Rational Subsets of Baumslag-Solitar Groups

Michaël Cadilhac
DePaul University, Chicago, IL, USA
michael@cadilhac.name

Dmitry Chistikov
Centre for Discrete Mathematics and its Applications (DIMAP) & Department of Computer Science, University of Warwick, Coventry, UK
d.chistikov@warwick.ac.uk

Georg Zetzsche
Max Planck Institute for Software Systems (MPI-SWS), Kaiserslautern, Germany
georg@mpi-sws.org

Abstract

We consider the rational subset membership problem for Baumslag-Solitar groups. These groups form a prominent class in the area of algorithmic group theory, and they were recently identified as an obstacle for understanding the rational subsets of $\mathrm{GL}(2, \mathbb{Q})$.

We show that rational subset membership for Baumslag-Solitar groups $\mathrm{BS}(1, q)$ with $q \geq 2$ is decidable and PSPACE-complete. To this end, we introduce a word representation of the elements of $\mathrm{BS}(1, q)$: their pointed expansion (PE), an annotated q-ary expansion. Seeing subsets of $\mathrm{BS}(1, q)$ as word languages, this leads to a natural notion of PE-regular subsets of $\mathrm{BS}(1, q)$: these are the subsets of $\mathrm{BS}(1, q)$ whose sets of PE are regular languages. Our proof shows that every rational subset of $\mathrm{BS}(1, q)$ is PE-regular.

Since the class of PE-regular subsets of $\mathrm{BS}(1, q)$ is well-equipped with closure properties, we obtain further applications of these results. Our results imply that (i) emptiness of Boolean combinations of rational subsets is decidable, (ii) membership to each fixed rational subset of $\mathrm{BS}(1, q)$ is decidable in logarithmic space, and (iii) it is decidable whether a given rational subset is recognizable. In particular, it is decidable whether a given finitely generated subgroup of $\mathrm{BS}(1, q)$ has finite index.

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness; Theory of computation → Formal languages and automata theory; Theory of computation → Models of computation

Keywords and phrases Rational subsets, Baumslag-Solitar groups, decidability, regular languages, pointed expansion

Digital Object Identifier 10.4230/LIPIcs.ICALP.2020.116

Category Track B: Automata, Logic, Semantics, and Theory of Programming

Related Version `https://arxiv.org/abs/2006.11898`

Acknowledgements The research for this work was carried out in part at the Autobóz Research Camp in 2019 in Firbush, Scotland. The authors are very grateful to Piotr Hofman for comments that led to important insights and to Michael Blondin as well as to all other participants for discussions.

1 Introduction

Subsets of groups. Regular languages are an extremely versatile tool in algorithmics on sets of finite words. This is mainly due to two reasons. First, they are robust in terms of representations and closure properties: They can be described by finite automata, by recognizing morphisms, and by monadic second-order logic and they are closed under Boolean and an abundance of other operations. Second, many properties (such as emptiness) are easily decidable using finite automata.

47th International Colloquium on Automata, Languages, and Programming (ICALP 2020).
Editors: Artur Czumaj, Anuj Dawar, and Emanuela Merelli; Article No. 116; pp. 116:1–116:16
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Given this success, there have been several attempts to develop an analogous notion for subsets of (infinite, finitely generated) groups. Adapting the notion of recognizing morphism yields *recognizable subsets* of a group G. They are closed under Boolean operations, and problems such as membership or emptiness are decidable. However, since they are merely unions of cosets of finite-index normal subgroups, their expressiveness is severely limited.

Another notion is that of *rational subsets*, which transfer (non-deterministic) finite automata to groups. Starting with pioneering work by Benois [6] in 1969, they have matured into an important tool in group theory. Rational subsets are quite expressive: They include finitely generated submonoids and are closed under (finite) union, pointwise product, and Kleene star. Moreover, they have been applied successfully to solving equations in groups [11, 9], as well as in other settings [2, 34].

The high expressiveness of rational subsets comes at the cost of undecidability of decision problems for many groups. The most fundamental one is the *membership problem* for rational subsets: Given a rational subset R of a group G and an element $g \in G$, does g belong to R? Understanding for which groups this problem is decidable received significant attention over the last two decades, see [24] for a survey. Unfortunately, the rational subsets do not quite reach the level of robustness of regular languages. In general, the class of rational subsets of a group is not closed under Boolean operations, and the papers [25, 4] study for which groups the rational subsets form a Boolean algebra.

Baumslag-Solitar groups. A prominent class of groups is that of *Baumslag-Solitar groups* $\mathrm{BS}(p, q)$. For each $p, q \in \mathbb{N}$, the group is defined as $\mathrm{BS}(p, q) = \langle a, t \mid ta^pt^{-1} = a^q\rangle$. They were introduced in 1962 by Baumslag and Solitar to provide an example of a two-generator one-relator group that is non-Hopfian. They recently came into focus from the algorithmic perspective in a paper by Kharlampovich, López, and Miasnikov [22], which shows that solvability of equations is decidable in $\mathrm{BS}(1, q)$. They have also been studied from several other perspectives, such as the decidability and complexity of the word problem [28, 14, 35], the conjugacy problem [14, 35], tiling problems [1], and computing normal forms [13, 18, 17].

More specifically to our setting, the Baumslag-Solitar groups have recently been identified by Diekert, Potapov, and Semukhin [15] as a stumbling block in solving rational subset membership in the group $\mathrm{GL}(2, \mathbb{Q})$, that is, the group of invertible 2×2 matrices over $\mathbb{Q}$. They show that any subgroup of $\mathrm{GL}(2, \mathbb{Q})$ containing $\mathrm{GL}(2, \mathbb{Z})$ is either of the form $\mathrm{GL}(2, \mathbb{Z}) \times \mathbb{Z}^k$ for $k \geq 1$ or contains $\mathrm{BS}(1, q)$ as a subgroup for some $q \geq 2$. Rational subset membership for $\mathrm{GL}(2, \mathbb{Z}) \times \mathbb{Z}^k$ is today a matter of standard arguments [24], because $\mathrm{GL}(2, \mathbb{Z})$ is virtually free. Therefore, making significant progress towards decidability in larger subgroups requires understanding rational subsets of $\mathrm{BS}(1, q)$.

One can represent the elements of $\mathrm{BS}(1, q)$ as pairs (r, m), where r is a number in $\mathbb{Z}[\frac{1}{q}]$, say $r = \pm\sum_{i=-n}^{n} a_iq^i$ for $a_{-n}, a_{-n+1}, \ldots, a_n \in \{0, \ldots, q-1\}$,[1] and $m \in \mathbb{Z}$. Here, one can think of m as a *cursor* pointing to a position in the q-ary expansion $a_nq^n + \cdots + a_{-n}q^{-n}$. Then the action of the generators of $\mathrm{BS}(1, q)$ is as follows. Multiplication by t or t^{-1} moves the cursor to the left or the right, respectively. Multiplication by a adds q^m; likewise, multiplication by a^{-1} subtracts q^m. Thus, from an automata-theoretic perspective, one can view the rational subset membership problem as the reachability problem for an extended version of one-counter automata. Instead of storing a natural number, such an automaton stores a number $r \in \mathbb{Z}[\frac{1}{q}]$. Moreover, instead of instructions "increment by 1" and "decrement

[1] $\mathbb{Z}[\frac{1}{q}]$ denotes (the additive group of) the smallest subring of $(\mathbb{Q}, +, \cdot)$ containing $\mathbb{Z}$ and $1/q$; as a set, it consists of all rational numbers of the form $n \cdot q^j$, $n, j \in \mathbb{Z}$.

by 1", it has an additional $\mathbb{Z}$-counter m that determines the value to be added in the next update. Then, performing "increment" on r will add q^m and "decrement" on r will subtract q^m. The $\mathbb{Z}$-counter m supports the classical "increment" and "decrement" instructions.

Contribution. Our *first main contribution* is to show is that for each group $\mathrm{BS}(1,q)$, the rational subset membership problem is decidable and PSPACE-complete. To this end, we show that each rational subset can be represented by a regular language of finite words that encode elements of $\mathrm{BS}(1,q)$ in the natural way: For (r,m) as above, we encode each digit a_i by a letter; and we decorate the digits at position 0 and at position m. We call this encoding the *pointed expansion* (PE) of (r,m). This leads to a natural notion of subsets of $\mathrm{BS}(1,q)$, which we call *PE-regular*. We regard the introduction of this notion as the *second main contribution* of this work.

The class of PE-regular subsets of $\mathrm{BS}(1,q)$ has several properties that make them a promising tool for decision procedures for $\mathrm{BS}(1,q)$: First, our proof shows that it effectively includes the large class of rational subsets, in particular any finitely generated submonoid. Second, they form an effective Boolean algebra. Third, due to them being regular languages of words, they inherit many algorithmic tools from the setting of free monoids. We apply these properties to obtain *three applications of our main results.*

1. Membership in each fixed rational subset can be decided in logarithmic space.
2. We show that it is decidable whether a given PE-regular subset (and thus a given rational subset) is recognizable. Recognizability of rational subsets is rarely known to be decidable for groups: The only examples known to the authors are free groups, for which decidability was shown by Sénizergues [31] (and simplified by Silva [33]) and free abelian groups (this follows from [19, Theorem 3.1]). Since (i) finitely generated subgroups are rational subsets and (ii) a subgroup of any group G is recognizable if and only if it has finite index in G, our result implies that it is decidable whether a given finitely generated subgroup of $\mathrm{BS}(1,q)$ has finite index. Studying decidability of this finite index problem in groups was recently proposed by Kapovich [12, Section 4.3].
3. Our results imply that emptiness of Boolean combinations (hence inclusion, equality, etc.) of rational subsets is decidable. (We also show that the rational subsets of $\mathrm{BS}(1,q)$ are not closed under intersection.) This is a strong decidability property that already fails for groups as simple as $F_2 \times \mathbb{Z}$ (this follows from [20, Theorem 6.3]), where F_2 is the free group over two generators, and hence for $\mathrm{GL}(2,\mathbb{Z}) \times \mathbb{Z}^k$, $k \geq 1$.

Finally, we remark that since $\mathrm{BS}(1,q)$ is isomorphic to the group of all matrices $\left(\begin{smallmatrix} q^m & r \\ 0 & 1 \end{smallmatrix}\right)$ for $m \in \mathbb{Z}$ and $r \in \mathbb{Z}[\frac{1}{q}]$, our results can be interpreted as solving the rational subset membership problem for this subgroup of $\mathrm{GL}(2,\mathbb{Q})$.

Related work. It is well-known that membership in a given finitely generated subgroup, called the *generalized word problem* of $\mathrm{BS}(1,q)$, is decidable. This is due to a general result of Romanovskiĭ, who showed in [29] and [30] that solvable groups of derived length two have a decidable generalized word problem (it is an easy exercise to show that $\mathrm{BS}(1,q)$ is solvable of derived length two for each $q \in \mathbb{N}$).

Another restricted version of rational subset membership is the *knapsack problem*, which was introduced by Myasnikov, Nikolaev, and Ushakov [27]. Here, one is given group elements $g_1, \ldots, g_k, g$ and is asked whether there exist $x_1, \ldots, x_k \in \mathbb{N}$ with $g_1^{x_1} \cdots g_k^{x_k} = g$. A recent paper on the knapsack problem in Baumslag-Solitar groups by Dudkin and Treyer [16] left open whether the knapsack problem is decidable in $\mathrm{BS}(1,q)$ for $q \geq 2$. This was settled very recently in [26], where one expresses solvability of $g_1^{x_1} \cdots g_k^{x_k} = g$ in a variant of Büchi arithmetic. A slight extension of that proof yields a regular language as above for the set $S = \{g_1^{x_1} \cdots g_k^{x_k} \mid x_1, \ldots, x_k \in \mathbb{N}\}$. Note that each element g_i moves the cursor either to the left (i.e. increases m), to the right (i.e. decreases m), or not at all. Thus, in a product $g_1^{x_1} \cdots g_k^{x_k}$, the cursor direction is reversed at most $k-1$ times. The challenge of our translation from rational subsets to PE-regular subsets is to capture products where the cursor changes direction an unbounded number of times.

Finally, closely related to rational subsets, there is another approach to group-theoretic problems via automata: One can represent finitely generated subgroups of free groups using *Stallings graphs*. Due to the special setting of free groups, they behave in many ways similar to automata over words and are thus useful for decision procedures [21]. Stallings graphs have recently been extended to semidirect products of free groups and free abelian groups by Delgado [10]. However, this does not include products $\mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$ and is restricted to subgroups.

2 Basic notions

Automata, rational subsets, and regular languages. Since we work with automata over finite words and over groups, we define automata over a general monoid M. A subset $S \subseteq M$ is *recognizable* if there is a finite monoid F and a morphism $\varphi\colon M \to F$ such that $S = \varphi^{-1}(\varphi(S))$. If M is a group, one can equivalently require F to be a finite group.

For a subset $S \subseteq M$, we write $\langle S \rangle$ or S^* for the submonoid *generated by* S, i.e. the set of elements that can be written as a (possibly empty) product of elements of S. In particular, the neutral element $1 \in M$ always belongs to $\langle S \rangle = S^*$. A *generating set* is a subset $\Sigma \subseteq M$ such that $M = \langle \Sigma \rangle$. We say that M is *finitely generated (f.g.)* if it has a finite generating set. Suppose M is finitely generated and fix a finite generating set Σ. An *automaton over* M is a tuple $\mathcal{A} = (Q, \Sigma, E, q_0, q_f)$, where Q is a finite set of *states*, $E \subseteq Q \times \Sigma \times Q$ is a finite set of *edges*, $q_0 \in Q$ is its *initial state*, and $q_f \in Q$ is its *final state*. A *run (in $\mathcal{A}$)* is a sequence $\rho = (p_0, a_1, p_1) \cdots (p_{m-1}, a_m, p_m)$, where $(p_{i-1}, a_i, p_i) \in E$ for $i \in [1, m]$. It is *accepting* if $p_0 = q_0$ and $p_m = q_f$. By $[\rho]$, we denote the *production* of ρ, that is, the element $a_1 \cdots a_m \in M$. Two runs are *equivalent* if they start in the same state, end in the same state, and have the same production. For a set of runs P, we denote $[P] = \{[\rho] \mid \rho \in P\}$.

The subset *accepted by* $\mathcal{A}$ is $\mathsf{L}(\mathcal{A}) = \{[\rho] \mid \rho \text{ is an accepting run in } \mathcal{A}\}$. A subset $R \subseteq M$ is called *rational* if it is accepted by some automaton over M. It is a standard fact that the family of rational subsets of M does not depend on the chosen generating set Σ. Rational subsets of a free monoid Γ^* for some alphabet Γ are also called *regular languages*. If $M = \Gamma^* \times \Delta^*$ for alphabets Γ, Δ, then rational subsets of M are also called *rational transductions*. If $T \subseteq \Gamma^* \times \Delta^*$ and $L \subseteq \Gamma^*$, then we set $TL = \{v \in \Delta^* \mid \exists u \in L\colon (u, v) \in T\}$. It is well-known that if $L \subseteq \Gamma^*$ is regular and $T \subseteq \Gamma^* \times \Delta^*$ is rational, then TL is regular as well [7].

Baumslag-Solitar groups. The *Baumslag-Solitar groups* are the groups $\mathrm{BS}(p, q)$ for $p, q \in \mathbb{N}$, where $\mathrm{BS}(p, q) = \langle a, t \mid ta^pt^{-1} = a^q \rangle$. They were introduced in 1962 by Baumslag and Solitar [3] to provide an example of a non-Hopfian group with two generators and one defining relation. In this paper, we focus on the case $p = 1$. In this case, there is a well-known isomorphism $\mathrm{BS}(1, q) \cong \mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$ and we will identify the two groups. Here, $\mathbb{Z}[\frac{1}{q}]$ is the additive group of number nq^i with $n, i \in \mathbb{Z}$, and $\rtimes$ denotes semidirect product. Building this semidirect product requires us to specify an automorphism φ_m of $\mathbb{Z}[\frac{1}{q}]$ for each $m \in \mathbb{Z}$, which is given by $\varphi_m(nq^i) = q^m \cdot nq^i$.

For readers not familiar with semidirect products, we give an alternative self-contained definition of $\mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$. The elements of this group are pairs (r, m), where $r \in \mathbb{Z}[\frac{1}{q}]$ and $m \in \mathbb{Z}$. The multiplication is defined as

$$(r, m)(r', m') = (r + q^m \cdot r', m + m').$$

We think of an element (r, m) as representing a number r in $\mathbb{Z}[\frac{1}{q}]$ together with a cursor m to a position in the q-ary expansion of r. Multiplying an element (r, m) by the pair $(1, 0)$ from the right means adding 1 at the position in r given by m, hence adding q^m to r and leaving the cursor unchanged: we have $(r, m)(1, 0) = (r + q^m, m)$. Multiplying by $(0, 1)$ moves the cursor one position to the left: $(r, m)(0, 1) = (r, m + 1)$. It is easy to see that $\mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$ is generated by the set $\{(1, 0), (-1, 0), (0, 1), (0, -1)\}$. The isomorphism $\mathrm{BS}(1, q) \xrightarrow{\sim} \mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$ mentioned above maps a to $(1, 0)$ and t to $(0, 1)$. Since we identify $\mathrm{BS}(1, q)$ and $\mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$, we will have $a = (1, 0)$ and $t = (0, 1)$. In particular, a can be thought of as "add"/"increment", and t as "move". We regard elements of the subgroup $\mathbb{Z}[\frac{1}{q}] \times \{0\}$ of $\mathrm{BS}(1, q)$ as elements of $\mathbb{Z}[\frac{1}{q}]$, i.e., integers or rational fractions with denominator q^i, $i \geq 1$.

Rational subset membership. Unless specified otherwise, automata over $\mathrm{BS}(1, q)$ will use the generating set $\Sigma = \{a, a^{-1}, t, t^{-1}\} = \{(1, 0), (-1, 0), (0, 1), (0, -1)\}$. The central decision problem of this work is the *rational subset membership problem for* $\mathrm{BS}(1, q)$:
Given An automaton $\mathcal{A}$ over $\mathrm{BS}(1, q)$ and an element $g \in \mathrm{BS}(1, q)$ as a word over Σ.
Question Does g belong to $\mathsf{L}(\mathcal{A})$?

Automata over BS(1, *q*). In the following definitions, let $\mathcal{A} = (Q, \Sigma, E, q_0, q_f)$ be an automaton over $\mathrm{BS}(1, q)$. For a run ρ of $\mathcal{A}$, recall that $[\rho] \in \mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$ is the *production* of ρ. Moreover, if $[\rho] = (r, m)$ with $r \in \mathbb{Z}[\frac{1}{q}]$ and $m \in \mathbb{Z}$, then we define $\mathsf{pos}(\rho) = m$, and call this the *final position* of ρ. More generally, the *position* at a particular point in ρ is the final position of the corresponding prefix of ρ. By $\mathsf{pmax}(\rho)$, we denote the maximal value of $\mathsf{pos}(\pi)$ where π is a prefix of ρ. Analogously, $\mathsf{pmin}(\rho)$ is the minimal value of $\mathsf{pos}(\pi)$ where π is a prefix of ρ. A run ρ is *returning* if $\mathsf{pos}(\rho) = 0$. It is *returning-left* if in addition $\mathsf{pmin}(\rho) = 0$. Note that for a returning run ρ, we have $[\rho] \in \mathbb{Z}[\frac{1}{q}]$ and if ρ is returning-left, we have $[\rho] \in \mathbb{Z}$. Let $|\rho|$ be the length of the run ρ as a word over E. We will often write ρ_i assuming $\rho = \rho_1\rho_2 \ldots \rho_\ell$ where each $\rho_i \in E$ and $\ell = |\rho|$. A run is a *cycle* if it is returning and starts and ends in the same state. The *thickness* of a run ρ is defined as the greatest number of times a position is seen:

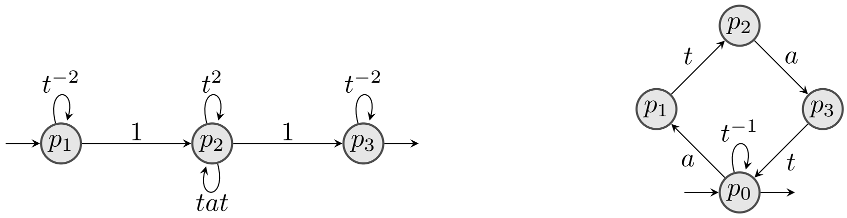
$$\mathsf{thickness}(\rho) = \max_{n \in \mathbb{Z}} |\{i \mid \mathsf{pos}(\rho_1 \cdots \rho_i) = n\}| \ .$$

We call a run *k-thin* if its thickness is at most k.

We let $\mathsf{Runs}(\mathcal{A})$ (resp. $\mathsf{Ret}(\mathcal{A})$, $\mathsf{RetL}(\mathcal{A})$) be the set of all accepting runs (resp. accepting returning runs, accepting returning-left runs) of $\mathcal{A}$. We add k in subscript to restrict the set to k-thin runs; for instance, $\mathsf{Ret}_k(\mathcal{A})$ is the set of k-thin returning runs. Further, we write $\mathsf{Runs}_k^{p \to p'}(\mathcal{A})$ for k-thin runs that start in p and end in p', and use the similar notations $\mathsf{Ret}_k^{p \to p'}(\mathcal{A})$ and $\mathsf{RetL}_k^{p \to p'}(\mathcal{A})$.

Seeing $\{0, \ldots, q-1\}$ as an alphabet, write Φ_q for letters from this alphabet with possibly a $\bullet$ subscript (e.g., $0_\bullet$), a $\triangleleft$ superscript (e.g., $0^\triangleleft$), or both (e.g., $0_\bullet^\triangleleft$). For $v = (r, n) \in \mathrm{BS}(1, q)$, we write $\mathsf{pe}(v)$ for its base-q *pointed expansion* (or just *expansion*) as a word in $\pm\Phi_q^*$, where the subscript $\bullet$ and the superscript $\triangleleft$ appear only once, the former representing the radix point, the latter indicating the value of n. That is, if $r = \sum_{i=-k_2}^{k_1} a_i q^i$, with $k_1, k_2 \geq 0$, $\mathsf{pe}(v)$ is the following word:

$$\pm a_{k_1} \cdots a_1 (a_0)_\bullet a_{-1} \cdots a_{-k_2} \ ,$$



(a) Automaton over $\mathrm{BS}(1,q)$ from Example 3.7. **(b)** Automaton over $\mathrm{BS}(1,2)$ from Example 4.1.

Figure 1 Example automata over $\mathrm{BS}(1,q)$.

where $\triangleleft$ is added to a_n. We tacitly assume a uniqueness condition: the expansion $\mathsf{pe}(v)$ of an element $v \in \mathrm{BS}(1,q)$ is the shortest that abides by the definition. Expansions are read by automata in the left to right direction, i.e., from most to least significant digit.

▶ **Definition 2.1.** *We say that a subset of* $R \subseteq \mathrm{BS}(1,q)$ *is* PE-regular*, where PE stands for pointed expansion, if the word language* $\{\mathsf{pe}(v) \mid v \in R\}$ *is regular.*

We remark that basic properties of regular languages support the transformation of noncanonical expansions of elements $\mathrm{BS}(1,q)$, i.e., those with zeros on the left or right, into canonical ones, $\mathsf{pe}(v)$. Finally, recall that we identify each $r \in \mathbb{Z}[\frac{1}{q}]$ with $(r,0) \in \mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$. Hence, for $r \in \mathbb{Z}[\frac{1}{q}]$, $\mathsf{pe}(r)$ is the q-ary expansion of r (with $\triangleleft$ as an additional decoration at the radix point).

3 Main results

Our first main result is that one can translate rational subsets into PE-regular subsets.

▶ **Theorem 3.1.** *Every rational subset of* $\mathrm{BS}(1,q)$ *is effectively PE-regular.*

This will be shown in Section 4. Since membership is decidable for regular languages and given $g \in \mathrm{BS}(1,q)$ as a word over $\{a, a^{-1}, t, t^{-1}\}$, one can compute $\mathsf{pe}(g)$, Theorem 3.1 implies that rational subset membership is decidable. Our next main result is that the problem is PSPACE-complete.

▶ **Theorem 3.2.** *The rational subset membership problem for* $\mathrm{BS}(1,q)$ *is* PSPACE*-complete.*

This is shown in Section 5. We shall also conclude that membership to each fixed rational subset is decidable in logspace.

▶ **Theorem 3.3.** *For each fixed rational subset of* $\mathrm{BS}(1,q)$*, membership is decidable in logarithmic space.*

The proof can also be found in Section 5. Note that, in particular, membership to each fixed subgroup of $\mathrm{BS}(1,q)$ is decidable in logarithmic space. Another application of Theorem 3.1 is that one can decide whether a given rational subset of $\mathrm{BS}(1,q)$ is recognizable.

▶ **Theorem 3.4.** *Given a PE-regular subset R of* $\mathrm{BS}(1,q)$*, it is decidable whether R is recognizable.*

This is shown in Section 6. Since a subgroup of any group H is recognizable if and only if it has finite index in H (see, e.g. [2, Prop. 3.2]), we obtain:

▶ **Corollary 3.5.** *Given a f.g. subgroup of* $\mathrm{BS}(1,q)$*, it is decidable whether it has finite index.*

We close this section by showing that regular subsets of $\mathrm{BS}(1,q)$ are robust in terms of closure properties.

▶ **Proposition 3.6.** *The PE-regular subsets of* $\mathrm{BS}(1,q)$ *form an effective Boolean algebra. Moreover, for PE-regular subsets* $R, S \subseteq \mathrm{BS}(1,q)$*, the sets* $RS = \{rs \mid r \in R,\ s \in S\}$ *and* $R^{-1} = \{r^{-1} \mid r \in R\}$ *are PE-regular as well.*

The proof is straightforward. Together with Theorem 3.1, this implies that emptiness of Boolean combinations (hence inclusion, equality) is decidable for rational subsets. To further highlight the advantages of PE-regular subsets, we also show that the rational subsets of $\mathrm{BS}(1,q)$ are not closed under intersection.

▶ **Example 3.7** (Intersection of rational subsets)**.** Let R be the set accepted by the automaton in Figure 1a. The automaton first moves an even number of positions to the right (p_1) and then an even number of positions to the left while adding 1 in a subset of the odd positions (p_2). Finally, it goes an even number of positions to the left again. Note that $(r,m) \in R$ if and only if $r = \sum_{i\in A} q^{2i+1}$ for some finite $A \subseteq \mathbb{Z}$ and $m \in 2\mathbb{Z}$. Now consider the rational sets aR and Ra and their intersection $I = aR \cap Ra$. Note that $(r,m) \in aR$ if and only if $r = 1 + \sum_{i\in A} q^{2i+1}$ and $m \in 2\mathbb{Z}$ for some finite $A \subseteq \mathbb{Z}$. Moreover, $(r,m) \in Ra$ if and only if $r = q^m + \sum_{i\in A} q^{2i+1}$ and $m \in 2\mathbb{Z}$ for some finite $A \subseteq \mathbb{Z}$. Therefore, we have $(r,m) \in I$ if and only if $r = 1 + \sum_{i\in A} q^{2i+1}$ and $m = 0$ for some finite $A \subseteq \mathbb{Z}$. Since I only contains elements with cursor 0, but carries non-zero digits in positions that are arbitrarily far to the right, it follows that I is not rational.

However, the PE-regular subsets of $\mathrm{BS}(1,q)$ are not closed under iteration.

▶ **Example 3.8** (Iteration of PE-regular subsets)**.** The subset $A = \{(1 + 2^{-i}, 0) \mid i \geq 1\}$ of $\mathrm{BS}(1,2)$ is PE-regular, because $\mathsf{pe}(A) = 1\overset{\triangleleft}{\bullet}0^*1$ is a regular language. However, the set A^* is not PE-regular: one can show that for each $n \geq 1$, we have $n = \min\{m \in \mathbb{N} \mid (m + 2^{-1} + \cdots + 2^{-n}, 0) \in A^*\}$.[2] Therefore, for each $n \geq 1$, a word in $\mathsf{pe}(A^*)$ with 1^n to the right of the radix point can have an integer part of n and cannot have a smaller integer part. This implies that $\mathsf{pe}(A^*)$ is not regular.

4 Every rational subset of BS(1, *q*) is effectively PE-regular

In this section, we prove Theorem 3.1. We first illustrate our approach on an example.

▶ **Example 4.1.** Consider the automaton over $\mathrm{BS}(1,2)$ in Figure 1b. In its only initial and final state p_0, it has a choice of two operations: (i) move the cursor one position to the right (i.e. multiplication by t^{-1}) or (ii) perform the increment on two neighbouring cells and stop one position left of them (i.e. multiplication by $atat$). The automaton can perform these operations arbitrarily many times in any order.

We shall prove that the automaton accepts

$$R = \{(3n \cdot 2^{m-2k}, m) \mid n \in \mathbb{N},\ k \in \mathbb{N},\ m \in \mathbb{Z},\ 0 \geq m - 2k,\ 3n \cdot 2^{m-2k} \geq f(m,k)\}\ ,$$

where

$$f(m,k) = \sum_{i=1}^{k} 3 \cdot 2^{m-2i} = \sum_{j=m-2k}^{m-1} 2^j = 2^m - 2^{m-2k}\ .$$

[2] We denote $\mathbb{N} = \{0, 1, 2, \ldots\}$.

The language $\mathsf{pe}(R)$ is regular. Indeed, note that the number $f(m,k)$ has a particularly simple binary representation. A pointed expansion of (r,m) belongs to $\mathsf{pe}(R)$ if there is a position $m-2k \leq 0$ such that reading the digits left of position $m-2k$ yields a number (namely $3n$) that (a) is divisible by 3 and (b) lies above a bound with a simple binary expansion.

Let us now prove that the automaton accepts R. Let ρ be an accepting run producing (r,m). Choose $k \in \mathbb{N}$ so that $\mathsf{pmin}(\rho) = m-2k$ or $\mathsf{pmin}(\rho) = m-2k+1$ (depending on whether $m-\mathsf{pmin}(\rho)$ is even or odd). Then $0 \geq \mathsf{pmin}(\rho) \geq m-2k$. Each time operation (ii) is performed from position $\ell \in \mathbb{Z}$, the update is $(r,m) \to (r+3\cdot 2^{\ell}, m+2)$.

Now, once ρ visits position $\mathsf{pmin}(\rho)$, in order to eventually reach a position $\ell > \mathsf{pmin}(\rho)$, the operation (ii) must be performed on some position $\geq \ell-2$. In particular, to reach position m, it must be performed at some position $m_1 \geq m-2$. If $m_1 > \mathsf{pmin}(\rho)$, to reach m_1, it must also be performed at some position $m_2 \geq m-4$, etc. Therefore, ρ has to perform (ii) at positions $m_i \geq m-2i$ for each i with $m > m-2i \geq \mathsf{pmin}(\rho)-1$. In other words, it has to do this for each $i = 1,\ldots,k$. Each time ρ performs (ii) at m_i, it adds $3\cdot 2^{m_i}$. Moreover, each extra time ρ performs (ii), it adds a multiple of $3\cdot 2^{m-2k}$, because $\mathsf{pmin}(\rho) \geq m-2k$. Thus, the number produced in total is some $3n\cdot 2^{m-2k}$ where

$$3n\cdot 2^{m-2k} \geq \sum_{i=1}^{k} 3\cdot 2^{m_i} \geq \sum_{i=1}^{k} 3\cdot 2^{m-2i} = f(m,k) \ .$$

Conversely, suppose $n \in \mathbb{N}$ and $k \in \mathbb{N}$, $m \in \mathbb{Z}$, $0 \geq m-2k$, and $3n\cdot 2^{m-2k} \geq f(m,k)$. The automaton first moves to position $m-2k$ using operation (i). Then, it performs operations (ii), (i), and (i) again, ℓ times in a loop (we specify ℓ later). That way, it adds $3\ell\cdot 2^{m-2k}$. Then, it moves to position m by applying operation (ii) exactly k times. Hence, it applies (ii) at positions $m-2i$ for $i=1,\ldots,k$ and each time, it adds $3\cdot 2^{m-2i}$. In total, the effect is

$$3\ell\cdot 2^{m-2k} + \sum_{i=1}^{k} 3\cdot 2^{m-2i} = 3\ell\cdot 2^{m-2k} + f(m,k) \ .$$

Since $3n\cdot 2^{m-2k} \geq f(m,k)$ and $f(m,k)$ is an integer multiple of $3\cdot 2^{m-2k}$, we can choose $\ell \in \mathbb{N}$ so as to produce $3n\cdot 2^{m-2k}$.

Following this example, we first show that any run has the same production as a thin (i.e. bounded thickness) run in which thin returning-left cycles are inserted (p. 9); in the example, such a cycle applies operations (ii), (i), and (i). We then prove that the productions of thin runs form a PE-regular set (p. 10); in the example, the thin run moves to the right to position $\mathsf{pmin}(\rho)$ using operation (i) and then left to $m \geq \mathsf{pmin}(\rho)$ using operations (i) and (ii). Finally, we show that iterating returning-left thin cycles also leads to a PE-regular set (p. 11); in the example, this is how we get all numbers divisible by 3 above a particular bound. We combine these three statements to prove Theorem 3.1.

In combining the thin run with cycles, we will need to ensure that the cycles are anchored on the correct state. To this end, we introduce an annotated version of $\mathsf{pe}([\rho])$ as follows. Let $\mathcal{A}$ be an automaton over $\mathrm{BS}(1,q)$ with state set Q. Let ρ be a run in $\mathcal{A}$ starting and ending in arbitrary states and with $[\rho] = (r,m)$. Letting $\bar{Q} = \{\bar{p} \mid p \in Q\}$ be a copy of Q, we define $\mathsf{sv}(\rho)$, the *state view* of ρ, to be the word over the alphabet $\Phi_q \cup Q \cup \bar{Q} \cup \{\pm\}$ built as follows. First, write: $\mathsf{pe}([\rho]) = \pm a_{k_1}\cdots a_1 a_0 a_{-1}\cdots a_{-k_2}$,where a_0 has subscript $\bullet$. Second, let $P_i \in (Q\cup\bar{Q})^{|Q|}$, for $i \in \{-k_2,\ldots,k_1\}$, be a word that contains all the states of Q once in a fixed ordering of Q, either with a bar or not; the states without a bar are exactly those that visit position i in ρ. That is, p appears in P_i iff there is a prefix of ρ ending in p whose final position is i. The state view of ρ is then:

$$\mathsf{sv}(\rho) = \pm a_{k_1} \cdot P_{k_1} \cdots a_0 \cdot P_0 \cdot a_{-1} \cdot P_{-1} \cdots a_{-k_2} \cdot P_{-k_2} \ .$$

We naturally extend sv to sets of runs.

Any run is equivalent to a thin run augmented with thin returning-left cycles. We now focus on two properties of runs: the states they visit in the automaton and the final position of their prefixes. To that end, we introduce the following notions. For Q a finite set, a *position path* is a word $\pi \in (Q \times \mathbb{Z})^*$. We extend the analogy with graphs calling elements of $Q \times \mathbb{Z}$ *vertices*, talking of the vertices *visited* by a position path, and using the notion of (position) *subpaths* and *cycles*. The *thickness* of a position path π is defined as:

$$\mathsf{thickness}(\pi) = \max_{n \in \mathbb{Z}} |\{i \mid \pi_i = (q, n) \text{ for some } q\}| \ .$$

▶ **Lemma 4.2.** *Let Q be a finite set and $\pi \in (Q \times \mathbb{Z})^*$ be a position path. For any subset V' of the vertices visited by π, there exists a subpath π' of π such that:*
1. *π' starts and ends with the same vertices as π,*
2. *π' visits all the vertices in V',*
3. *$\mathsf{thickness}(\pi') \leq |Q| \cdot (1 + 2|V'|)$,*
4. *$\pi - \pi'$ consists only of cycles.*

Proof (sketch). We first consider a shortest subpath π' of π from the initial to the final vertices of π – this implies that π' has thickness at most $|Q|$. We then treat each missing vertex from V' in turn, and add to π' a subpath from π that is a cycle and includes that vertex. Each of these iterations can augment the thickness of π' by at most $2|Q|$. ◀

▶ **Corollary 4.3.** *Let $\mathcal{A}$ be an automaton over* $\mathrm{BS}(1, q)$ *with state set Q, and let $k = |Q| + 2|Q|^2$. Any run of $\mathcal{A}$ is equivalent to a run in $\mathsf{Runs}_k(\mathcal{A})$ on which, for each state p appearing in the run, cycles from $\mathsf{RetL}_k^{p \to p}(\mathcal{A})$ are inserted at an occurrence of p with smallest position.*

Conversely, any run built by taking a run in $\mathsf{Runs}_k(\mathcal{A})$ and inserting cycles from $\mathsf{RetL}_k^{p \to p}(\mathcal{A})$ at an occurrence of p is a run of $\mathcal{A}$.

Proof. The converse is clear, we thus focus on the first direction.

(Step 1: Decomposing a run into a thin run and cycles.) Let $\rho \in \mathsf{Runs}(\mathcal{A})$, and extract from it a position path $\pi = \pi_0 \cdots \pi_{|\rho|}$ as follows. We let, $\pi_0 = (q_0, 0)$ and for all $i \geq 1$:

$$\pi_i = (p, n) \text{ where } \rho_i = (\cdot, \cdot, p) \text{ and } n = \mathsf{pos}(\rho_1 \cdots \rho_i) \ .$$

For each state p visited by ρ, let $n_p = \min\{n \mid \text{there exists } i \text{ such that } \pi_i = (p, n)\}$; in words, n_p is the smallest final position of a prefix of ρ ending in p. Using $V' = \{(p, n_p) \mid \rho \text{ visits } p\}$, Lemma 4.2 provides a position path π' of thickness $\leq k = |Q| + 2|Q|^2$ visiting all of V'.

From π', we can obtain the corresponding subpath ρ' of ρ that has the same starting and ending state and positions as ρ, and such that ρ is made of ρ' onto which cycles are added. The thickness of ρ' is bounded by k, but the cycles can be of any thickness.

(Step 2: Thinning the cycles.) Consider a cycle β that gets added to ρ' to form ρ, say at position i (after initial i moves, $\rho'_1 \cdots \rho'_i$), and assume that $\mathsf{thickness}(\beta) > k$. Since a position is repeated more than $k > |Q|$ times, there is a cycle β' *within* β with $\mathsf{thickness}(\beta') \leq k$;

write then $\beta = \alpha \cdot \beta' \cdot \alpha'$. Let p be the state in β' that has the smallest position, that is, p is the ending state of the prefix γ of β' with final position $\mathsf{pmin}(\beta')$; write $\beta' = \gamma \cdot \gamma'$. By definition, we have $\mathsf{pos}(\rho'_1 \cdots \rho'_i \alpha\gamma) \geq n_p$. Note that $\gamma' \cdot \gamma$ is in $\mathsf{RetL}_k^{p \to p}(\mathcal{A})$. We now remove β' from β and then insert $\gamma' \cdot \gamma$ at the position j in ρ' that is such that $\rho'_1 \cdots \rho'_j$ ends in p with final position n_p. For the contribution of $\gamma' \cdot \gamma$ to be the same as that of β' in the original path, we insert it q^d times, where $d = \mathsf{pos}(\rho'_1 \cdots \rho'_i \alpha\gamma) - n_p$.

This shows that if any cycle added to ρ' is of thickness $> k$, then a subcycle of it can be moved to another position of ρ' as a returning-left cycle. Iterating this process, all the cycles added to ρ' will thus be of thickness $\leq k$. Moreover, if an added cycle β is not returning-left after these operations, or if it does not sit at an occurrence of its initial state with *smallest* position, this means that we can decompose it just as above as $\gamma \cdot \gamma'$, with γ reaching $\mathsf{pmin}(\beta)$, and move $\gamma' \cdot \gamma$, a returning-left cycle, to an appropriate position in ρ' as before. ◀

Intermezzo: reflecting on Corollary 4.3. Before we continue with the proof, we want to illustrate how crucial the previous corollary is. Lemma 4.2 tells us that we can obtain every run from a thin run by then adding cycles. This already simplifies the structure of $\mathsf{Runs}(\mathcal{A})$: indeed, inserting cycles at a certain position in a run $\rho \in \mathsf{Runs}(\mathcal{A})$ corresponds (in algebraic terms) to adding to $[\rho]$ a subset of $\mathbb{Z}[\frac{1}{q}]$ closed under addition, i.e., a submonoid. (Closure under addition follows from the observation that any two returning cycles from each $\mathsf{Ret}_k^{p \to p}(\mathcal{A})$ can be concatenated.)

Sometimes one can conclude that every submonoid of a monoid has a simple structure. For example, every submonoid M of $\mathbb{Z}$ is semilinear and hence a PE-regular subset of $\mathbb{Z}[\frac{1}{q}]$. Unfortunately, the situation in $\mathbb{Z}[\frac{1}{q}]$ is not as simple as in $\mathbb{Z}$: One can show that $\mathbb{Z}[\frac{1}{q}]$ has uncountably many submonoids. Thus, $\mathbb{Z}[\frac{1}{q}]$ has submonoids with undecidable membership problem; moreover, there is no hope for a finite description for every submonoid as in $\mathbb{Z}$. Thus, we need to look at our specific submonoids. A simple observation similar to Lemma 4.2 allows us to obtain every run from a thin part by adding *thin* cycles. Hence, the submonoids that we add are of the form $[\mathsf{Ret}_k^{p \to p}(\mathcal{A})]^*$. It is not hard to show (see Lemma 4.4) that $[\mathsf{Ret}_k^{p \to p}(\mathcal{A})]$ is always a PE-regular set. Thus, one may hope to prove that the regularity of $[\mathsf{Ret}_k^{p \to p}(\mathcal{A})]$ implies regularity of $[\mathsf{Ret}_k^{p \to p}(\mathcal{A})]^*$. (This was an approach to rational subset membership proposed by the third author of this work in [12, Section 4.7].) However, Example 3.8 tells us that even for PE-regular $R \subseteq \mathrm{BS}(1,q)$, the set R^* may not be PE-regular.

Therefore, Corollary 4.3 is the key insight of our proof. It says that a run can be decomposed into a thin part and thin *returning-left* cycles. Since returning-left cycles produce integers, this will lead us to submonoids of $\mathbb{Z}$.

Sets of thin runs are PE-regular.

▶ **Lemma 4.4.** *Let $\mathcal{A}$ be an automaton over* $\mathrm{BS}(1,q)$*, p, p' be states of $\mathcal{A}$, and $k > 0$. The sets* $\mathsf{sv}(\mathsf{Runs}_k^{p \to p'}(\mathcal{A})), \mathsf{sv}(\mathsf{Ret}_k^{p \to p'}(\mathcal{A}))$*, and* $\mathsf{sv}(\mathsf{RetL}_k^{p \to p'}(\mathcal{A}))$ *are effectively regular.*

Proof (sketch). We see $\mathcal{A}$ as a two-way automaton, and apply a construction similar to the classical proof that two-way automata are no more expressive than one-way automata [32]. This transforms $\mathcal{A}$ into a one-way automaton over the alphabet $\{-1, 0, 1\}^k$, where each component tracks a 1-thin partial run. It is a classical exercise to show that automata can compute the addition of numbers in a given base; this can be extended to *signed-digit* expansions, in which negative digits can be used [8, Section 2.2.2.2]. We thus rely on this to compute the sum, componentwise, of these partial runs. Adding state information to that construction is straightforward, so that we obtain automata for state views. ◀

Iterations of returning-left thin cycles are PE-regular. It is well-known that for every set $S \subseteq \mathbb{N}$ the generated monoid $S^* = \{s_1 + \cdots + s_m \mid s_1, \ldots, s_m \in S, m \geq 0\}$ is eventually identical with $\gcd(S) \cdot \mathbb{N}$. In other words, the set $(\gcd(S) \cdot \mathbb{N}) \setminus S^*$ is finite and we may define $F(S) = \max((\gcd(S) \cdot \mathbb{N}) \setminus S^*)$. The number $F(S)$ is called the *Frobenius number* of S. With this, we have $S^* = \{n \in S^* \mid n \leq F(S)\} \cup \{n \in \gcd(S) \cdot \mathbb{N} \mid n > F(S)\}$. If $S \subseteq -\mathbb{N}$, then we set $F(S) := F(-S)$. Now consider an arbitrary set $S \subseteq \mathbb{Z}$. If S contains both a positive and a negative number, then $S^* = \gcd(S) \cdot \mathbb{Z}$ and we set $F(S) := 0$. We shall use the following well-known fact [36].

▶ **Lemma 4.5.** *If $S = \{n_1, \ldots, n_k\}$ with $0 < n_1 < \cdots < n_k$, then $F(S) \leq n_k^2$.*

▶ **Lemma 4.6.** *For every automaton $\mathcal{A}$ over $\mathrm{BS}(1,q)$, the language $\mathsf{pe}([\mathsf{RetL}_k^{p\to p}(\mathcal{A})]^*)$ is effectively regular.*

Proof. Recall that we identify each $r \in \mathbb{Z}[\frac{1}{q}]$ with $(r, 0) \in \mathbb{Z}[\frac{1}{q}]$. In particular, for $n \in \mathbb{Z}$, $\mathsf{pe}(n)$ is the same as $\mathsf{pe}((n,0))$.

Denote $S = [\mathsf{RetL}_k^{p\to p}(\mathcal{A})]$. We first consider the case $S \subseteq \mathbb{N}$ and $S \neq \emptyset$. Suppose we can compute $\gcd(S)$ and a bound $B \in \mathbb{N}$ with $B \geq F(S)$. Then we have

$$S^* = \underbrace{\{n \in S^* \mid n \leq B\}}_{=:X} \cup \underbrace{\{n \in \gcd(S) \cdot \mathbb{N} \mid n > B\}}_{=:Y} \tag{1}$$

and it suffices to show that $\mathsf{pe}(X)$ and $\mathsf{pe}(Y)$ are effectively regular. Note that X is finite and can be computed by finding all $n \leq B$ with $n \in S$ (recall that membership in S is decidable because $\mathsf{sv}(\mathsf{RetL}_k^{p\to p}(\mathcal{A}))$ is effectively regular by Lemma 4.4) and building sums. Moreover, $\mathsf{pe}(Y)$ is regular because the set $L_0 = \mathsf{pe}(\gcd(S) \cdot \mathbb{N})$ is effectively regular and so is $L_1 = \{\mathsf{pe}(n) \mid n \in \mathbb{N},\ n > B\}$, and hence $\mathsf{pe}(Y) = L_0 \cap L_1$.

Thus, it remains to compute $\gcd(S)$ and some $B \geq F(S)$. For the former, find any $r \in S$ and consider its decomposition $r = p_1^{e_1} \cdots p_m^{e_m}$ into prime powers. For each $i \in [1, m]$, we compute $d_i \in [0, e_i]$ and $n_i \in S$ such that (i) $S \subseteq p_i^{d_i} \cdot \mathbb{N}$, and (ii) $n_i \in S \setminus p_i^{d_i+1} \cdot \mathbb{N}$. Since for $d \in \mathbb{N}$, we can construct an automaton for $\mathsf{pe}(S \cap d \cdot \mathbb{N})$, these d_i and n_i can be computed. Observe that $\gcd(S) = p_1^{d_1} \cdots p_m^{d_m}$. Let $T = \{r, n_1, \ldots, n_k\}$. Observe that $\gcd(T) = \gcd(S)$, and hence T^* and S^* are ultimately identical. Since $T \subseteq S$, this means $F(S) \leq F(T)$. By Lemma 4.5, we have $F(T) \leq (\max\{r, n_1, \ldots, n_k\})^2$, which yields our bound B.

The case $S \subseteq -\mathbb{N}$ is analogous to $S \subseteq \mathbb{N}$. If S contains a positive and a negative number, then $S^* = \gcd(S) \cdot \mathbb{Z}$, so it suffices to just compute $\gcd(S)$. This is done as above. Finally, deciding between these three cases is easy. This completes the proof. ◀

Wrapping up: Proof of Theorem 3.1. Let $\mathcal{A}$ be an automaton over $\mathrm{BS}(1,q)$ with state set Q. Corollary 4.3 indicates that the set of productions of accepting runs is the same as the set of productions of k-thin runs in which thin cycles are introduced.

By Lemma 4.4, $\mathsf{sv}(\mathsf{Runs}_k(\mathcal{A}))$ is a regular language L. For any state p of $\mathcal{A}$, let $L_p = \mathsf{pe}\big([\mathsf{RetL}_k^{p\to p}(\mathcal{A})]^*\big)$, a regular language by Lemma 4.6. For padding purposes, let $s \in Q$ be some state, and let h be the morphism from $(\Phi_q \cup \{\pm\})^*$ to $(\Phi_q \cup Q \cup \{\pm\})^*$ defined, for any $a \in \Phi_q$, by $h(a) = as^{|Q|}$, and $h(+) = +$, $h(-) = -$. Define now L'_p to be the image by h of the version of L_p where arbitrary 0's are added after the sign, and at the end of the number (these 0's do not change the value represented).

Consider now the language R over the alphabet $(\Phi_q \cup Q \cup \bar{Q} \cup \{\pm\})^{|Q|+1}$ whose projection on the first component is the language L, and the other components correspond to the languages L'_p, for each $p \in Q$. The first component indicates in particular the states of $\mathcal{A}$

that visited that location; to synchronize the different components of R, we ensure that the letter annotated with $\bullet$ in L'_p is aligned with a letter from L that is followed by p – that is, the starting position of L'_p is at a position in L that is seen while being in the state p.

Finally, an automaton can do the componentwise addition in base q, collapsing the $|Q|+1$ components into a single one. The radix point is given by the digit with $\bullet$ of L, i.e., in the first component; and similarly for $\triangleleft$. The resulting language, thanks to Corollary 4.3, is the language of the pointed expansions of all runs in $\mathsf{Runs}(\mathcal{A})$. ◀

5 Complexity

Computing pointed expansions. In this section, we prove Theorems 3.2 and 3.3. For the upper bounds in Theorems 3.2 and 3.3, we shall rely on the fact that, given an element $g \in \mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$ as a word over $\Sigma = \{a, a^{-1}, t, t^{-1}\}$, one can compute the pointed expansion $\mathsf{pe}(g)$ in logarithmic space. This is a direct consequence of a result of Elder, Elston, and Ostheimer [18, Proposition 32]. They show that given a word w over Σ, one can compute in logarithmic space an equivalent word of one of the forms (i) t^i, (ii) $(a^{\eta_0})^{t^{\alpha_0}}(a^{\eta_1})^{t^{\alpha_1}} \cdots (a^{\eta_k})^{t^{\alpha_k}} t^i$ or (iii) $(a^{-\eta_0})^{t^{\alpha_0}}(a^{-\eta_1})^{t^{\alpha_1}} \cdots (a^{-\eta_k})^{t^{\alpha_k}} t^i$, where $i \in \mathbb{Z}$, $k \in \mathbb{N}$, $0 < \eta_j < q$ for $j \in [0,k]$, and $\alpha_0 > \cdots > \alpha_k$. Here, x^y stands for $y^{-1}xy$ in the group. Since these normal forms denote the elements (i) $(0,i)$, (ii) $(\sum_{j=0}^{k} \eta_j q^{-\alpha_j}, i)$ and (iii) $(-\sum_{j=0}^{k} \eta_j q^{-\alpha_j}, i)$, respectively, it is easy to turn these normal forms into $\mathsf{pe}(w)$ using logarithmic space.

This allows us to prove Theorem 3.3: For every rational subset $R \subseteq \mathrm{BS}(1,q)$, the language $\mathsf{pe}(R)$ is a regular language. In particular, there exists a deterministic automaton $\mathcal{B}$ for $\mathsf{pe}(R)$. Therefore, given $g \in \mathrm{BS}(1,q)$ as a word over $\{a, a^{-1}, t, t^{-1}\}$, we compute $\mathsf{pe}(g)$ in logspace and then check membership of $\mathsf{pe}(g)$ in $\mathsf{L}(\mathcal{B})$, which is decidable in logarithmic space.

PSPACE-completeness. The PSPACE lower bound in Theorem 3.2 is a reduction from the intersection nonemptiness of finite-state automata, a well-known PSPACE-complete problem [23]. For the PSPACE upper bound, we strengthen Theorem 3.1 by constructing a polynomial-size representation of an exponential size automaton for the resulting regular language. A *succinct finite automaton* is a tuple $\mathcal{S} = (n, \Gamma, (\varphi_x)_{x \in \Gamma \cup \{\varepsilon\}}, p_0, p_f\})$, where $n \in \mathbb{N}$ is its *bit length*, Γ is its *input alphabet*, $\varphi_x(\mathsf{v}_1, \ldots, \mathsf{v}_n, \mathsf{v}'_1, \ldots, \mathsf{v}'_n)$ is a formula from propositional logic with free variables $\mathsf{v}_1, \ldots, \mathsf{v}_n, \mathsf{v}'_1, \ldots, \mathsf{v}'_n$ for each $x \in \Gamma \cup \{\varepsilon\}$, $p_0 \in \{0,1\}^n$ is its *initial state*, and $p_f \in \{0,1\}^n$ is its *final state*. The *size* of $\mathcal{S}$ is defined as $|\mathcal{S}| = n + \sum_{x \in \Gamma \cup \{\varepsilon\}} |\varphi_x|$, where $|\varphi|$ denotes the length of the formula φ.

Moreover, $\mathcal{S}$ *represents* the automaton $\mathcal{A}(\mathcal{S})$, which is defined as follows. It has the state set $\{0,1\}^n$, initial state p_0, and final state p_f. For states $p = (b_1, \ldots, b_n), p' = (b'_1, \ldots, b'_n) \in \{0,1\}^n$ and $x \in \Gamma \cup \{\varepsilon\}$, there is an edge (p, x, q) in $\mathcal{A}(\mathcal{S})$ if and only if $\varphi_x(b_1, \ldots, b_n, b'_1, \ldots, b'_n)$ holds. We define the *language accepted by* $\mathcal{S}$ as $\mathsf{L}(\mathcal{S}) = \mathsf{L}(\mathcal{A}(\mathcal{S}))$.

We allow ε-edges in succinct automata, and with Boolean formulas, one can encode steps in a Turing machine. Thus, a succinct automaton of polynomial size can simulate a polynomial space Turing machine with a one-way read-only input tape. Our descriptions of succinct automata will therefore be in the style of polynomial space algorithms. We show:

▶ **Theorem 5.1.** *Given a rational subset $R \subseteq \mathrm{BS}(1,q)$, one can construct in polynomial space a polynomial-size succinct automaton accepting* $\mathsf{pe}(R)$.

This allows us to decide rational subset membership in PSPACE: Given an automaton $\mathcal{A}$ over $\mathrm{BS}(1,q)$ and an element g as a word over $\{a, a^{-1}, t, t^{-1}\}$, we construct a succinct automaton $\mathcal{B}$ for $\mathsf{pe}(\mathsf{L}(\mathcal{A}))$ and the pointed expansion $\mathsf{pe}(g)$ in logarithmic space. Since membership in succinct automata is well-known to be in PSPACE, we can check whether $\mathsf{pe}(g) \in \mathsf{L}(\mathcal{B})$.

Constructing succinct automata. It remains to prove Theorem 5.1. The construction of a succinct automaton for $\mathsf{pe}(R)$ proceeds with the same steps as in Section 4. For most of these steps, our constructions already yield small succinct automata (e.g., one for $\mathsf{pe}([\mathsf{RetL}_k^{p\to p'}(\mathcal{A})])$ in Lemma 4.4). The exception is Lemma 4.6 – in which case the key ingredient is as follows.

▶ **Proposition 5.2.** *Given an automaton $\mathcal{A}$ over $\mathrm{BS}(1,q)$, a state p of $\mathcal{A}$, and $k \in \mathbb{N}$ in unary, one can compute in polynomial space the number $\gcd([\mathsf{RetL}_k^{p\to p}(\mathcal{A})])$ and a bound $B \geq F([\mathsf{RetL}_k^{p\to p}(\mathcal{A})])$. Both are at most exponential in k and the size of $\mathcal{A}$.*

Our bound on F extends the bound for automatic sets in $\mathbb{N}$ [5, Lemma 4.5] to thin two-way computations. Before proving Proposition 5.2, let us show how it implies Theorem 5.1.

Proof of Theorem 5.1. The constructions in Lemma 4.4 and Theorem 3.1, immediately yield a polynomial-size succinct automaton for $\mathsf{pe}(R)$ once a succinct automaton for each $\mathsf{pe}([\mathsf{RetL}_k^{p\to p}(\mathcal{A})]^*)$ is found. For the latter, we proceed as in Lemma 4.6. Let $S = [\mathsf{RetL}_k^{p\to p}(\mathcal{A})]$ and compute $\gcd(S)$ and a bound $B \geq F(S)$ using Proposition 5.2. Then, by Equation (1) on page 11, it suffices to construct a succinct automaton for $\mathsf{pe}(X)$ and one for $\mathsf{pe}(Y)$. For $\mathsf{pe}(X)$, we use the fact that we can construct a succinct automaton $\mathcal{B}$ for $\mathsf{pe}(S)$. Our automaton for $\mathsf{pe}(X)$ proceeds as follows. With ε-transitions, it runs $\mathcal{B}$ to successively guess numbers $\leq B$ from S and stores each of them temporarily in its state. Such a number requires $O(\log(B))$ bits. In another $O(\log(B))$ bits, it stores the sum of the numbers guessed so far. This continues as long as the sum is at most B. Then, our automaton reads the resulting sum from the input. This automaton clearly accepts $\mathsf{pe}(X)$.

For $\mathsf{pe}(Y)$, we have to construct a succinct automaton that accepts any number $> B$ that is divisible by $\gcd(S)$. Since $\gcd(S)$ is available as a number with polynomially many digits, we can construct a succinct automaton accepting $\mathsf{pe}(\gcd(S)\cdot\mathbb{N})$: It keeps the remainder modulo $\gcd(S)$ of the currently read prefix. This requires $O(\log(\gcd(S))$ many bits. Since B also has polynomially many digits, we can construct a succinct automaton for $\{n \in \mathbb{N} \mid n > B\}$. An automaton for the intersection then accepts $\mathsf{pe}(Y)$. ◀

It is easy to see that the number produced by a returning-left run is at most exponential in the length of the run. The exact bound will not be important.

▶ **Lemma 5.3.** *If ρ is a run in $\mathsf{RetL}_k(\mathcal{A})$ of length ℓ, then $|[\rho]| \leq q^{2\ell}$.*

The main ingredient for Proposition 5.2 will be Lemma 5.4. We write $\rho \ll \rho'$ if $|\rho| < |\rho'|$. Moreover, for $d \in \mathbb{Z}$, we write $\rho \ll_d \rho'$ if $\rho \ll \rho'$ and for some $\ell \in \mathbb{Z}$, we have $[\rho'] = \ell \cdot [\rho] + d$.

▶ **Lemma 5.4.** *There is a polynomial f such that the following holds. Let $\mathcal{A}$ be an n-state automaton over $\mathrm{BS}(1,q)$ and let p, p' be two states of $\mathcal{A}$. Let $\rho_{11} \in \mathsf{RetL}_k^{p\to p'}(\mathcal{A})$ with $|\rho_{11}| > f(n,k)$. There exist runs $\rho_{00}, \rho_{10}, \rho_{01} \in \mathsf{RetL}_k^{p\to p'}(\mathcal{A})$ and $d \in \mathbb{Z}$ so that:*

$$\begin{array}{ccc} \rho_{01} & \ll_d & \rho_{11} \\ \mathbin{\rotatebox[origin=c]{90}{$\ll$}} & & \mathbin{\rotatebox[origin=c]{90}{$\ll$}} \\ \rho_{00} & \ll_d & \rho_{10} \end{array} \qquad (2)$$

Here, one shows that a long run can be shortened independently in two ways: Going left in the diagram (2), and going down. Shortening the run by "going left" changes the production of the run by the same difference, up to a factor ℓ that may differ in the two rows. Lemma 5.5 applies Lemma 5.4 to construct small numbers in $[\mathsf{RetL}_k(\mathcal{A})]$ that are not divisible by a given m. Later, these numbers allow us to compute $\gcd([\mathsf{RetL}_k^{p\to p}(\mathcal{A})])$ and bound $F([\mathsf{RetL}_k^{p\to p}(\mathcal{A})])$.

▶ **Lemma 5.5.** *There is a polynomial f such that the following holds. Let $m \in \mathbb{Z}$. Let $\mathcal{A}$ be an n-state automaton over* $\mathrm{BS}(1,q)$ *and let p, p' be two states of $\mathcal{A}$. Suppose there is a number in $[\mathsf{RetL}_k^{p\to p'}(\mathcal{A})]$ not divisible by m; then there is also an $s \in [\mathsf{RetL}_k^{p\to p'}(\mathcal{A})]$ not divisible by m such that $|s| \leq q^{f(n,k)}$.*

Proof. Let f be the polynomial from Lemma 5.4. Let $\rho \in \mathsf{RetL}_k^{p\to p'}(\mathcal{A})$ be of minimal length such that m does not divide $[\rho]$. Suppose $|\rho| > f(n,k)$. Write $\rho_{11} = \rho$ and apply Lemma 5.4. By minimality of ρ_{11}, we get $[\rho_{00}] \equiv [\rho_{10}] \equiv [\rho_{01}] \equiv 0 \bmod m$. In particular, $\rho_{00} \ll_d \rho_{10}$ implies $d \equiv 0 \bmod m$. However, since $\rho_{01} \ll_d \rho_{11}$ and $[\rho_{11}] \not\equiv 0 \bmod m$, we get $d \not\equiv 0 \bmod m$, a contradiction. Hence, $|\rho| \leq f(n,k)$ and thus $|[\rho]| \leq q^{2f(n,k)}$ by Lemma 5.3. ◀

With Lemma 5.5 in hand, one can show Proposition 5.2 similarly to Lemma 4.6.

6 Recognizability

In this section, we prove Theorem 3.4. We first present a characterization of recognizability that is easily checkable for PE-regular subsets. It is well-known that a subset S of $\mathbb{Z}$ is recognizable if and only if there is a $k \in \mathbb{Z} \setminus \{0\}$ such that for every $s \in \mathbb{Z}$, we have $s \in S$ if and only if $s + k \in S$. Our characterization is an analog for Baumslag-Solitar groups.

A subset $S \subseteq \mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$ is called *k-periodic* if for every $s \in \mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$, we have (i) $s \in S$ if and only if $s(0,k) \in S$ and (ii) for every $\ell \in \mathbb{Z}$, we have $s \in S$ if and only if $s(q^\ell - q^{\ell+k}, 0) \in S$. In other words, membership in S is insensitive to (i) moving the cursor k positions and (ii) replacing a power of q by another power of q whose exponent differs by k. The set S is *periodic* if it is k-periodic for some $k \geq 1$. We show the following:

▶ **Proposition 6.1.** *A subset $S \subseteq \mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$ is recognizable if and only if S is periodic.*

The fact that recognizable sets are periodic is an easy exercise. For the converse, we show that the subgroup H of $G = \mathbb{Z}[\frac{1}{q}] \rtimes \mathbb{Z}$ generated by $(0,k)$ and all $(q^\ell - q^{\ell+k}, 0)$ for $\ell \in \mathbb{Z}$ is normal and the quotient G/H is finite. Then, S is recognized by the projection $G \to G/H$.

To decide whether a PE-regular $R \subseteq \mathrm{BS}(1,q)$ is recognizable, we show effective regularity of the set $N \subseteq \{\mathtt{a}\}^*$ of all words $\mathtt{a}^k$ such that R is *not* k-periodic. Then, we just have to check whether N contains all words $\mathtt{a}^k$ with $k \geq 1$, which is clearly decidable. Since R is PE-regular, the set $D = R(G\backslash R)^{-1} \cup (G\backslash R)R^{-1}$ is effectively PE-regular (Proposition 3.6). Then R is not k-periodic if and only if $(0,k) \in D$ or $(q^\ell - q^{\ell+k}, 0) \in D$ for some $\ell \in \mathbb{Z}$. The element $(0,k)$ has the pointed expansion $0^{\triangleleft}0^{k-1}0_{\bullet}$. The pointed expansions of $(q^\ell - q^{\ell+k}, 0)$ for $\ell \in \mathbb{Z}$ are exactly those words obtained from words $-0^r(q-1)^{k-1}0^s$ for $r, s \in \mathbb{N}$ by decorating one of the digits with $^{\triangleleft}$ and with $_{\bullet}$, and removing leading or trailing 0's. Therefore, it is easy to see that $T_1 = \{(0^{\triangleleft}0^{k-1}0_{\bullet}, \mathtt{a}^k) \mid k \geq 1\}$ and $T_2 = \{(\mathsf{pe}((q^\ell - q^{\ell+k}, 0)), \mathtt{a}^k) \mid \ell \in \mathbb{Z},\ k \geq 1\}$ are rational transductions. This implies that $N = T_1(\mathsf{pe}(D)) \cup T_2(\mathsf{pe}(D)) \subseteq \mathtt{a}^*$ is effectively regular. Then clearly, R is not k-periodic if and only if $\mathtt{a}^k \in N$.

References

1 Nathalie Aubrun and Jarkko Kari. Tiling problems on Baumslag-Solitar groups. In *Proceedings of Machines, Computations and Universality 2013 (MCU 2013)*, pages 35–46, 2013. doi:10.4204/EPTCS.128.12.

2 Laurent Bartholdi and Pedro V. Silva. Rational subsets of groups. *CoRR*, abs/1012.1532, 2010. Chapter 23 of the handbook AutoMathA (to appear). arXiv:1012.1532.

3 Gilbert Baumslag and Donald Solitar. Some two-generator one-relator non-Hopfian groups. *Bulletin of the American Mathematical Society*, 68(3):199–201, 1962. `doi:10.1090/S0002-9904-1962-10745-9`.

4 Galina Aleksandrovna Bazhenova. Rational sets in finitely generated nilpotent groups. *Algebra and Logic*, 39(4):215–223, 2000. `doi:10.1007/BF02681647`.

5 Jason P. Bell, Kathryn Hare, and Jeffrey Shallit. When is an automatic set an additive basis? *Proceedings of the American Mathematical Society, Series B*, 5(6):50–63, 2018. `doi:10.1090/bproc/37`.

6 Michèle Benois. Parties rationnelles du groupe libre. *CR Acad. Sci. Paris*, 269:1188–1190, 1969.

7 Jean Berstel. *Transductions and Context-Free Languages*. Teubner, 1979.

8 Valérie Berthé and Michel Rigo, editors. *Combinatorics, automata, and number theory*, volume 135 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2010.

9 Laura Ciobanu and Murray Elder. Solutions sets to systems of equations in hyperbolic groups are EDT0L in PSPACE. In *Proceedings of the 46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, pages 110:1–110:15, 2019. `doi:10.4230/LIPIcs.ICALP.2019.110`.

10 Jordi Delgado Rodríguez. *Extensions of free groups: algebraic, geometric, and algorithmic aspects*. PhD thesis, Universitat Politècnica de Catalunya. Facultat de Matemàtiques i Estadística, 2017.

11 Volker Diekert, Claudio Gutierrez, and Christian Hagenah. The existential theory of equations with rational constraints in free groups is PSPACE -complete. *Information and Computation*, 202(2):105–140, 2005. `doi:10.1016/j.ic.2005.04.002`.

12 Volker Diekert, Olga Kharlampovich, Markus Lohrey, and Alexei G. Myasnikov. Algorithmic Problems in Group Theory (Dagstuhl Seminar 19131). *Dagstuhl Reports*, 9(3):83–110, 2019. `doi:10.4230/DagRep.9.3.83`.

13 Volker Diekert and Jürn Laun. On computing geodesics in Baumslag-Solitar groups. *International Journal on Algebra and Computation*, 21(1-2):119–145, 2011. `doi:10.1142/S0218196711006108`.

14 Volker Diekert, Alexei G. Myasnikov, and Armin Weiß. Conjugacy in Baumslag's group, generic case complexity, and division in power circuits. In *Proceedings of 11th Latin American Symposium on Theoretical Informatics (LATIN 2014)*, pages 1–12, 2014. `doi:10.1007/978-3-642-54423-1_1`.

15 Volker Diekert, Igor Potapov, and Pavel Semukhin. Decidability of membership problems for flat rational subsets of $\mathrm{GL}(2, \mathbb{Q})$ and singular matrices, 2019. `arXiv:1910.02302`.

16 F. A. Dudkin and A. V. Treyer. Knapsack problem for Baumslag–Solitar groups. *Siberian Journal of Pure and Applied Mathematics*, 18:43–55, 2018. `doi:10.33048/pam.2018.18.404`.

17 Murray Elder. A linear-time algorithm to compute geodesics in solvable Baumslag–Solitar groups. *Illinois Journal of Mathematics*, 54(1):109–128, 2010. `doi:10.1215/ijm/1299679740`.

18 Murray Elder, Gillian Elston, and Gretchen Ostheimer. On groups that have normal forms computable in logspace. *Journal of Algebra*, 381:260–281, 2013. `doi:10.1016/j.jalgebra.2013.01.036`.

19 Seymour Ginsburg and Edwin H. Spanier. Bounded regular sets. *Proceedings of the American Mathematical Society*, 17(5):1043–1049, 1966. `doi:10.2307/2036087`.

20 Oscar H. Ibarra. Reversal-bounded multicounter machines and their decision problems. *Journal of the ACM*, 25(1):116–133, 1978. `doi:10.1145/322047.322058`.

21 Ilya Kapovich and Alexei Myasnikov. Stallings foldings and subgroups of free groups. *Journal of Algebra*, 248(2):608–668, 2002. `doi:10.1006/jabr.2001.9033`.

22 Olga Kharlampovich, Laura López, and Alexei Miasnikov. Diophantine problem in some metabelian groups, 2019. `arXiv:1903.10068`.

23 Dexter Kozen. Lower bounds for natural proof systems. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science (FOCS 1977)*, pages 254–266, 1977. doi:10.1109/SFCS.1977.16.

24 Markus Lohrey. The rational subset membership problem for groups: a survey. In C. M. Campbell, M. R. Quick, E. F. Robertson, and C. M. Roney-Dougal, editors, *Groups St Andrews 2013*, volume 422 of *Lond. Math. S.*, pages 368–389, Cambridge, United Kingdom, 2016. Cambridge University Press. doi:10.1017/CBO9781316227343.024.

25 Markus Lohrey and Géraud Sénizergues. Rational subsets in HNN-extensions and amalgamated products. *International Journal on Algebra and Computation*, 18(1):111–163, 2008. doi:10.1142/S021819670800438X.

26 Markus Lohrey and Georg Zetzsche. Knapsack in metabelian Baumslag-Solitar groups, 2020. arXiv:2002.03837.

27 Alexei Myasnikov, Andrey Nikolaev, and Alexander Ushakov. Knapsack problems in groups. *Mathematics of Computation*, 84:987–1016, 2015. doi:10.1090/S0025-5718-2014-02880-9.

28 David Robinson. *Parallel Algorithms for Group Word Problems*. PhD thesis, Department of Mathematics, University of Califoria, San Diego, 1993.

29 N. S. Romanovskiĭ. Some algorithmic problems for solvable groups. *Algebra and Logic*, 13:13–16, 1974. doi:10.1007/BF01462922.

30 N. S. Romanovskiĭ. The occurrence problem for extensions of abelian groups by nilpotent groups. *Siberian Mathematical Journal*, 21:273–276, 1980. doi:10.1007/BF00968275.

31 Géraud Sénizergues. On the rational subsets of the free group. *Acta Informatica*, 33(3):281–296, 1996. doi:10.1007/s002360050045.

32 John C Shepherdson. The reduction of two-way automata to one-way automata. *IBM Journal of Research and Development*, 3(2):198–200, 1959. doi:10.1147/rd.32.0198.

33 Pedro V. Silva. Free group languages: Rational versus recognizable. *RAIRO—Theoretical Informatics and Applications*, 38(1):49–67, 2004. doi:10.1051/ita:2004003.

34 Pedro V. Silva. An automata-theoretic approach to the study of fixed points of endomorphisms. In Ventura E. González-Meneses J., Lustig M., editor, *Algorithmic and Geometric Topics Around Free Groups and Automorphisms*, Advanced Courses in Mathematics—CRM Barcelona, pages 1–42. Birkhäuser, 2017. doi:10.1007/978-3-319-60940-9_1.

35 Armin Weiß. *On the Complexity of Conjugacy in Amalgamated Products and HNN Extensions*. PhD thesis, Institut für Formale Methoden der Informatik, Universität Stuttgart, 2015. doi:10.18419/opus-3538.

36 Herbert S. Wilf. A circle-of-lights algorithm for the “money-changing problem”. *The American Mathematical Monthly*, 85(7):562–565, 1978. doi:10.2307/2320864.