean combinations of linear inequalities to constraint valuations of the variables from $\mathbb{Z}_p$, whereas Dolzmann and Sturm [8] as well as Lipshitz [20] only allow restricted constraints of the form $v_p(u) \leq v_p(v)$.

Büchi arithmetic. Let $p \geq 2$ be a fixed integer. *Büchi arithmetic* of base p is the first-order theory of the structure $(\mathbb{N}, 0, 1, +, =, V_p)$, where the constants 0 and 1 and the relations $+$ and $=$ are interpreted in their natural semantics, and V_p is the unary function mapping every non-zero integer x to the largest power of p that divides x without remainder as defined in the introduction. For the purpose of this paper, as in [4] we define $V_p(0) = 1$, though other definitions such $V_p(0) = \infty$ are possible, but they do not change the sets of numbers definable in Büchi arithmetic. The decidability of Büchi arithmetic rests on the fact that Büchi arithmetic is an automatic structure in the sense of [14, 16, 2]. While full Büchi arithmetic is TOWER-complete [25], its existential fragment is only NP-complete [11].

Main decision problems and results. Both Büchi arithmetic and linear arithmetic constraints over p-adic integers are defined with respect to a fixed base $p \in \mathbb{N}$. In this paper, we treat p as a parameter, and, for a given formula Φ of *existential* Büchi arithmetic or *existential* linear arithmetic constraints over p-adic integers mentioning p, are interested in the following two decision problems:

- p-EXISTENCE: Is Φ satisfiable for *some* $p \geq 2$?
- p-UNIVERSALITY: Is Φ satisfiable for *every* $p \geq 2$?

When Φ is a formula of linear arithmetic constraints over p-adic integers, p above is additionally restricted to be a prime number. For the complexity of those decision problems, we stipulate that the V_p and v_p functions count as a single symbol in $\langle \Phi \rangle$ for any formula Φ. Note that p-universality and p-existence are not the complement of one and another: the formula $x \neq 2 \lor V_p(x) = 2$ of Büchi arithmetic has a solution for $p = 2$, but its negation is not p-universal. As the main results of this paper, we show:

▶ **Theorem 1.** *For both Büchi arithmetic and linear arithmetic constraints over p-adic integers, p-existence and p-universality are decidable in NEXP and coNEXP, respectively.*

▶ **Theorem 2.** *Deciding p-universality for Büchi arithmetic is coNEXP-hard.*

Further general notation. For an arbitrary set A, we write $\#A$ for its cardinality. If A is infinite, then $\#A = \infty$. For $a, b \in \mathbb{Z}$, we write $[a, b]$ for the set $\{a, a+1, \dots, b\}$. Given a matrix $\mathbf{A} \in \mathbb{Z}^{n \times d}$ with components $a_{i,j} \in \mathbb{Z}$ ($i \in [1, n]$ and $j \in [1, d]$), the ∞-norm of $\mathbf{A}$ is defined as $\|\mathbf{A}\|_\infty \overset{\text{def}}{=} \max_{i=1,j=1}^{n,d} |a_{i,j}|$. We extend $\|.\|_\infty$ to vectors in $\mathbb{Z}^d$ by viewing them as elements of $\mathbb{Z}^{d \times 1}$. The $(1, \infty)$-norm of $\mathbf{A}$ is defined as $\|\mathbf{A}\|_{1,\infty} \overset{\text{def}}{=} \max_{i=1}^{n} \sum_{j=1}^{d} |a_{i,j}|$. Given a finite set $A \subseteq \mathbb{Z}^n$ of d integer vectors, we write $A^{\mathbf{M}}$ to denote the $n \times d$ matrix whose columns are the vectors in A, ordered following a lexicographic ordering. When clear from the context, we shall abbreviate $A^{\mathbf{M}}$ simply as $\mathbf{A}$. We write $\|A\|_\infty$ for $\|\mathbf{A}\|_\infty$.

Let $S\colon \mathbf{A}\cdot\boldsymbol{x}\geq\boldsymbol{c}$ be a system of linear inequalities with $\mathbf{A}\in\mathbb{Z}^{n\times d}$ and $\boldsymbol{c}\in\mathbb{Z}^n$. We write $[\![S]\!]$ for the solution set of S, that is the set of all $\boldsymbol{v}\in\mathbb{Z}^d$ such that $\mathbf{A}\cdot\boldsymbol{v}\geq\boldsymbol{c}$. We use $[\![S]\!]_{\geq 0}$ as a shorthand for $[\![S]\!]\cap\mathbb{N}^d$. Moreover, we define $\|S\|\overset{\text{def}}{=}\max(\|\mathbf{A}\|_\infty,\|\boldsymbol{c}\|_\infty)$. Finally, given a formula Φ of either Büchi arithmetic or linear arithmetic constraints over p-adic integers, we write $\|\Phi\|$ for the maximum absolute value of an integer appearing in Φ.

Deterministic p-automata and linear Diophantine equations. A central technical tool underlying the results of Theorem 1 are p-automata, a class of finite-state automata whose languages encode sets of natural numbers, see e.g. [4]. Given an integer $p\geq 2$, a *p-automaton* is a deterministic automaton over an alphabet $\Sigma_p^d := [0,p-1]^d$ for some positive integer d. A finite word $w=\boldsymbol{u}_k\cdots\boldsymbol{u}_0\in(\Sigma_p^d)^*$ over Σ_p^d can be seen as encoding a d-tuple of non-negative integers in base p. We consider a *msd-first encoding* $[\![\cdot]\!]^*$, in which the most significant digit is on the left. Formally, $[\![w]\!]^*\in\mathbb{N}^d$ is defined as $\sum_{j=0}^k p^k\cdot\boldsymbol{u}_j$. Also note that for $w=\varepsilon$, the empty word, we have $[\![w]\!]^*=\mathbf{0}$.

Following [29], we define a p-automaton whose language is the msd-first encoding of all non-negative integer solutions of a system of linear equations.

▶ **Definition 3.** *Let $S\colon \mathbf{A}\cdot\boldsymbol{x}=\boldsymbol{c}$ be a system of linear Diophantine equations with $\mathbf{A}\in\mathbb{Z}^{n\times d}$ and $\boldsymbol{c}\in\mathbb{Z}^n$. We define a p-automaton corresponding to S as $\mathcal{A}_p^*(S)\overset{\text{def}}{=}(Q,\Sigma_p^d,\delta,\boldsymbol{q}_0,F)$ with a set of* states $Q=\mathbb{Z}^n$, transitions $\delta(\boldsymbol{q},\boldsymbol{u})=p\cdot\boldsymbol{q}+\mathbf{A}\cdot\boldsymbol{u}$ *for all $\boldsymbol{q}\in Q$ and $\boldsymbol{u}\in\Sigma_p^d$,* initial state $\boldsymbol{q}_0=\mathbf{0}$, *and* final state $F=\{\boldsymbol{c}\}$.

For states $\boldsymbol{s},\boldsymbol{t}\in Q$ and $\boldsymbol{u}\in\Sigma_p^d$, we write $\boldsymbol{s}\xrightarrow{\boldsymbol{u}}_{\mathbf{A},p}\boldsymbol{t}$ whenever $\delta(\boldsymbol{s},\boldsymbol{u})=\boldsymbol{t}$. This notation is extended to words in the usual way: for a word $w\in(\Sigma_p^d)^*$, $\boldsymbol{s}\xrightarrow{w\cdot\boldsymbol{u}}_{\mathbf{A},p}\boldsymbol{t}$ whenever there is $\boldsymbol{q}\in Q$ such that $\boldsymbol{s}\xrightarrow{w}_{\mathbf{A},p}\boldsymbol{q}\xrightarrow{\boldsymbol{u}}_{\mathbf{A},p}\boldsymbol{t}$. We write $\boldsymbol{s}\rightarrow_{\mathbf{A},p}\boldsymbol{t}$ if $\boldsymbol{s}\xrightarrow{w}_{\mathbf{A},p}\boldsymbol{t}$ holds for some $w\in(\Sigma_p^d)^*$, and omit the subscripts $\mathbf{A}$ or p from $\rightarrow_{\mathbf{A},p}$ when clear from the context.

As usual, under *regular* acceptance condition, a finite word $w\in(\Sigma_p^d)^*$ is accepted by the automaton $\mathcal{A}=\mathcal{A}_p^*(S)$ whenever $\boldsymbol{q}_0\xrightarrow{w}\boldsymbol{f}$ for some $\boldsymbol{f}\in F$. The language $\mathcal{L}^*(\mathcal{A})$ of $\mathcal{A}$ is the set of all words that are accepted by $\mathcal{A}$. Even though the automaton $\mathcal{A}$ has infinitely many states, $\mathcal{L}^*(\mathcal{A})$ is a regular language since only finitely many *live states* can reach an accepting state.

▶ **Proposition 4** ([11], Prop. 5). *Given the automaton $\mathcal{A}_p^*(S)$, only states $\boldsymbol{q}\in Q$ such that $\|\boldsymbol{q}\|_\infty\leq\max(\|\mathbf{A}\|_{1,\infty},\|\boldsymbol{c}\|_\infty)$ can reach an accepting state.*

Proposition 4 implies a bound on the cardinality of the set L of live states of the p-automaton $\mathcal{A}_p^*(S)$ as defined in Definition 3:

$$\#L\leq 2^n\cdot\max(\|\mathbf{A}\|_{1,\infty},\|\boldsymbol{c}\|_\infty)^n \tag{1}$$

Observe that Proposition 4 also gives us a first key insight into deciding p-universality, as it shows that the set of live states of a p-automaton $\mathcal{A}_p^*(S)$ does *not* depend on the base p, but only on the system S. Deciding reachability in a p-automaton reduces to finding non-negative solutions to a certain system of Diophantine equations, as shown by the following proposition.

▶ **Proposition 5** ([11]). *Given $\boldsymbol{s},\boldsymbol{t}\in Q$, $k\in\mathbb{N}$ and $w\in(\Sigma_p^d)^k$, $\boldsymbol{s}\xrightarrow{w}\boldsymbol{t}$ iff $\boldsymbol{t}=p^k\cdot\boldsymbol{s}+\mathbf{A}\cdot[\![w]\!]^*$.*

In view of the bounds on the set of live states given in (1), the length of the shortest word w witnessing $\boldsymbol{s}\rightarrow\boldsymbol{t}$ is exponential in $\langle S\rangle$. Of course, when $\boldsymbol{s}=\mathbf{0}$ and $\boldsymbol{t}=\boldsymbol{c}$, this bound is non-optimal, as von zur Gathen and Sieveking [27] have shown that any feasible system of linear Diophantine equations S has a solution whose bit-size is polynomially bounded in $\langle S\rangle$. However, in the context of the p-universality problem, this bound on w is sufficient for us to establish the complexity upper bounds given by Theorem 1.

ω-regular acceptance condition and systems of equations over p-adic integers. A similar connection as in the previous paragraph can be established for systems of equations over p-adic integers [11]. In this setting, we consider infinite words $w = \boldsymbol{u}_0\boldsymbol{u}_1 \cdots \in (\Sigma_d^p)^\omega$ over Σ_p^d and view them as *lsd-first encodings* of d-tuples of p-adic integers in which the least significant digit is on the left. Formally, we define $[\![w]\!]^\omega \in \mathbb{Z}_p^d$ as $\sum_{j=0}^\infty p^j \cdot \boldsymbol{u}_j$.

Let $S\colon \mathbf{A} \cdot \boldsymbol{x} = \boldsymbol{c}$ be a system of linear equations with $\mathbf{A} \in \mathbb{Z}^{n \times d}$ and $\boldsymbol{c} \in \mathbb{Z}^n$, and let $w = \boldsymbol{u}_0\boldsymbol{u}_1 \cdots \in (\Sigma_d^p)^\omega$. We have $\mathbf{A} \cdot [\![w]\!]^\omega = \boldsymbol{c}$ if and only if $\mathbf{A} \cdot [\![w]\!]^\omega = \boldsymbol{c} \bmod p^k$ for all $k \in \mathbb{N}$. It follows, and was also discussed in [11], that $\mathbf{A} \cdot [\![w]\!]^\omega = \boldsymbol{c}$ if and only if for every $k \in \mathbb{N}$ there is $\boldsymbol{r} \in \mathbb{Z}^n$ such that $\mathbf{A} \cdot [\![\boldsymbol{u}_{k-1}\boldsymbol{u}_{k-2} \dots \boldsymbol{u}_0]\!]^* + \boldsymbol{r} \cdot p^k = \boldsymbol{c}$. By Proposition 5, the right hand side of this double implication expresses that the state $\boldsymbol{r}$ can reach $\boldsymbol{c}$ in the p-automaton $\mathcal{A}_p^*(S)$ by reading the word $\boldsymbol{u}_{k-1}\boldsymbol{u}_{k-2} \dots \boldsymbol{u}_0$. So, $\mathbf{A} \cdot [\![w]\!]^\omega = \boldsymbol{c}$ is satisfied whenever the Büchi automaton obtained from $\mathcal{A}_p^*(S)$ by reversing every transition and making all states accepting has a non-empty language for the initial state $\boldsymbol{c}$. This ω-regular acceptance condition can equivalently be formulated as follows:

▶ **Proposition 6.** *For all $w = \boldsymbol{u}_0\boldsymbol{u}_1 \cdots \in (\Sigma_d^p)^\omega$, $\mathbf{A} \cdot [\![w]\!]^\omega = \boldsymbol{c}$ iff there is $\boldsymbol{r} \in \mathbb{Z}^n$ and a strictly ascending sequence $(\lambda_i)_{i \in \mathbb{N}}$ such that $\boldsymbol{r} \xrightarrow{\boldsymbol{u}_{\lambda_0 - 1} \cdots \boldsymbol{u}_0}_{\mathbf{A},p} \boldsymbol{c}$ and $\boldsymbol{r} \xrightarrow{\boldsymbol{u}_{\lambda_{j+1}} \cdots \boldsymbol{u}_{\lambda_j}}_{\mathbf{A},p} \boldsymbol{r}$ for all $j \in \mathbb{N}$.*

Semi-linear set and ultimately periodic sets. Together with p-automata, to prove Theorem 1 we rely on well-known connections between solutions of systems of linear Diophantine equations and semi-linear sets. For $\boldsymbol{b} \in \mathbb{Z}^d$ and a finite set $P \subseteq \mathbb{Z}^d$ consisting of n elements, $L(\boldsymbol{b}, P)$ defines the *linear set* $\{\boldsymbol{x} \in \mathbb{Z}^d : \boldsymbol{x} = \boldsymbol{b} + \mathbf{P} \cdot \boldsymbol{\lambda} \text{ for some } \boldsymbol{\lambda} \in \mathbb{N}^n\}$. For a finite set $B \subseteq \mathbb{Z}^d$, $L(B, P)$ defines the *hybrid-linear set* $\bigcup_{\boldsymbol{b} \in B} L(\boldsymbol{b}, P)$. A *semi-linear set* is a finite union of hybrid-linear sets.

We use the following bound on the magnitude of the *bases* B and *periods* P of the set of solutions of a system of linear Diophantine equations, which is derived from [23].

▶ **Proposition 7** ([6], Prop. 4)**.** *Let $\mathbf{A} \in \mathbb{Z}^{n \times d}$, $\boldsymbol{c} \in \mathbb{Z}^n$ and $S\colon \mathbf{A} \cdot \boldsymbol{x} = \boldsymbol{c}$. Then $[\![S]\!]_{\geq 0} = L(B, P)$ where $\|B\|_\infty \leq ((d+1) \cdot \|\mathbf{A}\|_\infty + \|\boldsymbol{c}\|_\infty + 1)^n$ and $\|P\|_\infty \leq (d \cdot \|\mathbf{A}\|_\infty + 1)^n$.*

Eventually, deciding p-existence and p-universality reduces to characterising the set of bases p for which an existential formula of Büchi arithmetic (or linear arithmetic constraints over p-adic integers) is satisfiable. This leads us to consider semi-linear sets in $\mathbb{N}$, which are equivalent to *ultimately periodic sets*, i.e., sets of definable as $F \cup L(T, q)$, where $q \in \mathbb{N}$ is the *period* of the ultimately periodic set, $F \subseteq \mathbb{N}$ is a finite set such that $\max F < \min T$, and $T \subseteq [t, t + q - 1]$, where $t \in \mathbb{N}$ is the *threshold* of the ultimately periodic set.

Following [28], the essential building block leading to this change of representation, from one-dimensional semi-linear sets to ultimately periodic sets, is given by the proposition below.

▶ **Proposition 8.** *Let $M = L(B, P) \subseteq \mathbb{N}$. Then M is an ultimately periodic set with period $\gcd P$ and threshold bounded by $\|B\|_\infty + \|P\|_\infty^2$.*

We recall bounds on union, intersection and set difference of ultimately periodic sets.

▶ **Proposition 9.** *Let M and N be two ultimately periodic sets with periods and thresholds respectively (p_1, t_1) and (p_2, t_2). Then, $M \cup N$, $M \cap N$ and $M \setminus N$ are ultimately periodic sets with period $\operatorname{lcm}(p_1, p_2)$ and threshold $\max(t_1, t_2)$.*

Since $\mathbb{N} = L(0, 1)$, Proposition 9 shows that the complement $\mathbb{N} \setminus M$ of an ultimately periodic set M is itself ultimately periodic, and has the same period and threshold as M.

For linear arithmetic constraints over p-adic integers, p-existence and p-universality restrict p to be prime numbers. We handle this restriction by using a variant of Linnik's theorem [19] to guarantee the existence of small primes on arithmetic progressions.

▶ **Proposition 10.** *There is a constant $c > 0$ such that for all co-prime $b, q \in \mathbb{N}$ there is some $p \in L(b, q) \cap \mathbb{P}$ such that $p \leq c \cdot (b \cdot r)^5$.*

Proof. Under the assumption that $b \in [1, q-1]$, Linnik's theorem states that $L(b, q)$ contains a prime in $[1, d \cdot p^L]$ for fixed $d > 0$ and $L \in \mathbb{N}$. The best known bound for L is 5, as shown by Xylouris in [30]. To get rid of this additional restriction on b, consider a prime $s \in [b+1, 2(b+1)]$, whose existence follows from Bertrand's postulate [22]. From the primality of $s > b$ and the co-primality of b and q, we derive $\gcd(b, s \cdot q) = 1$ and $b < s \cdot r$. We can now safely apply Linnik's theorem, and derive that $L(b, s \cdot r) \subseteq L(b, q)$ contains a prime bounded by $c \cdot (b \cdot r)^5$ for some constant $c > 0$. ◀

Observe that every element of $L(b, q)$ is by definition divided by $\gcd(b, q)$. Hence, in the case where b and q are not co-prime, the only possible prime number appearing in $L(b, q)$ is b.

3 Exponential witnesses for p-existence and p-universality

For an existential formula Φ of either Büchi arithmetic or linear arithmetic constraints over p-adic integers, parametric in their base p, we write $\mathcal{B}(\Phi)$ for the set of bases $p \geq 2$ for which Φ is satisfiable. Note that, in defining $\mathcal{B}(\Phi)$ for linear arithmetic constraints over p-adic integers, we temporarily lift the primality condition on p. In this section, we establish the following result, which represents a crucial step in showing that the p-existence and p-universality problems are decidable in NEXP and coNEXP, respectively, proven in Section 4.

▶ **Theorem 11.** *Let Φ be an existential formula from Büchi arithmetic (resp. from linear arithmetic constraints over p-adic integers).*

- *If it exists, the smallest base $p \geq 2$ (resp. p prime) in $\mathcal{B}(\Phi)$ is bounded by $2^{2^{\mathcal{O}(\langle\Phi\rangle^2)}}$,*
- *If it exists, the smallest base $p \geq 2$ (resp. p prime) not in $\mathcal{B}(\Phi)$ is bounded by $2^{2^{\mathcal{O}(\langle\Phi\rangle^2)}}$.*

Whereas members of $\mathcal{B}(\Phi)$ are certificates of p-existence, a certificate for the non-universality of $\mathcal{B}(\Phi)$ can be retrieved from the "bases complement" $\overline{\mathcal{B}(\Phi)} \overset{\text{def}}{=} \mathbb{N} \setminus (\mathcal{B}(\Phi) \cup \{0, 1\})$. Consequently, a proof of Theorem 11 follows as soon as we show the following proposition.

▶ **Proposition 12.** *$\mathcal{B}(\Phi)$ is an ultimately periodic set with period and threshold in $2^{2^{\mathcal{O}(\langle\Phi\rangle^2)}}$.*

By Proposition 9, this result implies that $\overline{\mathcal{B}(\Phi)}$ is an ultimately periodic set with the same period and threshold as $\mathcal{B}(\Phi)$. Notice that for linear arithmetic constraints over p-adic integers the primality of the certificates can be obtained by an application of Linnik's theorem: consider the ultimately periodic representation $F \cup L(T, q)$ of $\mathcal{B}(\Phi)$, and suppose that it contains a prime. If $F \cup T$ contains a prime, then it is bounded by $2^{2^{\mathcal{O}(\langle\Phi\rangle^2)}}$. Otherwise, there is some $t \in T$ such that $L(t, q)$ contains a prime. So, t and q are co-prime (as $t \notin \mathbb{P}$), and by Linnik's theorem $L(t, q)$ has a prime bounded by $2^{2^{\mathcal{O}(\langle\Phi\rangle^2)}}$. Analogously, if $\mathcal{B}(\Phi)$ avoids a prime, then $\overline{\mathcal{B}(\Phi)}$ has a prime in $2^{2^{\mathcal{O}(\langle\Phi\rangle^2)}}$.

Proof of Proposition 12: Büchi arithmetic. Let Φ be a formula of existential Büchi arithmetic with parametric base p. To show Proposition 12, we introduce an abstraction of p-automata that we call *support graphs*. Support graphs are graphs that, while being independent from the base p, may correspond to paths of a p-automaton for a linear system $S\colon \mathbf{A}\cdot\boldsymbol{x} = \boldsymbol{c}$, and integrate auxiliary systems of inequalities that we use to enforce the satisfaction of formulae of the form $V_p(x) = y$, again independently of the choice of p.

▶ **Definition 13.** *Let $n \in \mathbb{N}$, and consider a tuple of variables $\boldsymbol{x}$. A* support graph *on $(n, \boldsymbol{x})$ is a finite directed graph (V, E) with vertices $V \subseteq \mathbb{Z}^n$ and edges E of the form $\boldsymbol{s} \to_T \boldsymbol{t}$, where $\boldsymbol{s}, \boldsymbol{t} \in V$ and T is a system of linear inequalities with variables from $\boldsymbol{x}$.*

A support graph can have multiple edges over the same two vertices, labelled with different systems of linear inequalities. We evaluate a support graph to the set of bases p for which it can be embedded into a p-automaton. Given $\boldsymbol{s}, \boldsymbol{t} \in V$ and a matrix $\mathbf{A} \in \mathbb{Z}^{n\times d}$, we define

$$[\![\boldsymbol{s} \to_T \boldsymbol{t}]\!]_{\mathbf{A}} \overset{\text{def}}{=} \{z \in \mathbb{N} : \boldsymbol{t} = \boldsymbol{s}\cdot z + \mathbf{A}\cdot\boldsymbol{x}, z \geq 2 \text{ and } \|\boldsymbol{x}\|_\infty < z, \text{ for some } \boldsymbol{x} \in [\![T]\!]_{\geq 0}\}.$$

Notice that $[\![\boldsymbol{s} \to_\top \boldsymbol{t}]\!]_{\mathbf{A}}$, where $\top$ is a (trivial) system of inequalities such that $[\![\top]\!]_{\geq 0} = \mathbb{N}^d$, corresponds to the set of bases $p \geq 2$ for which the p-automaton $\mathcal{A}_p^*(S)$ has a one-step transition from $\boldsymbol{s}$ to $\boldsymbol{t}$. As we only look at non-negative values for z and $\boldsymbol{x}$, we can introduce slack variables to translate the inequalities $z \geq 2$, $\|\boldsymbol{x}\|_\infty < z$, as well as all the ones in T, into equalities. This allows us to apply Proposition 7, followed by Proposition 8, to characterise $[\![\boldsymbol{s} \to_T \boldsymbol{t}]\!]_{\mathbf{A}}$ as an ultimately periodic set. Below, let $\|\boldsymbol{s} \to_T \boldsymbol{t}\|_\infty \overset{\text{def}}{=} \max(\|\boldsymbol{s}\|_\infty, \|\boldsymbol{t}\|_\infty, \|T\|)$.

▶ **Lemma 14.** *Let $\mathbf{A} \in \mathbb{Z}^{n\times d}$, $\boldsymbol{s}, \boldsymbol{t} \in \mathbb{Z}^n$ and let T be a linear system of m inequalities. The set $[\![\boldsymbol{s} \to_T \boldsymbol{t}]\!]_{\mathbf{A}}$ is an ultimately periodic set with period and threshold bounded by $U^{\mathcal{O}(k\log k)}$, where $k = n + d + m$ and $U = \max(2, \|\mathbf{A}\|_\infty, \|\boldsymbol{s} \to_T \boldsymbol{t}\|_\infty)$.*

Given a support graph $\mathcal{G}$ with edges $e_1, \dots, e_\ell$, we write $[\![\mathcal{G}]\!]_{\mathbf{A}}$ for $\bigcap_{i\in[1,\ell]} [\![e_i]\!]_{\mathbf{A}}$, i.e. the set of $p \geq 2$ such that, for every edge $\boldsymbol{s} \to_T \boldsymbol{t}$ of $\mathcal{G}$, the transition $\boldsymbol{s} \xrightarrow{\boldsymbol{u}}_{p,\mathbf{A}} \boldsymbol{t}$ holds for some tuple $\boldsymbol{u} \in \Sigma_p^d$ satisfying T. By Proposition 9 and Lemma 14, $[\![\mathcal{G}]\!]_{\mathbf{A}}$ is ultimately periodic.

To prove Proposition 12, we first translate the formula Φ (possibly by introducing slack variables to replace inequalities with equalities) in a disjunctive normal form with $2^{\mathcal{O}(\langle\Phi\rangle)}$ disjuncts have the form $\mathbf{A}\cdot\boldsymbol{x} = \boldsymbol{c} \wedge \bigwedge_{i\in I} V_p(x_i) = y_i$, where $\mathbf{A} \in \mathbb{Z}^{n\times d}$, $\boldsymbol{c} \in \mathbb{Z}^n$, and all variables are among the ones in $\boldsymbol{x}$. We further manipulate each of these disjuncts by considering all linear orderings among the variables y_i $(i \in I)$. Variables that are set to be equal in an ordering can be substituted accordingly, so that Φ is found to be equivalent to a disjunction of $2^{\mathcal{O}(\langle\Phi\rangle\log\langle\Phi\rangle)}$ formulae of size $\mathcal{O}(\langle\Phi\rangle)$ that have the form

$$\mathbf{A}\cdot\boldsymbol{x} = \boldsymbol{c} \wedge \textstyle\bigwedge_{(i,j)\in J} V_p(x_i) = y_j \wedge \bigwedge_{j\in[1,m]} y_j < y_{j-1} \qquad (2)$$

where $J \subseteq I \times [0, m]$ is a binary relation that is functional and surjective on its first component.

Let ψ be a formula of the form in (2). We aim at characterising $\mathcal{B}(\psi)$ as an ultimately periodic set. Recall that, by Proposition 5, solutions of the system $S\colon \mathbf{A}\cdot\boldsymbol{x} = \boldsymbol{c}$ are values $[\![w]\!]^* \in \mathbb{N}^d$ for some $w \in (\Sigma_p^d)^*$ such that $\boldsymbol{0} \xrightarrow{w}_{\mathbf{A},p} \boldsymbol{c}$. Moreover, a constraint $V_p(x) = y$ restricts the variables x and y to be such that, in their base-p msd representation, $y \in \{0\}^\ell \cdot \{1\} \cdot \{0\}^r$ and $x \in [0, p-1]^\ell \cdot [1, p-1] \cdot \{0\}^r$, for some $\ell, r \in \mathbb{N}$. Consequently, in order for $[\![w]\!]^*$ to be a solution of (2), the word w must admit a decomposition $w_0 \cdot \boldsymbol{u}_0 \cdot w_1 \cdots w_m \cdot \boldsymbol{u}_m \cdot w_{m+1}$ such that $\boldsymbol{u}_0, \dots, \boldsymbol{u}_m \in \Sigma_p^d$,

$$\boldsymbol{0} = \boldsymbol{s}_0 \xrightarrow{w_0}_{\mathbf{A},p} \boldsymbol{t}_0 \xrightarrow{\boldsymbol{u}_0}_{\mathbf{A},p} \boldsymbol{s}_1 \;\cdots\; \boldsymbol{s}_m \xrightarrow{w_m}_{\mathbf{A},p} \boldsymbol{t}_m \xrightarrow{\boldsymbol{u}_m}_{\mathbf{A},p} \boldsymbol{s}_{m+1} \xrightarrow{w_{m+1}}_{\mathbf{A},p} \boldsymbol{t}_{m+1} = \boldsymbol{c}, \qquad (3)$$

where $\boldsymbol{s}_1, \dots, \boldsymbol{s}_{m+1}, \boldsymbol{t}_0, \boldsymbol{t}_m$ are (intermediate) live states, and every $\boldsymbol{u}_j$ and w_j with $j \in [0, m]$ shall satisfy the constraints induced by the function V_p together with the ordering on the variables y_j. In particular, following the aforementioned decomposition for the variables

x and y appearing in a constraint $V_p(x) = y$, for every $j \in [0, m]$ the values of $\boldsymbol{u}_j$ for the variables $x_1, \dots, x_{\#I}$ and $y_1, \dots, y_m$ shall satisfy the system U_j:

$$\begin{cases} y_j = 1, & x_i \geq 1 \ \ : i \in I \text{ and } V_p(x_i) = y_j \text{ occurs in (2)}, \\ y_k = 0 \ \ : k \in [1, m] \setminus \{j\}, & x_i = 0 \ \ : i \in I \text{ and } V_p(x_i) = y_k \text{ occurs in (2) for some } k < j. \end{cases}$$

whereas at each position of the word w_j ($j \in [0, m+1]$) shall satisfy the system W_j:

$$\begin{cases} y_k = 0 \ \ : k \in [1, m], \qquad x_i = 0 \ \ : i \in I \text{ and } V_p(x_i) = y_k \text{ occurs in (2), for some } k < j. \end{cases}$$

Hence, paths as in (3) can be abstracted into support graphs with vertices from the set of live states of $\mathcal{A}_p^*(S)$ and having the form

$$\mathbf{0} = \boldsymbol{s}_0 \to^{j_0}_{W_0} \boldsymbol{t}_0 \to_{U_0} \boldsymbol{s}_1 \ \dots \ \boldsymbol{s}_m \to^{j_m}_{W_m} \boldsymbol{t}_m \to_{U_m} \boldsymbol{s}_{m+1} \to^{j_{m+1}}_{W_{m+1}} \boldsymbol{t}_{m+1} = \boldsymbol{c}, \tag{4}$$

where $\boldsymbol{s} \to^{j}_{T} \boldsymbol{t}$ is short for a path of length j going from $\boldsymbol{s}$ to $\boldsymbol{t}$, and with arrows labelled by the system of inequalities T, and for every $i \in [0, m+1]$, j_i is the length of w_i.

▶ **Lemma 15.** *Let ψ be a formula as in* (2)*.*

- *For every support graph $\mathcal{G}$ of the form described in* (4)*, $[\![\mathcal{G}]\!]_{\mathbf{A}} \subseteq \mathcal{B}(\psi)$.*
- *For every $p \in \mathcal{B}(\psi)$ there is a support graph $\mathcal{G}$ as in* (4) *such that $p \in [\![\mathcal{G}]\!]_{\mathbf{A}}$.*

Proof. Let $\mathbb{G}$ be the set of support graphs of the form in (4). The lemma equivalently states that $\mathcal{B}(\psi) = \bigcup_{\mathcal{G} \in \mathbb{G}} [\![\mathcal{G}]\!]_{\mathbf{A}}$. Below, we refer to the first and second points in the lemma as the two inclusions $\supseteq$ and $\subseteq$ of this equality.
($\supseteq$): Let $\mathcal{G}$ be a support graph in $\mathbb{G}$, and consider $p \in [\![\mathcal{G}]\!]_{\mathbf{A}}$. Notice that, by definition, this means that for every edge $\boldsymbol{s} \to_T \boldsymbol{t}$ of $\mathcal{G}$ there is $\boldsymbol{u} \in \Sigma_p^d$ such that $\boldsymbol{s} \xrightarrow{\boldsymbol{u}}_{\mathbf{A},p} \boldsymbol{t}$ and $\boldsymbol{u} \in [\![T]\!]$. From (4), there is a path

$$\mathbf{0} = \boldsymbol{s}_0 \xrightarrow{w_0}_{\mathbf{A},p} \boldsymbol{t}_0 \xrightarrow{\boldsymbol{u}_0}_{\mathbf{A},p} \boldsymbol{s}_1 \ \dots \ \boldsymbol{s}_m \xrightarrow{w_m}_{\mathbf{A},p} \boldsymbol{t}_m \xrightarrow{\boldsymbol{u}_m}_{\mathbf{A},p} \boldsymbol{s}_{m+1} \xrightarrow{w_{m+1}}_{\mathbf{A},p} \boldsymbol{t}_{m+1} = \boldsymbol{c},$$

where $w \overset{\text{def}}{=} w_0 \cdot \boldsymbol{u}_0 \cdot w_1 \cdot \ldots \cdot w_m \cdot \boldsymbol{u}_m \cdot w_{m+1} \in (\Sigma_p^d)^*$, $\boldsymbol{u}_0, \dots, \boldsymbol{u}_m \in \Sigma_p^d$, every $\boldsymbol{u}_j$ satisfies U_j and every symbol in w_j satisfies W_j. Clearly, $\mathbf{A} \cdot [\![w]\!] = \boldsymbol{c}$. From the definition of the systems $U_0, \dots, U_m$ and $W_0, \dots, W_{m+1}$, we obtain that in w the value for the variable y_j ($j \in [0, m]$) has a base-p *msd* representation of the form $\{0\}^{\ell_j} \cdot \{1\} \cdot \{0\}^{r_j}$, for some $\ell_j, r_j \in \mathbb{N}$ such that $\ell_j + 1 + r_j$ corresponds to the length of w. This means that every y_j is a power of p. Moreover, $r_{j-1} > r_j$ for every $j \in [1, m]$, and therefore $y_j < y_{j-1}$. Lastly, consider $(i, j) \in J$, so that $V_p(x_i) = y_j$ appears in ψ. The systems $U_0, \dots, U_m$ and $W_0, \dots, W_{m+1}$ force the base-p msd encoding of x_i to belong to the language $[0, p-1]^{\ell_j} \cdot [1, p-1] \cdot \{0\}^{r_j}$. We conclude that the formula $V_p(x_i) = y_j$ holds. So, ψ is satisfiable with respect to the base p, i.e., $p \in \mathcal{B}(\psi)$.
($\subseteq$): Follows conversely to the other inclusion. Suppose ψ satisfiable with respect to the base p. Consider a word $w \in (\Sigma_p^d)^*$ such that $[\![w]\!]^*$ is a solution of ψ. From $\bigwedge_{(i,j) \in J} V_p(x_i) = y_j$ we conclude that the base-p *msd* encodings of x_i and y_j belong to $\{0\}^{\ell_j} \cdot \{1\} \cdot \{0\}^{r_j}$ and $[0, p-1]^{\ell_j} \cdot [1, p-1] \cdot \{0\}^{r_j}$, respectively, for some ℓ_j and r_j such that $\ell_j + 1 + r_j$ corresponds to the length of w. From $y_j < y_{j-1}$ ($j \in [1, m]$), $r_{j-1} > r_j$. Hence, w admits a decomposition $w_0 \cdot \boldsymbol{u}_0 \cdot w_1 \cdot \ldots \cdot w_m \cdot \boldsymbol{u}_m \cdot w_{m+1}$ such that $\boldsymbol{u}_0, \dots, \boldsymbol{u}_m \in \Sigma_p^d$,

$$\mathbf{0} = \boldsymbol{s}_0 \xrightarrow{w_0}_{\mathbf{A},p} \boldsymbol{t}_0 \xrightarrow{\boldsymbol{u}_0}_{\mathbf{A},p} \boldsymbol{s}_1 \ \dots \ \boldsymbol{s}_m \xrightarrow{w_m}_{\mathbf{A},p} \boldsymbol{t}_m \xrightarrow{\boldsymbol{u}_m}_{\mathbf{A},p} \boldsymbol{s}_{m+1} \xrightarrow{w_{m+1}}_{\mathbf{A},p} \boldsymbol{t}_{m+1} = \boldsymbol{c},$$

where $\boldsymbol{s}_1, \dots, \boldsymbol{s}_{m+1}, \boldsymbol{t}_0, \boldsymbol{t}_m$ are live states. Moreover, every $\boldsymbol{u}_j$ is a solution of U_j and every symbol in the word w_j is a solution of W_j. Let $\mathcal{G}$ be the support graph with edges $\boldsymbol{t}_0 \to_{U_0} \boldsymbol{s}_1, \dots, \boldsymbol{t}_m \to_{U_j} \boldsymbol{s}_{m+1}$ together with $\boldsymbol{i} \to_{W_j} \boldsymbol{i}'$, for every $j \in [0, m+1]$ and every two states $\boldsymbol{i}, \boldsymbol{i}'$ such that $\boldsymbol{i} \to_{\mathbf{A},p} \boldsymbol{i}'$ appears in the path going from $\boldsymbol{s}_j$ to $\boldsymbol{t}_j$. The graph $\mathcal{G}$ is of the form in (4), and $p \in [\![\mathcal{G}]\!]_{\mathbf{A}}$. ◀

Only finitely many support graphs have the form described in (4), as they all have vertices from the finite set of live states of $\mathcal{A}_p^*(S)$, and edges with labels from a finite set of linear systems. So, Lemma 15 implies that the set $\mathcal{B}(\psi)$ is equivalent to a finite union of $[\![\mathcal{G}]\!]_{\mathbf{A}}$, for which we can obtain an ultimately periodic representation according to Proposition 9 and Lemma 14.

▶ **Lemma 16.** *Let ψ be as in (2). Then $\mathcal{B}(\psi)$ is ultimately periodic with threshold in $U^{\mathcal{O}(k \log k)}$ and period in $U^{\mathcal{O}(\ell \cdot k \log k)}$, where $U = \max(2, \|\mathbf{A}\|_{1,\infty}, \|\boldsymbol{c}\|_\infty)$, $k = n + 3d^2$ and $\ell = U^{4n}$.*

Proof. Let $\mathbb{G}$ be the finite family of support graphs such that $\mathcal{B}(\psi) = \bigcup_{\mathcal{G} \in \mathbb{G}} [\![\mathcal{G}]\!]_{\mathbf{A}}$, according to Lemma 15. Every edge $\boldsymbol{s} \to_T \boldsymbol{t}$ of a support graph $\mathcal{G} \in \mathbb{G}$ is such that $\|\boldsymbol{s}\|_\infty, \|\boldsymbol{t}\|_\infty \leq \max(\|\mathbf{A}\|_{1,\infty}, \|\boldsymbol{c}\|_\infty)$ and T is a system among $U_0, \dots, U_m, W_0, \dots, W_{m+1}$. Hence, all the graphs in $\bigcup_{j \in J} \mathbb{G}_j$ are built from a set E of $(2 \cdot \max(\|\mathbf{A}\|_{1,\infty}, \|\boldsymbol{c}\|_\infty))^{2n} \cdot (2m + 3) \leq \mathcal{O}(\ell)$ edges (note: $m \leq \#I \leq d^2$). Each possible linear system T labelling an edge in E has at most $2d^2$ inequalities, with coefficients and constants in $\{0, 1\}$. By Lemma 14, each edge $e \in E$ is such that $[\![e]\!]_{\mathbf{A}}$ is an ultimately periodic set with period and threshold bounded by $U^{\mathcal{O}(k \log k)}$. By Proposition 9, taking unions and intersections of sets $[\![e]\!]_{\mathbf{A}}$ with $e \in E$, as for instance $\mathcal{B}(\psi) = \bigcup_{\mathcal{G} \in \mathbb{G}} [\![\mathcal{G}]\!]_{\mathbf{A}}$, always yields an ultimately periodic set with threshold in $U^{\mathcal{O}(k \log k)}$ and period in $U^{\mathcal{O}(\ell \cdot k \log k)}$. ◀

The bounds U, k and ℓ established in Lemma 16 for the threshold and the period of $\mathcal{B}(\varphi)$ can be restated in terms of the size of the initial formula Φ as follows: $U \leq 2^{\mathcal{O}(\langle \Phi \rangle)}$, $k \leq \mathcal{O}(\langle \Phi \rangle^2)$ and $\ell \leq 2^{\mathcal{O}(\langle \Phi \rangle^2)}$. This is sufficient to conclude that Proposition 12 holds. Indeed, the formula Φ is equivalent to a disjunction $\bigvee_{k \in K} \psi_k$ of formulae ψ_k of the form in (2), with $\#K \leq 2^{\mathcal{O}(\langle \Phi \rangle \log \langle \Phi \rangle)}$. This means that the set $\mathcal{B}(\Phi)$ is the union of all $\mathcal{B}(\psi_k)$ with $k \in K$. We apply Proposition 9 to obtain a representation of $\mathcal{B}(\Phi)$ as an ultimately periodic set with threshold bounded by $2^{\mathcal{O}(\langle \Phi \rangle^3 \log \langle \Phi \rangle)}$ and period bounded by $2^{2^{\mathcal{O}(\langle \Phi \rangle^2)}}$.

Proof of Proposition 12: Linear arithmetic constraints over p-adic integers. We now establish Proposition 12 for the case of Φ being an existential formula of linear arithmetic constraints over p-adic integers with parametric base p. For brevity, all proofs in this section are relegated to Appendix B. The crucial difference from the proof of Proposition 12 for Büchi arithmetic is that, differently from the V_p function, the p-adic valuation v_p induces constraints related to the relative lengths of subwords of the infinite words accepted by the p-automaton. For instance, to satisfy the formula $v_p(u) = 3 \cdot v_p(v) \land v_p(v) \geq 1$, the base-$p$ lsd-first representation of u and v must obey the following constraints:

$$\begin{array}{rcllll} u & \in & \{0\}^i & \{0\} & \{0\}^{2i+1} & [1, p-1]\ (\Sigma_p)^\omega, \\ v & \in & \{0\}^i & [1, p-1] & (\Sigma_p)^{2i+1} & \Sigma_p \qquad (\Sigma_p)^\omega. \end{array}$$

where $i \geq 1$. In particular, we notice that the length of the maximal all-zeros prefix of v fixes the length of the maximal all-zeros prefix of u, and vice versa. This reflects in the proof of Proposition 12 where, instead of only considering support graphs that are linear structures in the sense of (4), we must consider arbitrary graphs and establish ultimately periodic representations of the lengths of their paths. To do so, we rely on the following result on the lengths of words accepted by a *nondeterministic finite automaton* (NFA) over a *unary alphabet*. Recall that an NFA is a tuple $\mathcal{A} = (Q, \delta, I, F)$ where Q is a finite set of states, $\delta \subseteq Q \times Q$ is a transition relation, and $I, F \subseteq Q$ are set of initial and final states, respectively.

▶ **Proposition 17** ([24]). *Given an unary NFA $\mathcal{A} = (Q, \delta, I, F)$ with $s = \#Q$, one can construct in time $\mathcal{O}(s^2(s + \#\delta))$ a set $L = \bigcup_{k \in K} L(b_k, q_k)$ characterising the lengths of the words accepted by $\mathcal{A}$, with $\#K \leq \mathcal{O}(s^2)$, $b_k \leq (2 \cdot s + 1) \cdot s$ and $q_k \leq s$.*

Of course, other modifications with respect to the treatment of Büchi arithmetic are required: the support graphs must take into account the ω-regular acceptance condition defined in Proposition 6, and we also have to deal with the two-sorted structure of the theory.

Moving to the proof of Proposition 12, similarly to the case of Büchi arithmetic we start by manipulating the formula Φ and obtain a disjunctive normal form where each of the $2^{\mathcal{O}(\langle\Phi\rangle \log \langle\Phi\rangle)}$ disjuncts are of size $\mathcal{O}(\langle\Phi\rangle)$ and have the following form:

$$\mathbf{A} \cdot \boldsymbol{u} = \boldsymbol{c} \land \mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d} \land \textstyle\bigwedge_{(i,j) \in J} v_p(u_i) = x_j \land \bigwedge_{j \in [1,r]} x_{j-1} < x_j \tag{5}$$

where $\mathbf{A} \in \mathbb{Z}^{n \times d}$, $\boldsymbol{c} \in \mathbb{Z}^n$, $\mathbf{B} \in \mathbb{Z}^{m \times e}$, $\boldsymbol{d} \in \mathbb{Z}^m$ and $J \subseteq I \times [0, r]$ is a binary relation that is functional and surjective on its first component. Each u_i with $i \in I$ is a variable among $\boldsymbol{u}$ interpreted over $\mathbb{Z}_p$, and each x_j with $j \in [0, r]$ is a variable among $\boldsymbol{x}$, interpreted over $\mathbb{Z}$. Notice that restricting the interpretation of $\boldsymbol{x}$ from $\overline{\mathbb{Z}}$ to $\mathbb{Z}$ is without loss of generality: when bringing the formula Φ in disjunctive normal form, we can introduce tautologies of the form $x < \infty \lor x = \infty$, for each of the variables x in $\boldsymbol{x}$. Then, following the axiom system presented in [17], disjuncts where $x = \infty$ holds can be easily modified so that x is eliminated.

According to Proposition 6, solutions of the system $S\colon \mathbf{A} \cdot \boldsymbol{u} = \boldsymbol{c}$ over the p-adic integers are values $[\![w]\!]^\omega \in \mathbb{Z}_p^d$ for some infinite word $w = \boldsymbol{u}_0\boldsymbol{u}_1 \cdots \in (\Sigma_p^d)^\omega$ such that there is a live state $\boldsymbol{r} \in \mathbb{Z}^n$ of the p-automaton $\mathcal{A}_p^*(S)$ and an infinite sequence $\lambda_0 < \lambda_1 < \dots$ for which $\boldsymbol{r} \xrightarrow{\boldsymbol{u}_{\lambda_0 - 1} \cdots \boldsymbol{u}_0}_{\mathbf{A},p} \boldsymbol{c}$ and $\boldsymbol{r} \xrightarrow{\boldsymbol{u}_{\lambda_{j+1}} \cdots \boldsymbol{u}_{\lambda_j}}_{\mathbf{A},p} \boldsymbol{r}$ for all $j \in \mathbb{N}$. Moreover, a constraint $v_p(u) = x$ restricts the variables u and x to be such that, in the base p lsd-first representation of u, we have $u \in \{0\}^x \cdot [1, p-1] \cdot [0, p-1]^\omega$. Consequently, in order for $([\![w]\!]^\omega, \boldsymbol{x})$ to be a solution of (5), in addition to $\mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$, the word w must have a prefix of the form $w_0 \cdot \boldsymbol{v}_0 \cdot w_1 \cdots w_r \cdot \boldsymbol{v}_r \cdot w_{r+1}$ such that $\boldsymbol{v}_0, \dots, \boldsymbol{v}_r \in \Sigma_p^d$, the word $w_0 \in (\Sigma_p^d)^*$ has length x_0, each $w_i \in (\Sigma_p^d)^*$ with $i \in [1, r]$ has length $x_i - (x_{i-1} + 1) \geq 0$, and

$$\boldsymbol{r} = \boldsymbol{s}_{r+1} \xrightarrow{(w_{r+1})^R}_{\mathbf{A},p} \boldsymbol{t}_{r+1} \xrightarrow{\boldsymbol{v}_r}_{\mathbf{A},p} \boldsymbol{s}_r \ \dots \ \boldsymbol{s}_1 \xrightarrow{(w_1)^R}_{\mathbf{A},p} \boldsymbol{t}_1 \xrightarrow{\boldsymbol{v}_0}_{\mathbf{A},p} \boldsymbol{s}_0 \xrightarrow{(w_0)^R}_{\mathbf{A},p} \boldsymbol{t}_0 = \boldsymbol{c} \tag{6}$$

where each $(w_i)^R$ with $i \in [0, r+1]$ is the reverse of the word w_i, and $\boldsymbol{r}$ is a live state of $\mathcal{A}_p^*(S)$ for which the ω-regular condition of Proposition 6 is satisfied. Following the decomposition for the variable u appearing in a constraint $v_p(u) = x$ given above, for every $j \in [0, r]$ the values of $\boldsymbol{v}_j$ for the variables $u_1, \dots, u_{\#I}$ shall satisfy the system U_j:

$$\begin{cases} u_i \geq 1 & : i \in I \text{ and } v_p(u_i) = x_j \text{ occurs in (5)}, \\ u_i = 0 & : i \in I \text{ and } v_p(u_i) = x_k \text{ occurs in (5), for some } k \in [j+1, r]. \end{cases}$$

whereas at each position of the word w_j ($j \in [0, r+1]$) shall satisfy the system W_j:

$$\begin{cases} u_i = 0 & : i \in I \text{ and } v_p(u_i) = x_k \text{ occurs in (5), for some } k \in [j, r]. \end{cases}$$

Given a formula ψ of the form in (5), we abstract paths in the p-automaton $\mathcal{A}_p^*(S)$ induced by infinite words such as the word w above, by introducing a family of support graphs, denoted by $\mathbb{G}(\psi)$. Each support graph $\mathcal{G} \in \mathbb{G}(\psi)$ has live states of $\mathcal{A}_p^*(S)$ as vertices, and its set of edges can be partitioned in the following sets, for some *intermediate live states* $\boldsymbol{r}, \boldsymbol{s}_0, \dots, \boldsymbol{s}_{r+1}, \boldsymbol{t}_0, \dots, \boldsymbol{t}_{r+1}$ such that $\boldsymbol{r} = \boldsymbol{s}_{r+1}$ and $\boldsymbol{t}_0 = \boldsymbol{c}$:

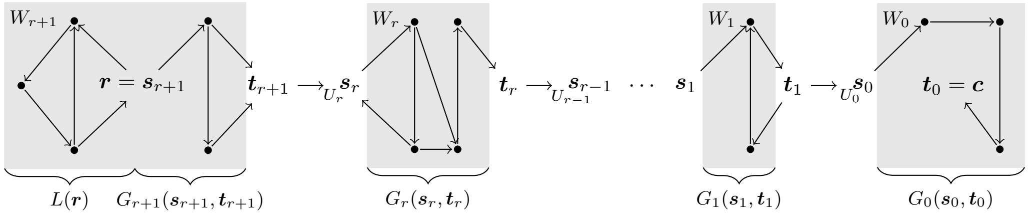
W_{r+1}
$\boldsymbol{r} = \boldsymbol{s}_{r+1}$
$\boldsymbol{t}_{r+1} \to_{U_r} \boldsymbol{s}_r$
W_r
$\boldsymbol{t}_r \to_{U_{r-1}} \boldsymbol{s}_{r-1}$ …
$\boldsymbol{s}_1$
W_1
$\boldsymbol{t}_1 \to_{U_0} \boldsymbol{s}_0$
W_0
$\boldsymbol{t}_0 = \boldsymbol{c}$
$L(\boldsymbol{r})$
$G_{r+1}(\boldsymbol{s}_{r+1}, \boldsymbol{t}_{r+1})$
$G_r(\boldsymbol{s}_r, \boldsymbol{t}_r)$
$G_1(\boldsymbol{s}_1, \boldsymbol{t}_1)$
$G_0(\boldsymbol{s}_0, \boldsymbol{t}_0)$

Figure 1 A support graphs for existential linear arithmetic constraints over p-adic integers.

- $C(\boldsymbol{r})$: a set of edges of the form $\boldsymbol{s} \to_{W_{r+1}} \boldsymbol{t}$ that describes a connected graph with a non-empty path from $\boldsymbol{r}$ to itself (as required by the ω-regular condition of Proposition 6),
- $G_j(\boldsymbol{s}_j, \boldsymbol{t}_j)$, with $j \in [0, r+1]$: a set of edges of the form $\boldsymbol{s} \to_{W_j} \boldsymbol{t}$ that describes a connected graph with a (possibly empty) path going from $\boldsymbol{s}_j$ to $\boldsymbol{t}_j$,
- $\{\boldsymbol{t}_{j+1} \to_{U_j} \boldsymbol{s}_j\}$, for every $j \in [0, r]$.

As the set of live states of $\mathcal{A}_p^*(S)$ and the set of all linear systems U_j and W_j considered are finite, so is the set $\mathbb{G}(\psi)$. Figure 1 depicts a support graph from $\mathbb{G}(\psi)$. We say that $\mathcal{G}$ *generates the length values* $(v_0, \dots, v_r) \in \mathbb{N}^{r+1}$ if $\mathcal{G}$ has a path of the form

$$\boldsymbol{t}_{r+1} \to_{U_r} \boldsymbol{s}_r \to_{W_r}^{v_r-(v_{r-1}+1)} \boldsymbol{t}_r \to_{U_{r-1}} \boldsymbol{s}_{r-1} \ \dots\ \boldsymbol{s}_1 \to_{W_1}^{v_1-(v_0+1)} \boldsymbol{t}_1 \to_{U_0} \boldsymbol{s}_0 \to_{W_0}^{v_0} \boldsymbol{t}_0 = \boldsymbol{c} \tag{7}$$

Exactly as in the case of Büchi arithmetic, we aim at characterising $\mathcal{B}(\psi)$ as a union over a subset B of $\{[\![\mathcal{G}]\!]_{\mathbf{A}} : \mathcal{G} \in \mathbb{G}(\psi)\}$. According to the definition of $\mathcal{G} \in \mathbb{G}(\psi)$, the set $[\![\mathcal{G}]\!]_{\mathbf{A}}$ consists of some of the bases $p \geq 2$ for which, if we disregard the constraints imposed by the system of inequalities $\mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$, the formula ψ is satisfiable. To account for this system, we need to characterise the set of all length values that can be generated from $\mathcal{G}$, and check whether $\mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d} \land \bigwedge_{j \in [0,r]} x_j = v_j$ can be satisfied with respect to one of these length values $(v_0, \dots, v_r)$. This check, which we now formalise, does not depend on the base-p, so that either $[\![\mathcal{G}]\!]_{\mathbf{A}} \subseteq \mathcal{B}(\psi)$ or $\mathcal{G}$ can be discarded when constructing the set B.

Consider $\mathcal{G} \in \mathbb{G}(\psi)$, with intermediate live states $\boldsymbol{s}_0, \dots, \boldsymbol{s}_{r+1} = \boldsymbol{r}, \boldsymbol{c} = \boldsymbol{t}_0, \dots, \boldsymbol{t}_{r+1}$. For every $j \in [0, r]$, we construct from the set of edges $G_j(\boldsymbol{s}_j, \boldsymbol{t}_j)$ the unary NFA (Q, δ, I, F) where Q is the set of live states appearing in some of the edges of $G_j(\boldsymbol{s}_j, \boldsymbol{t}_j)$, $I = \{\boldsymbol{s}_j\}$ and $F = \{\boldsymbol{t}_j\}$, and $\delta = \{(\boldsymbol{s}, \boldsymbol{t}) \in Q^2 : \boldsymbol{s} \to_{W_j} \boldsymbol{t} \in G_j(\boldsymbol{s}_j, \boldsymbol{t}_j)\}$. By Proposition 17, the set L_j of the lengths of the words accepted by this automaton is ultimately periodic. To obtain the length values that can be generated by $\mathcal{G}$, we combine the lengths of the sets $L_0, \dots, L_r$, and construct the following set $\oplus(L_0, \dots, L_r)$:

$$\oplus(L_0, \dots, L_r) \overset{\text{def}}{=} \left\{ \begin{pmatrix} \ell_0 \\ 1 + \ell_0 + \ell_1 \\ \dots \\ r + \sum_{i=0}^{r} \ell_i \end{pmatrix} \in \mathbb{N}^{r+1} : \ell_j \in L_j \text{ for all } j \in [0, r] \right\} \tag{8}$$

Below, we set $U \overset{\text{def}}{=} \max(2, \|\mathbf{A}\|_{1,\infty}, \|\boldsymbol{c}\|_\infty)$, so that the live states of S are at most U^{2n}.

▶ **Lemma 18.** *The set $\oplus(L_0, \dots, L_r)$ contains all length values generated by $\mathcal{G}$. One can construct in time $\mathcal{O}(r^2) \cdot U^{\mathcal{O}(n \cdot r)}$ a representation of $\oplus(L_0, \dots, L_r)$ as a semi-linear set $\bigcup_{k \in K} L(\boldsymbol{b}_k, P_k)$, where $\#K \leq U^{\mathcal{O}(n \cdot r)}$, $\#P \leq r+1$, $\|P\|_\infty \leq U^{2n}$ and $\|\boldsymbol{b}_k\|_\infty \leq \mathcal{O}(r \cdot U^{4n})$.*

As already stated, the set $L_{\mathcal{G}}$ allows us to characterise $\mathcal{B}(\psi)$ as a union over some of the sets in $\{[\![\mathcal{G}]\!]_{\mathbf{A}} : \mathcal{G} \in \mathbb{G}(\psi)\}$. This is formalised by the two following lemma, analogous to Lemma 15 established for Büchi arithmetic. For brevity, we write $\mathcal{G} \vdash \mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$ whenever there is a length value $(v_0, \dots, v_r) \in L_{\mathcal{G}}$ such that $\mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d} \land \bigwedge_{j \in [0,r]} x_i = v_i$ is satisfiable.

▶ **Lemma 19.** *Let ψ be a formula of the form given in* (5)*.*

- *Given $\mathcal{G} \in \mathbb{G}(\psi)$, if $\mathcal{G} \vdash \mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$ then $[\![\mathcal{G}]\!]_{\mathbf{A}} \subseteq \mathcal{B}(\psi)$.*
- *For every $p \in \mathcal{B}(\psi)$, there is $\mathcal{G} \in \mathbb{G}(\psi)$ such that $\mathcal{G} \vdash \mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$ and $p \in [\![\mathcal{G}]\!]_{\mathbf{A}}$.*

Following the case of Büchi arithmetic, we then express $\mathcal{B}(\psi)$ as an ultimately periodic set.

▶ **Lemma 20.** *Let ψ be as in* (5)*. The set $\mathcal{B}(\psi)$ is ultimately periodic, with threshold bounded by $U^{\mathcal{O}(k \log k)}$ and period bounded by $U^{\mathcal{O}(\ell \cdot k \log k)}$, with $k = n + 3d$ and $\ell = (r+2) \cdot U^{4n+1}$.*

Together, Proposition 9 and Lemma 20 yield Proposition 12, as we recall that Φ is equivalent to a disjunction $\bigvee_{k \in K} \psi_k$ of formulae ψ_k of the form in (5), with $\#K \leq 2^{\mathcal{O}(\langle \Phi \rangle \log \langle \Phi \rangle)}$.

4 Deciding satisfiability when the base p is large

In view of the magnitude of the bases p established in Theorem 11, in order to prove Theorem 1, i.e., to show that the p-existence and p-universality problems are decidable in NEXP and coNEXP, respectively, it is sufficient to show the following statement.

▶ **Theorem 21.** *Let Φ be an existential formula of linear arithmetic constraints over p-adic integers (resp. Büchi arithmetic) with parametric base p. Then satisfiability of Φ with respect to a given value $p \in \mathbb{P}$ (resp. $p \geq 2$) can be decided in time $2^{\mathcal{O}(\langle \Phi \rangle^3)} \cdot \mathcal{O}(\langle p \rangle)$.*

This result cannot directly be obtained from [11], where it is shown that satisfiability of Φ with the base p given in binary is decidable NP, as it only gives a coNEXP$^{\text{NP}}$ upper bound for p-universality when $\langle p \rangle$ is of exponential size. For our purposes, we require a decision procedure that runs in time polynomial in the size of the binary encoding of p provided as input. This can be a achieved by appealing to a strongly polynomial-time algorithm for the feasibility problem of a system of linear Diophantine inequalities in a fixed dimension established in [9].

▶ **Proposition 22** ([9])**.** *Let $S \colon \mathbf{A} \cdot \boldsymbol{x} \geq \boldsymbol{c}$ be a system of linear Diophantine inequalities, with $\mathbf{A} \in \mathbb{Z}^{n \times d}$ and $\boldsymbol{c} \in \mathbb{Z}^n$. Checking whether $[\![S]\!] \neq \emptyset$ can be decided using $d^{2.5d+o(d)} \cdot \langle S \rangle$ arithmetic operations, and space polynomial in $\langle S \rangle$.*

With this proposition at hand, proving Theorem 21 for both existential formulas of linear arithmetic constraints over p-adic integers and Büchi arithmetic, respectively, is not difficult. Any such formula can be converted in time $2^{\mathcal{O}(\langle \Phi \rangle \log \langle \Phi \rangle)}$ into a disjunctive normal form with disjuncts of the form given in (2) and (5), respectively. For every disjunct, we iterate in time $2^{\mathcal{O}(\langle \Phi \rangle^2)}$ over all decompositions of the form describe in (4) and (7), respectively, with each decomposition giving rise to family of systems of linear Diophantine equations whose number of variables is bounded by $\mathcal{O}(\langle \Phi \rangle^2)$ and whose coefficients are bounded by $\mathcal{O}(p + 2^{\langle \Phi \rangle})$. Proposition 22 then enables to decide any of such system with the required time bounds. Full details are deferred to Appendix C.

5 Büchi arithmetic: coNEXP lower bound for p-universality

Here, we prove Theorem 2 and show that the p-universality problem for existential Büchi arithmetic is coNEXP-hard. The proof is by a reduction from a coNEXP-complete generalisation of the quantified Boolean satisfiability problem, denoted by QOΠ_1-Sat (where QO stands for "quantified oracle"), that was introduced in [1] and later generalised in [21]. For $m \in \mathbb{N}$, let $\mathcal{F}_m$ denote the set of all m-ary Boolean functions.

The $\mathrm{QO}\Pi_1$-SAT problem takes as input a tuple (m, n, φ) where $m, n \in \mathbb{N}$ are written in unary, and φ is a Boolean combination of $x_1, \dots, x_m, y_1, \dots, y_n$ and $f(x_1, \dots, x_m)$. The input is accepted if and only if for all $f \in \mathcal{F}_m$ there are $x_1, \dots, x_m, y_1, \dots, y_n \in \{0, 1\}$ such that φ is a valid.

Our reduction from $\mathrm{QO}\Pi_1$-SAT to the p-universality problem of existential Büchi arithmetic follows an approach for showing coNEXP hardness of the Π_2-fragment of Presburger arithmetic [10, 12]. The main challenge is to show how to universally quantify over and suitably encode the doubly-exponential number of m-ary Boolean functions. Given $f \in \mathcal{F}_m$, we encode f via a number $z \in \mathbb{N}$ using a variant of Gödel encoding as follows:

$$f(b_0, \dots, b_{m-1}) = b \iff z \equiv b \bmod q,\ \text{ for all } q \in \mathbb{P} \cap [k^3, (k+1)^3), k = \textstyle\sum_{i=0}^{m-1} 2^i \cdot b_i\,. \quad (9)$$

Note that Ingham's theorem on prime gaps [15] guarantees that for sufficiently large $k \in \mathbb{N}$, there is at least one prime in the interval $[k^3, (k+1)^3)$. For technical convenience, to avoid adding a constant offset throughout our constructions, and as done in e.g. [10, 12], we apply Ingham's theorem as if it was true for all $k \in \mathbb{N}$.

▶ **Lemma 23.** *For every $f \in \mathcal{F}_m$ there is some $z \in \mathbb{N}$ encoding f as specified in* (9)*, and for all $i, j > 0$, p^i is a valid encoding if and only if p^j is a valid encoding.*

Proof. The first part immediately follows from the Chinese remainder theorem, and the second part follows from $b \in \{0, 1\}$ in (9). ◀

From [12] we can derive the existence of the following families of existential Presburger formulas polynomial-time computable in $m \in \mathbb{N}$ given in unary:

- $\Phi_m^{prime}(x)$ that evaluates to true if and only if $x < 2^m$ and $x \in \mathbb{P}$;
- $\Phi_m^{pow3}(x, y)$ that evaluates to true if and only if $x < 2^m$ and $y = x^3$;
- $\Phi_m^{mod}(x, y)$ that evaluates to true if and only if $y < 2^m$ and $x \equiv 0 \bmod y$; and
- $\Phi_m^{invalid}(x)$ that evaluates to true if and only if there are $q_1, q_2 \in \mathbb{P} \cap [k^3, (k+1)^3)$ for some $k < 2^m$ such that $(x \bmod q_1) \neq (x \bmod q_2)$ or $(x \bmod q_1) \notin \{0, 1\}$.

We define a family of existential formulas of Presburger arithmetic $\Phi_m^{fun}(\boldsymbol{x}, y, z)$ that hold if and only if $f(x_0, \dots, x_{m-1}) = y$, under the assumptions that z is a valid encoding of some $f \in \mathcal{F}_m$ as defined in Equation (9), $\boldsymbol{x} = (x_0, \dots, x_{m-1}) \in \{0, 1\}^m$ and $y \in \{0, 1\}$:

$$\begin{aligned}\Phi_m^{fun}(\boldsymbol{x}, y, z) \stackrel{\text{def}}{=} \exists x\, k\, k_0\, k_1 \colon k = \textstyle\sum_{i=0}^{m-1} 2^i \cdot x_i \land \Phi_m^{pow3}(k, k_0) \land \Phi_{m+1}^{pow3}(k+1, k_1) \\ \land k_0 \leq x < k_1 \land \Phi_{3m+1}^{prime}(x) \land \Phi_{3m+1}^{mod}(z - y, x)\,.\end{aligned}$$

For the final step of our reduction, given an instance $I = (m, n, \varphi)$ of $\mathrm{QO}\Pi_1$-SAT, denote by $\widetilde{\varphi}(\boldsymbol{x}, \boldsymbol{y}, y)$ the quantifier-free formula of Presburger arithmetic obtained from φ by replacing x_i and y_i by $x_i = 1$ and $y_i = 1$, respectively; $\neg x_i$ and $\neg y_i$ by $x_i = 0$ and $y_i = 0$, respectively; and $f(x_1, \dots, x_n)$ by $y = 1$ and $\neg f(x_1, \dots, x_n)$ by $y = 0$. We claim that I is a positive instance if and only if the following formula Ψ_p of existential Büchi arithmetic, with parametric base p and a single occurrence of a V_p function, is p-universal:

$$\begin{aligned}\Psi_p \stackrel{\text{def}}{=} \exists x\, \exists y \exists \boldsymbol{x}\, \exists \boldsymbol{y}\, \exists z \colon V_p(z) = z \land z > 1 \land 0 \leq x \leq 1 \land 0 \leq y \leq 1 \land \textstyle\bigwedge_{i \in [1,m]} 0 \leq x_i \leq 1 \\ \land \textstyle\bigwedge_{i \in [1,n]} 0 \leq y_i \leq 1 \land \Phi_m^{fun}(\boldsymbol{x}, y, z) \land \big(\Phi_m^{invalid}(z) \lor \widetilde{\varphi}(\boldsymbol{x}, \boldsymbol{y}, y)\big)\ ,\end{aligned}$$

where $\boldsymbol{x} = (x_1, \dots, x_m)$ and $\boldsymbol{y} = (y_1, \dots, y_n)$. Theorem 2 is now an immediate consequence of the following proposition.

▶ **Proposition 24.** *Let $I = (m, n, \varphi)$ be an instance of $QO\Pi_1$-Sat. Then I is a positive instance if and only if Ψ_p is p-universal.*

Proof. ($\Rightarrow$): By Lemma 23, any $f \in \mathcal{F}_m$ is encoded by some $p \in \mathbb{N}$. Choosing $z = p$ in Ψ_p and instantiating the x_j and y_j by those x_j and y_j making φ true for f, which exist by assumption, it follows that Ψ_p is p-universal.

($\Leftarrow$): Suppose that Ψ_p is p-universal. By Lemma 23, for every $f \in \mathcal{F}_m$ there is some valid encoding p of f, and by assumption Ψ_p evaluates to true in base p for some choice $z = p^i$. By Lemma 23, p is a valid encoding of f as well, and hence the same x_j and y_j that make Ψ_p for $z = p^i$ and *a fortiori* $z = p$ true also make φ true for f. Hence I is a positive instance. ◀

6 Conclusion

There remains the open problems to what extend the coNEXP upper bound for p-universality for the p-adic integers stated in Theorem 1 is tight. The coNEXP lower bound for p-universality for existential Büchi arithmetic together with the bounds on the ultimately periodic representation of the set of bases satisfying a given formula obtained in Proposition 12 gives strong evidence that, should it be possible to improve the coNEXP upper bound for the p-adic integers, a different approach not based on p-automata will likely be required. Likewise, we do not know whether the NEXP upper bounds for p-existence can be improved.

The coNEXP lower bound for p-universality for Büchi arithmetic crucially relies on the presence of disjunction and conjunctive as Boolean connectives. It would be interesting to better understand the complexity of the conjunctive fragments of the logics we consider, at present we cannot obtain any better upper bounds. In particular, Lechner et al. have shown that the restricted formulas obtained in Lipshitz' decidability proof are p-universal if and only if they are satisfied for all primes p singly exponentially bounded in the input [18].

Another interesting open problem is to settle the decidability status of p-universality for full Büchi arithmetic. Given that Büchi arithmetic does not have quantifier elimination [13], this problem will also likely require new approaches and techniques.

References

1 László Babai, Lance Fortnow, and Carsten Lund. Non-deterministic exponential time has two-prover interactive protocols. *Comput. Complex.*, 1:3–40, 1991. doi:10.1007/BF01200056.

2 Achim Blumensath and Erich Grädel. Automatic structures. In *Logic in Computer Science, LICS*, pages 51–62. IEEE Computer Society, 2000. doi:10.1109/LICS.2000.855755.

3 Véronique Bruyère. Entiers et automates finis. *Mémoire de fin d'études*, 1985.

4 Véronique Bruyère, Georges Hansel, Christian Michaux, and Roger Villemaire. Logic and p-recognizable sets of integers. *Bull. Belg. Math. Soc. Simon Stevin*, 1(2):191–238, 1994. doi:10.36045/bbms/1103408547.

5 J. Richard Büchi. Weak second-order arithmetic and finite automata. *Math. Logic Quart.*, 6(1-6):66–92, 1960. doi:10.1002/malq.19600060105.

6 Dmitry Chistikov and Christoph Haase. The taming of the semi-linear set. In *International Colloquium on Automata, Languages, and Programming, ICALP*, volume 55 of *LIPIcs*, pages 128:1–128:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPIcs.ICALP.2016.128.

7 Alan Cobham. On the base-dependence of sets of numbers recognizable by finite automata. *Math. Syst. Theory*, 3(2):186–192, jun 1969.

8 Andreas Dolzmann and Thomas Sturm. P-adic constraint solving. In *International Symposium on Symbolic and Algebraic Computation, ISSAC*, pages 151–158. ACM, 1999. doi:10.1145/309831.309894.

9 András Frank and Éva Tardos. An application of simultaneous Diophantine approximation in combinatorial optimization. *Combinatorica*, 7(1):49–65, 1987. doi:10.1007/BF02579200.

10 Erich Grädel. Dominoes and the complexity of subclasses of logical theories. *Ann. Pure Appl. Log.*, 43(1):1–30, 1989.

11 Florent Guépin, Christoph Haase, and James Worrell. On the existential theories of Büchi arithmetic and linear p-adic fields. In *Logic in Computer Science, LICS*, pages 1–10. IEEE, 2019. doi:10.1109/LICS.2019.8785681.

12 Christoph Haase. Subclasses of Presburger arithmetic and the weak EXP hierarchy. In *Computer Science Logic (CSL) and Logic in Computer Science (LICS), CSL-LICS*, pages 47:1–47:10. ACM, 2014. doi:10.1145/2603088.2603092.

13 Christoph Haase and Jakub Rózycki. On the expressiveness of Büchi arithmetic. In *Foundations of Software Science and Computation Structures, FOSSACS*, volume 12650 of *Lect. Notes Comp. Sci.*, pages 310–323. Springer, 2021. doi:10.1007/978-3-030-71995-1_16.

14 Bernard R. Hodgson. On direct products of automaton decidable theories. *Theor. Comput. Sci.*, 19(3):331 – 335, 1982. doi:10.1016/0304-3975(82)90042-1.

15 Albert E. Ingham. On the Estimation of $N(\sigma, T)$. *Q. J. Math.*, os-11(1):201–202, 01 1940. doi:10.1093/qmath/os-11.1.201.

16 Bakhadyr Khoussainov and Anil Nerode. Automatic presentations of structures. In *Logical and Computational Complexity, LCC*, volume 960 of *Lect. Notes Comp. Sci.*, pages 367–392. Springer, 1995. doi:10.1007/3-540-60178-3_93.

17 Aless Lasaruk and Thomas Sturm. Effective quantifier elimination for Presburger arithmetic with infinity. In *Computer Algebra in Scientific Computing, CASC*, volume 5743 of *Lect. Notes Comp. Sci.*, pages 195–212. Springer, 2009. doi:10.1007/978-3-642-04103-7_18.

18 Antonia Lechner, Joël Ouaknine, and James Worrell. On the complexity of linear arithmetic with divisibility. In *Logic in Computer Science, LICS*, pages 667–676. IEEE, 2015. doi:10.1109/LICS.2015.67.

19 Yuri V. Linnik. On the least prime in an arithmetic progression. I. The basic theorem. *Rec. Math. [Mat. Sbornik] N.S.*, 15(57):139–178, 1944.

20 L. Lipshitz. The diophantine problem for addition and divisibility. *T. Am. Math. Soc.*, 235:271–283, 1978. doi:10.2307/1998219.

21 Markus Lohrey. Model-checking hierarchical structures. *J. Comput. Syst. Sci.*, 78(2):461–490, 2012. doi:10.1016/j.jcss.2011.05.006.

22 Jaban Meher and M. Ram Murty. Ramanujan's proof of Bertrand's postulate. *Am. Math. Mon.*, 120(7):650–653, 2013.

23 Loïc Pottier. Minimal solutions of linear Diophantine systems: Bounds and algorithms. In *Rewriting Techniques and Applications, RTA*, volume 488 of *Lect. Notes Comp. Sci.*, pages 162–173. Springer, 1991. doi:10.1007/3-540-53904-2_94.

24 Zdeněk Sawa. Efficient construction of semilinear representations of languages accepted by unary nondeterministic finite automata. *Fundam. Informaticae*, 123(1):97–106, 2013. doi:10.3233/FI-2013-802.

25 Sylvain Schmitz. Complexity hierarchies beyond elementary. *ACM Trans. Comput. Theory*, 8(1):3:1–3:36, 2016. doi:10.1145/2858784.

26 Aleksei L Semenov. Presburgerness of predicates regular in two number systems. *Sib. Math. J.*, 18(2):289–300, 1977. doi:10.1007/BF00967164.

27 Joachim von zur Gathen and Malte Sieveking. A bound on solutions of linear integer equalities and inequalities. *P. Am. Math. Soc.*, 72(1):155–158, 1978. doi:10.2307/2042554.

28 Herbert S. Wilf. A circle-of-lights algorithm for the "money-changing problem". *Am. Math. Mon.*, 85(7):562–565, 1978. doi:10.1080/00029890.1978.11994639.

29 Pierre Wolper and Bernard Boigelot. On the construction of automata from linear arithmetic constraints. In *Tools and Algorithms for the Construction and Analysis of Systems, TACAS*, pages 1–19, 2000. doi:10.1007/3-540-46419-0_1.

30 Triantafyllos Xylouris. *Über die Nullstellen der Dirichletschen L-Funktionen und die kleinste Primzahl in einer arithmetischen Progression.* PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2011. URL: `http://hdl.handle.net/20.500.11811/5074`.

A Missing proofs from Section 2

▶ **Proposition 6.** *For all $w = \boldsymbol{u}_0\boldsymbol{u}_1\cdots \in (\Sigma_d^p)^\omega$, $\mathbf{A}\cdot[\![w]\!]^\omega = \boldsymbol{c}$ iff there is $\boldsymbol{r} \in \mathbb{Z}^n$ and a strictly ascending sequence $(\lambda_i)_{i\in\mathbb{N}}$ such that $\boldsymbol{r} \xrightarrow{\boldsymbol{u}_{\lambda_0-1}\cdots\boldsymbol{u}_0}_{\mathbf{A},p} \boldsymbol{c}$ and $\boldsymbol{r} \xrightarrow{\boldsymbol{u}_{\lambda_{j+1}}\cdots\boldsymbol{u}_{\lambda_j}}_{\mathbf{A},p} \boldsymbol{r}$ for all $j \in \mathbb{N}$.*

Proof. Simple reformulation of the fact that $\mathbf{A}\cdot[\![w]\!]^\omega = \boldsymbol{c}$ holds if and only if for every $k \in \mathbb{N}$ there is $\boldsymbol{v} \in \mathbb{Z}^n$ such that $\mathbf{A}\cdot[\![\boldsymbol{u}_{k-1}\boldsymbol{u}_{k-2}\dots\boldsymbol{u}_0]\!]^* + \boldsymbol{v}\cdot p^k = \boldsymbol{c}$, where $w = \boldsymbol{u}_0\boldsymbol{u}_1,\dots$. Indeed, as the set of live states of the p automaton for $\mathbf{A}\cdot\boldsymbol{u} = \boldsymbol{c}$ is finite (by Proposition 4), in the infinite sequence $\boldsymbol{v}_0, \boldsymbol{v}_1, \boldsymbol{v}_2, \dots$ where $\mathbf{A}\cdot[\![\boldsymbol{u}_{k-1}\boldsymbol{u}_{k-2}\dots\boldsymbol{u}_0]\!]^* + \boldsymbol{v}_k\cdot p^k = \boldsymbol{c}$ for every $k \in \mathbb{N}$, there must be a state $\boldsymbol{r}$ that appears infinitely often. ◀

B Missing proofs from Section 3

Below, given a semi-linear set $M = \bigcup_{i\in I} L(B_i, P_i)$, we define $\|M\| \stackrel{\text{def}}{=} \max_{i\in I}(\|B_i\|_\infty, \|P_i\|_\infty)$.

▶ **Lemma 14.** *Let $\mathbf{A} \in \mathbb{Z}^{n\times d}$, $\boldsymbol{s}, \boldsymbol{t} \in \mathbb{Z}^n$ and let T be a linear system of m inequalities. The set $[\![\boldsymbol{s} \to_T \boldsymbol{t}]\!]_{\mathbf{A}}$ is an ultimately periodic set with period and threshold bounded by $U^{\mathcal{O}(k\log k)}$, where $k = n + d + m$ and $U = \max(2, \|\mathbf{A}\|_\infty, \|\boldsymbol{s} \to_T \boldsymbol{t}\|_\infty)$.*

Proof. Recall that $[\![\boldsymbol{s} \to_T \boldsymbol{t}]\!]_{\mathbf{A}}$ is the set of $z \in \mathbb{N}$ for which there is $\boldsymbol{x} = (x_1, \dots, x_d) \in [\![T]\!]_{\geq 0}$ satisfying the following system S of linear inequalities:

$$\begin{cases} \boldsymbol{s}\cdot z + \mathbf{A}\cdot\boldsymbol{x} = \boldsymbol{t} \\ z \geq 2 \\ x_i \leq z - 1 & \text{for every } i \in [1, d]. \end{cases}$$

As usual, when restricting a system to its non-negative solutions, we can replace inequalities with equalities by introducing slack variables, obtaining the system S':

$$\begin{cases} \boldsymbol{s}\cdot z + \mathbf{A}\cdot\boldsymbol{x} = \boldsymbol{t} \\ z - z' = 2 \\ x_i + x_i' = z - 1 & \text{for every } i \in [1, d]. \end{cases}$$

where $z', x_1', \dots, x_d'$ are $(d+1)$ variables that shall be interpreted with non-negative integers. It is easy to verify that the set of solutions in $[\![S]\!]_{\geq 0}$ can be characterised by considering the set $[\![S']\!]_{\geq 0}$ and projecting away the dimensions relatives to the slack variables $z', x_1', \dots, x_d'$. Notice that S' has $2\cdot(d+1)$ variables and $n+d+1$ rows. A similar treatment can be applied to the system of inequalities T: by introducing at most m new slack variables, we obtain a system of equalities T' with m rows and $d + m$ variables. We apply Proposition 7, on the system made of S' and T' and conclude that $[\![S']\!]_{\geq 0} \cap [\![T']\!]_{\geq 0} = L(B, P)$ where

- $\|B\|_\infty \leq ((2\cdot(d+1) + (d+m) + 2)\cdot U + 1)^{(n+d+1)+(d+m)}$,
- $\|P\|_\infty \leq ((2\cdot(d+1) + (d+m))\cdot U + 1)^{(n+d+1)+(d+m)}$.

Hence, $\|L(B,P)\| \leq U^{\mathcal{O}(k\log k)}$. Projecting $L(B,P)$ on the variable z yields $[\![\boldsymbol{s} \to_T \boldsymbol{t}]\!]_{\mathbf{A}}$. So, $[\![\boldsymbol{s} \to_T \boldsymbol{t}]\!]_{\mathbf{A}} = L(C, Q) \subseteq \mathbb{N}$ with $\|L(C,Q)\| \leq U^{\mathcal{O}(k\log k)}$. We now change representation: by Proposition 8, we characterise $[\![\boldsymbol{s} \to_T \boldsymbol{t}]\!]_{\mathbf{A}}$ as an ultimately periodic set with period $p \leq \gcd Q \leq U^{\mathcal{O}(k\log k)}$ and threshold $t \leq \|C\|_\infty + \|Q\|_\infty^2 \leq U^{\mathcal{O}(k\log k)}$. ◀

▶ **Lemma 18.** *The set $\oplus(L_0, \dots, L_r)$ contains all length values generated by $\mathcal{G}$. One can construct in time $\mathcal{O}(r^2) \cdot U^{\mathcal{O}(n \cdot r)}$ a representation of $\oplus(L_0, \dots, L_r)$ as a semi-linear set $\bigcup_{k \in K} L(\boldsymbol{b}_k, P_k)$, where $\#K \leq U^{\mathcal{O}(n \cdot r)}$, $\#P \leq r+1$, $\|P\|_\infty \leq U^{2n}$ and $\|\boldsymbol{b}_k\|_\infty \leq \mathcal{O}(r \cdot U^{4n})$.*

Proof. It is relatively straightforward to see that $\oplus(L_0, \dots, L_r)$ corresponds to the set of length values generated by $\mathcal{G}$. First, let $(v_0, \dots, v_r) \in \mathbb{N}^{r+1}$ be a length value generated by $\mathcal{G}$, which by definition means that $\mathcal{G}$ has a path of the following form

$$\boldsymbol{t}_{r+1} \to_{U_r} \boldsymbol{s}_r \to_{W_r}^{v_r-(v_{r-1}+1)} \boldsymbol{t}_r \to_{U_{r-1}} \boldsymbol{s}_{r-1} \;\dots\; \boldsymbol{s}_1 \to_{W_1}^{v_1-(v_0+1)} \boldsymbol{t}_1 \to_{U_0} \boldsymbol{s}_0 \to_{W_0}^{v_0} \boldsymbol{t}_0 = \boldsymbol{c}.$$

Directly by definition of $G_j(\boldsymbol{s}_j, \boldsymbol{t}_j)$ and its related unary NFA, we have that $v_0 \in L_0$ and for every $j \in [1, r]$, $v_j - (v_j + 1) \in L_j$. By definition of $\oplus(L_1, \dots, L_r)$,

$$\begin{pmatrix} v_0 \\ 1 + v_0 + v_1 - (v_0 + 1) \\ \dots \\ i + v_0 + \sum_{j \in [1,i]} (v_j - (v_{j-1} + 1)) \\ \dots \\ r + v_0 + \sum_{j \in [1,r]} (v_j - (v_{j-1} + 1)) \end{pmatrix} = \begin{pmatrix} v_0 \\ v_1 \\ \dots \\ v_i \\ \dots \\ v_r \end{pmatrix} \in \oplus(L_1, \dots, L_r)$$

Conversely, suppose $(v_0, \dots, v_r) \in \oplus(L_1, \dots, L_r)$. This implies that there are $\ell_0, \dots, \ell_r$ such that, for every $j \in [0, r]$, $\ell_j \in L_j$ and $v_j = j + \sum_{i \in [0,j]} \ell_j$. By definition of L_j, we conclude that $\mathcal{G}$ has a path of the following form

$$\boldsymbol{t}_{r+1} \to_{U_r} \boldsymbol{s}_r \to_{W_r}^{\ell_r} \boldsymbol{t}_r \to_{U_{r-1}} \boldsymbol{s}_{r-1} \;\dots\; \boldsymbol{s}_1 \to_{W_1}^{\ell_1} \boldsymbol{t}_1 \to_{U_0} \boldsymbol{s}_0 \to_{W_0}^{\ell_0} \boldsymbol{t}_0 = \boldsymbol{c}.$$

We have $\ell_0 = v_0$ and, given $j \in [1, r]$, from $v_j = j + \sum_{i \in [0,j]} \ell_j = 1 + v_{j-1} + \ell_j$ we conclude that $\ell_j = v_j - (v_{j-1} + 1)$. Hence $(v_0, \dots, v_r)$ is a length value generated by $\mathcal{G}$.

Let us now show how to construct a semi-linear representation of $\oplus(L_0, \dots, L_r)$. Recall that, by Proposition 17, for every $i \in [0, r]$ we have $L_i = \bigcup_{j \in J_i} L(b_j, p_j)$ with $\#J_i \leq \mathcal{O}(U^{4n})$, $b_j \leq (2 \cdot U^{2n} + 1) \cdot U^{2n}$ and $p_j \leq U^{2n}$. To compute $\oplus(L_0, \dots, L_r)$ we iterate over all $(j_0, \dots, j_r) \in J_0 \times \dots \times J_r$ and construct a linear set representation for

$$\oplus(L(b_{j_0}, p_{j_0}), \dots, L(b_{j_r}, p_{j_r})) \overset{\text{def}}{=} \left\{ \begin{pmatrix} \ell_0 \\ 1 + \ell_0 + \ell_1 \\ \dots \\ r + \sum_{i=0}^{r} \ell_i \end{pmatrix} \in \mathbb{N}^{r+1} : \forall i \in [0, r],\ \ell_i \in L(b_{j_i}, p_{j_i}) \right\}.$$

The set $\oplus(L_0, \dots, L_r)$ is then the union over all such linear sets. Hence, consider $(j_0, \dots, j_r) \in J_0 \times \dots \times J_r$ and the linear sets $L(b_{j_0}, p_{j_0}), \dots, L(b_{j_r}, p_{j_r})$. Given $i \in [0, r]$, we write $\boldsymbol{q}_i \in \mathbb{N}^{r+1}$ for the vector having zero in the first $i-1$ components, and p_{j_i} in the last $r+1-i$ components. Moreover, let $\boldsymbol{b} \in \mathbb{N}^{r+1}$ be the vector with ith component set to $i + \sum_{k=0}^{i} b_{j_i}$. It is easy to see that $\oplus(L(b_{j_0}, p_{j_0}), \dots, L(b_{j_r}, p_{j_r})) = L(\boldsymbol{b}, \{\boldsymbol{q}_0, \dots, \boldsymbol{q}_r\})$. Due to the bound on each $L(b_{j_i}, p_{j_i})$, one can compute $L(\boldsymbol{b}, \{\boldsymbol{q}_0, \dots, \boldsymbol{q}_r\})$ in time $\mathcal{O}(r^2 \cdot \max_{i \in [0,r]}(\langle b_{j_i} \rangle, \langle p_{j_i} \rangle)) \leq \mathcal{O}(n \cdot r^2 \cdot \log U)$. Moreover, $\|\boldsymbol{q}_i\|_\infty \leq U^{2n}$ and $\|\boldsymbol{b}\|_\infty \leq \mathcal{O}(r \cdot U^{4n})$. When considering all $U^{\mathcal{O}(n \cdot r)}$ tuples, a semi-linear representation of $\oplus(L_0, \dots, L_r)$ can then be constructed in time $\mathcal{O}(r^2) \cdot U^{\mathcal{O}(n \cdot r)}$. ◀

▶ **Lemma 19.** *Let ψ be a formula of the form given in* (5).

- *Given $\mathcal{G} \in \mathbb{G}(\psi)$, if $\mathcal{G} \vdash \mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$ then $[\![\mathcal{G}]\!]_{\mathbf{A}} \subseteq \mathcal{B}(\psi)$.*
- *For every $p \in \mathcal{B}(\psi)$, there is $\mathcal{G} \in \mathbb{G}(\psi)$ such that $\mathcal{G} \vdash \mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$ and $p \in [\![\mathcal{G}]\!]_{\mathbf{A}}$.*

Proof. To show this result, we simply adapt the proof of Lemma 15. Let $\mathbb{G}$ be the set of support graphs $\mathcal{G}$ in $\mathbb{G}(\psi)$ such that $\mathcal{G} \vdash \mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$. The lemma equivalently states that $\mathcal{B}(\psi) = \bigcup_{\mathcal{G} \in \mathbb{G}} [\![\mathcal{G}]\!]_{\mathbf{A}}$. Below, we refer to the first and second points in the lemma as the two inclusions $\supseteq$ and $\subseteq$ of this equality.
($\supseteq$): Let $\mathcal{G}$ be a support graph in $\mathbb{G}$, and consider $p \in [\![\mathcal{G}]\!]_{\mathbf{A}}$. Moreover, let $(v_0, \dots, v_r) \in \mathbb{N}^{r+1}$ be a length value generated by $\mathcal{G}$ such that $\mathbf{B} \cdot \boldsymbol{c} \geq \boldsymbol{d} \wedge \bigwedge_{i \in [0,r]} x_i = v_i$ is satisfiable. This means that $\mathcal{G}$ has a path of the form

$$\boldsymbol{t}_{r+1} \rightarrow_{U_r} \boldsymbol{s}_r \rightarrow_{W_r}^{v_r-(v_{r-1}+1)} \boldsymbol{t}_r \rightarrow_{U_{r-1}} \boldsymbol{s}_{r-1} \ \dots \ \boldsymbol{s}_1 \rightarrow_{W_1}^{v_1-(v_0+1)} \boldsymbol{t}_1 \rightarrow_{U_0} \boldsymbol{s}_0 \rightarrow_{W_0}^{v_0} \boldsymbol{t}_0 = \boldsymbol{c}$$

Moreover, as $\mathcal{G} \in \mathbb{G}(\psi)$, there is $\boldsymbol{r} \in \mathbb{Z}^n$ such that $\mathcal{G}$ has a path from $\boldsymbol{r}$ to $\boldsymbol{t}_{r+1}$ as well as a path from $\boldsymbol{r}$ to itself. We recall that, by definition, for every edge $\boldsymbol{s} \rightarrow_T \boldsymbol{t}$ of $\mathcal{G}$ there is $\boldsymbol{u} \in \Sigma_p^d$ such that $\boldsymbol{s} \xrightarrow{\boldsymbol{u}}_{\mathbf{A},p} \boldsymbol{t}$ and $\boldsymbol{u} \in [\![T]\!]$. We conclude that there is an infinite word w such that $w = w_0 \cdot \boldsymbol{v}_0 \cdot w_1 \cdot \ldots \cdot w_r \cdot \boldsymbol{v}_r \cdot \overline{w}$, where

1. $\boldsymbol{v}_0, \dots, \boldsymbol{v}_r \in \Sigma_p^d$, and for all $j \in [0, r]$, the prefix $w_0 \cdot \boldsymbol{v}_0 \cdot \ldots \cdot w_j \in (\Sigma_p^d)^*$ of w has length v_j,
2. $\boldsymbol{t}_{r+1} \xrightarrow{\boldsymbol{v}_r}_{\mathbf{A},p} \boldsymbol{s}_r \ \dots \ \boldsymbol{s}_1 \xrightarrow{(w_1)^R}_{\mathbf{A},p} \boldsymbol{t}_1 \xrightarrow{\boldsymbol{v}_0}_{\mathbf{A},p} \boldsymbol{s}_0 \xrightarrow{(w_0)^R}_{\mathbf{A},p} \boldsymbol{t}_0 = \boldsymbol{c}$,
3. $\overline{w} = \boldsymbol{u}_0 \boldsymbol{u}_1 \dots$ is an infinite suffix of w for which there is an infinite sequence $\lambda_0 < \lambda_1 < \dots$ such that $\boldsymbol{r} \xrightarrow{\boldsymbol{u}_{\lambda_0 - 1} \dots \boldsymbol{u}_0}_{\mathbf{A},p} \boldsymbol{c}$ and for all $j \in \mathbb{N}$, $\boldsymbol{r} \xrightarrow{\boldsymbol{u}_{\lambda_{j+1}} \dots \boldsymbol{u}_{\lambda_j}}_{\mathbf{A},p} \boldsymbol{r}$,
4. By definition of $U_0, \dots, U_m$ and $W_0, \dots, W_{m+1}$, given $(i, j) \in J$, so that $v_p(u_i) = x_j$ appears in ψ, we have that $w_0 \cdot \boldsymbol{v}_0 \cdot \ldots \cdot w_j$ projected on the encoding of u_i is in $\{0\}^*$, and v_j projected on the encoding of u_i is in $[1, p-1]$.

Therefore, ψ admits a solution $([\![w]\!]^\omega, \boldsymbol{x})$, where in $\boldsymbol{x}$ for every $i \in [0, r]$ we have $x_i = v_i$. Indeed, from points 2 and 3, and by Proposition 6, $\mathbf{A} \cdot [\![w]\!]^\omega = \boldsymbol{c}$. From points 1 and 4, given $(i, j) \in J$, the *lsd* encoding of u_i is such that $v_p(u_i) = v_j$, hence $v_p(u_i) = x_j$. Lastly, by definition of $(v_0, \dots, v_r)$ we have $x_0 < x_1 < \dots < x_r$, and we know that $\mathbf{B} \cdot \boldsymbol{c} \geq \boldsymbol{d} \wedge \bigwedge_{i \in [0,r]} x_i = v_i$ is satisfiable. This means that ψ is satisfiable with respect to the base p, i.e. $p \in \mathcal{B}(\psi)$.
($\subseteq$): Follows conversely from the other inclusion. Briefly, suppose ψ satisfiable with respect to the base p. Consider a word $w \in (\Sigma_p^d)^\omega$ and $\boldsymbol{x} \in \mathbb{Z}^e$ such that $([\![w]\!]^\omega, \boldsymbol{x})$ is a solution of ψ. As already said in the body of the paper, this means that w must have a prefix of the form $w_0 \cdot \boldsymbol{v}_0 \cdot w_1 \cdot \ldots \cdot w_r \cdot \boldsymbol{v}_r \cdot w_{r+1}$ such that $\boldsymbol{v}_0, \dots, \boldsymbol{v}_r \in \Sigma_p^d$, the word $w_0 \in (\Sigma_p^d)^*$ has length x_0, each $w_i \in (\Sigma_p^d)^*$ with $i \in [1, r]$ has length $x_i - (x_{i-1} + 1) \geq 0$,

$$\boldsymbol{r} = \boldsymbol{s}_{r+1} \xrightarrow{(w_{r+1})^R}_{\mathbf{A},p} \boldsymbol{t}_{r+1} \xrightarrow{\boldsymbol{v}_r}_{\mathbf{A},p} \boldsymbol{s}_r \ \dots \ \boldsymbol{s}_1 \xrightarrow{(w_1)^R}_{\mathbf{A},p} \boldsymbol{t}_1 \xrightarrow{\boldsymbol{v}_0}_{\mathbf{A},p} \boldsymbol{s}_0 \xrightarrow{(w_0)^R}_{\mathbf{A},p} \boldsymbol{t}_0 = \boldsymbol{c}$$

and $\boldsymbol{r}$ is a live state of $\mathcal{A}_p^*(S)$ for which the ω-regular condition of Proposition 6 is satisfied. Moreover, for every $j \in [0, r]$ the values of $\boldsymbol{v}_j$ for the variables $u_1, \dots, u_{\#I}$ satisfy the system U_j, whereas at each position of the word w_j ($j \in [0, r+1]$) shall satisfy the system W_j.

Let $\mathcal{G}$ be the support graph with edges $\boldsymbol{t}_1 \rightarrow_{U_0} \boldsymbol{s}_0, \dots, \boldsymbol{t}_{r+1} \rightarrow_{U_j} \boldsymbol{s}_r$ together with $\boldsymbol{i} \rightarrow_{W_j} \boldsymbol{i}'$, for every $j \in [0, r+1]$ and every two states $\boldsymbol{i}, \boldsymbol{i}'$ such that $\boldsymbol{i} \rightarrow_{\mathbf{A},p} \boldsymbol{i}'$ appears in the path going from $\boldsymbol{s}_j$ to $\boldsymbol{t}_j$, and $\boldsymbol{i} \rightarrow_{W_{r+1}} \boldsymbol{i}'$, for every two states $\boldsymbol{i}, \boldsymbol{i}'$ such that $\boldsymbol{i} \rightarrow_{\mathbf{A},p} \boldsymbol{i}'$ appears in the path going from $\boldsymbol{r}$ to itself. By definition, $\mathcal{G} \in \mathbb{G}(\psi)$ and $p \in [\![\mathcal{G}]\!]_{\mathbf{A}}$. Moreover, $(x_0, \dots, x_r)$ is a length value generated by $\mathcal{G}$, which entails $\mathcal{G} \vdash \mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$. ◀

▶ **Lemma 20.** *Let ψ be as in* (5)*. The set $\mathcal{B}(\psi)$ is ultimately periodic, with threshold bounded by $U^{\mathcal{O}(k \log k)}$ and period bounded by $U^{\mathcal{O}(\ell \cdot k \log k)}$, with $k = n + 3d$ and $\ell = (r+2) \cdot U^{4n+1}$.*

Algorithm 1 Procedure for deciding satisfiability of a formula ψ of the form in (5).

```
 1: function SAT($\psi$)
 2:     $Q \leftarrow [-U,U]^n$ where $U = \max(\|\mathbf{A}\|_{1,\infty}, \|\boldsymbol{c}\|_\infty)$
 3:     for $(\boldsymbol{s}_0,\boldsymbol{t}_0,\dots,\boldsymbol{s}_{r+1},\boldsymbol{t}_{r+1}) \in Q^{2(r+1)}$ with $\boldsymbol{t}_0 = \boldsymbol{c}$ do
 4:         $L_0,\dots,L_r \leftarrow \emptyset$
 5:         for $j \in [0,r+1]$ do                ▷ below, $W_j$ and $U_j$ defined as in Section 3
 6:             $\mathcal{A}_j \leftarrow (Q,\delta,\{\boldsymbol{s}_j\},\{\boldsymbol{t}_j\})$ where $\delta = \{(\boldsymbol{s},\boldsymbol{t}) : p \in [\![\boldsymbol{s} \to_{W_j} \boldsymbol{t}]\!]_{\mathbf{A}}, \boldsymbol{s},\boldsymbol{t} \in Q\}$.
 7:             if $j > 0$ then check $p \in [\![\boldsymbol{t}_j \to_{U_j} \boldsymbol{s}_{j-1}]\!]_{\mathbf{A}}$
 8:             check if $\mathcal{A}_j$ has a path form $\boldsymbol{s}_j$ to $\boldsymbol{t}_j$ (i.e. $\mathcal{L}(\mathcal{A}_j) \neq \emptyset$)
 9:             if $j = r+1$ then check if $\mathcal{A}_j$ has a non-empty path form $\boldsymbol{s}_j$ to itself
10:             if $j \neq r+1$ then $L_j \leftarrow L$ where $L$ defined as in Proposition 17 w.r.t. $\mathcal{A}_j$
11:         end for
12:         if $\oplus(L_0,\dots,L_r) \cap [\![\mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}]\!] \neq \emptyset$ then return true    ▷ $\oplus(L_0,\dots,L_r)$ as in (8)
13:     end for
14:     return false
```

Proof. Analogous to the proof of Lemma 16. Consider the family of support graphs $\mathbb{G}(\psi)$. Since $\mathbb{G}(\psi)$ is finite, according to Lemma 19 we have $\mathcal{B}(\psi) = \bigcup_{\mathcal{G} \in G} [\![\mathcal{G}]\!]_{\mathbf{A}}$ for some $G \subseteq \mathbb{G}(\psi)$. Every edge $\boldsymbol{s} \to_T \boldsymbol{t}$ of a support graph $\mathcal{G} \in G$ is such that $\|\boldsymbol{s}\|_\infty, \|\boldsymbol{t}\|_\infty \leq \max(\|\mathbf{A}\|_{1,\infty}, \|\boldsymbol{c}\|_\infty)$ and T is a system among $U_0,\dots,U_r,W_0,\dots,W_{r+1}$. Hence, all the graphs in G are built from a set E of $(2 \cdot \max(\|\mathbf{A}\|_{1,\infty}, \|\boldsymbol{c}\|_\infty))^{2n} \cdot (2r+3) \leq \ell$ edges. Each possible linear system T labelling an edge in E has at most $2d$ inequalities, with coefficients and constants in $\{0,1\}$. By Lemma 14, each edge $e \in E$ is such that $[\![e]\!]_{\mathbf{A}}$ is an ultimately periodic set with period and threshold bounded by $U^{\mathcal{O}(k \log k)}$, where $k = n + 3d$. By Proposition 9, $\mathcal{B}(\psi) = \bigcup_{\mathcal{G} \in G} [\![\mathcal{G}]\!]_{\mathbf{A}}$ is an ultimately periodic set with threshold in $U^{\mathcal{O}(k \log k)}$ and period in $U^{\mathcal{O}(\ell \cdot k \log k)}$. ◀

C Missing proofs from Section 4

▶ **Theorem 21.** *Let Φ be an existential formula of linear arithmetic constraints over p-adic integers (resp. Büchi arithmetic) with parametric base p. Then satisfiability of Φ with respect to a given value $p \in \mathbb{P}$ (resp. $p \geq 2$) can be decided in time $2^{\mathcal{O}(\langle \Phi \rangle^3)} \cdot \mathcal{O}(\langle p \rangle)$.*

Proof. Consider an existential formula Φ of linear arithmetic constraints over p-adic integers, with parametric base q, and a concrete value $p \in \mathbb{P}$ for q. As done in Proposition 12, in time exponential in $\langle \Phi \rangle$ we manipulate Φ and obtain a disjunctive normal form where each of the $2^{\mathcal{O}(\langle \Phi \rangle \log \langle \Phi \rangle)}$ disjuncts are of size $\mathcal{O}(\langle \Phi \rangle)$ and the form in (5). We then iterate through each disjunct ψ of Φ, calling the procedure $\textsc{sat}(\psi)$ described in Algorithm 1. If the procedure returns true for some disjunct ψ, then Φ is satisfiable, otherwise it is unsatisfiable. Let us fix a disjunct ψ of Φ. We argue that $\textsc{sat}(\psi)$ is a decision procedure for the satisfiability problem of ψ that runs in time $2^{\mathcal{O}(\langle \psi \rangle^3)} \cdot \mathcal{O}(\langle p \rangle)$. As Φ has $2^{\mathcal{O}(\langle \Phi \rangle \log \langle \Phi \rangle)}$ many disjuncts of size $\mathcal{O}(\langle \Phi \rangle)$, this is sufficient to establish Theorem 21.

Let us sketch the correctness of the algorithm, which essentially follows by replaying some of the arguments that led to Proposition 12. Indeed, line 3 iterates through all possible tuples of intermediate live states, as done when considering the set of automatic support graphs $\mathbb{G}(\psi)$. Let $(\boldsymbol{s}_0,\boldsymbol{t}_0,\dots,\boldsymbol{s}_{r+1},\boldsymbol{t}_{r+1})$ be one of these tuples. For every $j \in [0,r+1]$, the algorithm considers the unary NFA $\mathcal{A}_j$ (line 6) whose arrows $(\boldsymbol{s},\boldsymbol{t})$ corresponds to edges $\boldsymbol{s} \to_{W_j} \boldsymbol{t}$ of automatic support graphs for which $p \in [\![\boldsymbol{s} \to_{W_j} \boldsymbol{t}]\!]_{\mathbf{A}}$. The fundamental relation between $\mathcal{A}_j$ and the p-automaton $\mathcal{A}_p^*(S)$ for the system $S\colon \mathbf{A} \cdot \boldsymbol{u} = \boldsymbol{c}$ is as follows:

$$(\boldsymbol{s}, \boldsymbol{t}) \in \delta \text{ if and only if there is } \boldsymbol{u} \in [0, p-1]^n \text{ such that } \boldsymbol{s} \xrightarrow{\boldsymbol{u}}_{\mathbf{A},p} \boldsymbol{t} \text{ and } \boldsymbol{u} \in [\![W_j]\!].$$

Therefore, if the checks performed in lines 7 and 8 are satisfied for all $j \in [0, r+1]$ (else, a new tuple in $Q^{2(r+1)}$ is considered), we conclude that $\mathcal{A}_p^*(S)$ contains a path of the form in (6). According to Proposition 6, line 9 then tests for the existence of a non-empty path in $\mathcal{A}_{r+1}$ (equivalently, $\mathcal{A}_p^*(S)$) going from $\boldsymbol{s}_{r+1}$ to itself. If such a path is found, then there is an infinite word $w \in (\Sigma_p^d)^\omega$ and $\boldsymbol{x} \in \mathbb{N}^e$ such that $([\![w]\!]^*, \boldsymbol{x})$ is a solution for the subformula $\mathbf{A} \cdot \boldsymbol{u} = \boldsymbol{c} \wedge \bigwedge_{(i,j) \in J} v_p(u_i) = x_j \wedge \bigwedge_{j \in [1,r]} x_{j-1} < x_j$ of ψ. Lines 10 and 12 provides further analysis needed to check for the satisfaction of the system $\mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$. Again following what done for Proposition 12, line 10 computes the set L_j of lengths of words accepted by $\mathcal{A}_j$, that are then combined in the set $\oplus(L_0, \dots, L_r)$. Note that $\oplus(L_0, \dots, L_r)$ is empty if so is one set among $L_0, \dots, L_r$. If one element of $\oplus(L_0, \dots, L_r)$ satisfies $\mathbf{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$, then ψ is satisfiable, and the procedure returns true (line 12).

Let us discuss the time complexity of $\textsc{sat}(\psi)$. First of all, the running time of the loop in line 5 is in $2^{\mathcal{O}(\langle\psi\rangle^2)} \cdot \mathcal{O}(\langle p \rangle)$. Indeed, notice that given $\boldsymbol{s}, \boldsymbol{t} \in Q$ and a system $T \in \{U_j, W_j, W_{r+1} : j \in [0, r]\}$, the membership problem $p \in [\![\boldsymbol{s} \to_T \boldsymbol{t}]\!]_{\mathbf{A}}$ correspond to the non-emptiness problem $[\![R]\!] \neq \emptyset$ of the linear system $R\colon \boldsymbol{t} = \boldsymbol{s} \cdot p + \mathbf{A} \cdot \boldsymbol{x} \wedge \|\boldsymbol{x}\| < p \wedge \boldsymbol{x} \in [\![T]\!]$, which by Proposition 22 can be decided in time $\langle\psi\rangle^{\mathcal{O}(\langle\psi\rangle)} \cdot \mathcal{O}(\langle p \rangle)$. Hence, by iterating over all pairs $(\boldsymbol{s}, \boldsymbol{t}) \in Q^2$, one constructs the automaton $\mathcal{A}_j$ in time $2^{\mathcal{O}(\langle\psi\rangle^2)} \cdot \mathcal{O}(\langle p \rangle)$ (line 6). Line 7 runs in time $\langle\psi\rangle^{\mathcal{O}(\langle\psi\rangle)} \cdot \mathcal{O}(\langle p \rangle)$. Lines 8 and 9, can be implemented with a depth-first search on the automaton $\mathcal{A}_j$, which takes time bounded by the cardinality of the transition relation δ, that is bounded by $2^{\mathcal{O}(\psi)^2}$. Directly from Proposition 17, line 10 can be performed in time $2^{\mathcal{O}(\langle\psi\rangle^2)}$. Let us now look at line 12. By Lemma 18, computing $\oplus(L_0, \dots, L_r)$ as a semi-linear set can be done in time $2^{\mathcal{O}(\langle\psi\rangle^3)}$. This set is of the form $\bigcup_{k \in K} L(\boldsymbol{b}_k, P_k) \subseteq \mathbb{N}^{r+1}$, with $\#K \leq 2^{\mathcal{O}(\langle\psi\rangle^3)}$, $\|\boldsymbol{b}_k\|_\infty, \|P_k\|_\infty \leq 2^{\mathcal{O}(\langle\psi\rangle^2)}$ and $\#P_k \leq r+1$. Hence, to check for the non-emptiness of the intersection in line 12, it is sufficient to iterate over all $k \in K$ and check for the satisfiability of the system $\boldsymbol{y} = \boldsymbol{b_k} + \boldsymbol{P}_k \cdot \boldsymbol{\lambda} \wedge \boldsymbol{B} \cdot \boldsymbol{x} \geq \boldsymbol{d}$, where $\boldsymbol{y}$ is a subset of the variables appearing in $\boldsymbol{x}$. By Proposition 22, this can be done in time $\langle\psi\rangle^{\mathcal{O}(\langle\psi\rangle)}$. Overall, line 12 can be evaluated in $2^{\mathcal{O}(\langle\psi\rangle^3)}$, and hence the running time for the body of the for loop of line 3 is bounded by $2^{\mathcal{O}(\langle\psi\rangle^3)} \cdot \mathcal{O}(\langle p \rangle)$. As this loop iterates over $(2U)^{2(r+1)} \leq 2^{\mathcal{O}(\langle\psi\rangle^2)}$ tuples, the running time of $\textsc{sat}(\psi)$ is in $2^{\mathcal{O}(\langle\psi\rangle^3)} \cdot \mathcal{O}(\langle p \rangle)$. ◀

In order to establish Theorem 21 for the case of existential Büchi arithmetic, it is sufficient to bring Φ into disjunctive normal form with disjuncts of the form in Equation (2), and call on each disjunct ψ the procedure $\textsc{sat}(\psi)$ subject to the following updates: (I) reflecting (3), the iteration on line 3 is on tuples such that $\boldsymbol{s}_0 = \boldsymbol{0}$ and $\boldsymbol{t}_{r+1} = \boldsymbol{c}$, and the test in line 7 becomes $j > 0$ and $p \notin [\![\boldsymbol{t}_{j-1} \to_{U_j} \boldsymbol{s}_j]\!]_{\mathbf{A}}$; (II) the systems U_j and W_j in lines 6 and 7 are defined as introduced in Section 3 and (III) lines 4, 9, 10 and 12 are removed. The proof of Theorem 21 can be easily updated accordingly.