Knapsack in hyperbolic groups

Markus Lohrey

Universität Siegen, Siegen, Germany
lohrey@eti.uni-siegen.de

Abstract. Recently knapsack problems have been generalized from the integers to arbitrary finitely generated groups. The knapsack problem for a finitely generated group G is the following decision problem: given a tuple $(g, g_1, \ldots, g_k)$ of elements of G, are there natural numbers $n_1, \ldots, n_k \in \mathbb{N}$ such that $g = g_1^{n_1} \cdots g_k^{n_k}$ holds in G? Myasnikov, Nikolaev, and Ushakov proved that for every hyperbolic group, the knapsack problem can be solved in polynomial time. In this paper, it is shown that for every hyperbolic group G, the knapsack problem belongs to the complexity class LogCFL, and it is LogCFL-complete if G contains a free group of rank two. Moreover, it is shown that for every hyperbolic group G and every tuple $(g, g_1, \ldots, g_k)$ of elements of G the set of all $(n_1, \ldots, n_k) \in \mathbb{N}^k$ such that $g = g_1^{n_1} \cdots g_k^{n_k}$ in G is effectively semilinear.

1 Introduction

In [22], Myasnikov, Nikolaev, and Ushakov initiated the investigation of discrete optimization problems, which are usually formulated over the integers, for arbitrary (possibly non-commutative) groups. One of these problems is the *knapsack problem* for a finitely generated group G: The input is a sequence of group elements $g_1, \ldots, g_k, g \in G$ (specified by finite words over the generators of G) and it is asked whether there exists a tuple $(n_1, \ldots, n_k) \in \mathbb{N}^k$ such that $g_1^{n_1} \cdots g_k^{n_k} = g$ in G. For the particular case $G = \mathbb{Z}$ (where the additive notation $n_1 \cdot g_1 + \cdots + n_k \cdot g_k = g$ is usually preferred) this problem is NP-complete (resp., TC^0-complete) if the numbers $g_1, \ldots, g_k, g \in \mathbb{Z}$ are encoded in binary representation [11,9] (resp., unary notation [2]).

In [22], Myasnikov et al. encode elements of the finitely generated group G by words over the group generators and their inverses, which corresponds to the unary encoding of integers. There is also an encoding of words that corresponds to the binary encoding of integers, so called straight-line programs, and knapsack problems under this encoding have been studied in [18]. In this paper, we only consider the case where input words are explicitly represented. Here is a list of known results concerning the knapsack problem:

- Knapsack can be solved in polynomial time for every hyperbolic group [22]. In [4] this result was extended to free products of any finite number of hyperbolic groups and finitely generated abelian groups.

- There are nilpotent groups of class 2 for which knapsack is undecidable. Examples are direct products of sufficiently many copies of the discrete Heisenberg group $H_3(\mathbb{Z})$ [12], and free nilpotent groups of class 2 and sufficiently high rank [20].
- Knapsack for $H_3(\mathbb{Z})$ is decidable [12]. In particular, together with the previous point it follows that decidability of knapsack is not preserved under direct products.
- Knapsack is decidable for every co-context-free group [12], i.e., groups where the set of all words over the generators that do not represent the identity is a context-free language. Lehnert and Schweitzer [14] have shown that the Higman-Thompson groups are co-context-free.
- Knapsack belongs to NP for all virtually special groups (finite extensions of subgroups of graph groups) [19]. The class of virtually special groups is very rich. It contains all Coxeter groups, one-relator groups with torsion, fully residually free groups, and fundamental groups of hyperbolic 3-manifolds. For graph groups (also known as right-angled Artin groups) a complete classification of the complexity of knapsack was obtained in [19]: If the underlying graph contains an induced path or cycle on 4 nodes, then knapsack is NP-complete; in all other cases knapsack can be solved in polynomial time (even in LogCFL).
- Decidability of knapsack is preserved by finite extensions, HNN-extensions over finite associated subgroups and amalgamated products over finite subgroups [18].

In this paper we further investigate the knapsack problem in hyperbolic groups. The definition of hyperbolic groups requires that all geodesic triangles in the Cayley-graph are δ-slim for a constant δ; see Section 3 for details. The class of hyperbolic groups has several alternative characterizations (e.g., it is the class of finitely generated groups with a linear Dehn function), which gives hyperbolic groups a prominent role in geometric group theory. Moreover, in a certain probabilistic sense, almost all finitely presented groups are hyperbolic [8,23]. Also from a computational viewpoint, hyperbolic groups have nice properties: it is known that the word problem and the conjugacy problem can be solved in linear time [3,10]. As mentioned above, knapsack can be solved in polynomial time for every hyperbolic group [22]. Our first main result of this paper provides a precise characterization of the complexity of knapsack for hyperbolic groups: for every hyperbolic group, knapsack belongs to LogCFL, which is the class of all problems that are logspace-reducible to a context-free language. LogCFL has several alternative characterizations, see Section 5 for details. The LogCFL upper bound for knapsack in hyperbolic groups improves the polynomial upper bound shown in [22], and also generalizes a result from [15], stating that the word problem for a hyperbolic group is in LogCFL. For hyperbolic groups that contain a copy of a non-abelian free group (such hyperbolic groups are called non-elementary) it follows from [19] that knapsack is LogCFL-complete. Hyperbolic groups that contain no copy of a non-abelian free group (so called elementary hyperbolic groups) are known to be virtually cyclic, in which case knapsack can be shown to be in $\mathsf{NL} \subseteq \mathsf{LogCFL}$.

In Section 6 we prove our second main result: for every hyperbolic group G and every tuple $(g, g_1, \ldots, g_k)$ of elements of G the set of all $(n_1, \ldots, n_k) \in \mathbb{N}^k$ such that $g = g_1^{n_1} \cdots g_k^{n_k}$ in G is effectively semilinear. In other words: the set of all solutions of a knapsack instance in G is semilinear. Groups with this property are also called knapsack-semilinear. For the special case $G = \mathbb{Z}$ this is well-known (the set of solutions of a linear equation is Presburger definable and hence semilinear). Clearly, knapsack is decidable for every knapsack-semilinear group (due to the effectiveness assumption). In a series of recent papers it turned out that the class of knapsack-semilinear groups is surprisingly rich. It contains all virtually special groups [16] and all co-context-free group [12] and is closed under the following constructions: going to a finitely generated subgroup (this is trivial), going to a finite group extension [18], HNN-extensions over finite associated subgroups [18], amalgamated free products over finite subgroups [18], direct products (this follows from the closure of semilinear sets under intersection), and restricted wreath products [5].

Our proof of the knapsack-semilinearity of a hyperbolic group shows an additional quantitative statement: If the group elements $g, g_1, \ldots, g_k$ are represented by words over the generators and the total length of these words is N, then the set $\{(n_1, \ldots, n_k) \in \mathbb{N}^k \mid g = g_1^{n_1} \cdots g_k^{n_k} \text{ in } G\}$ has a semilinear representation, where all vectors only contain integers of size at most $p(N)$. Here, $p(x)$ is a fixed polynomial that only depends on G. Groups with this property are called knapsack-tame in [19]. In [19], it is shown that the class of knapsack-tame groups is closed under free products and direct products with $\mathbb{Z}$.

Missing proofs can be found in the long version [17].

2 General notations

We assume that the reader is familiar with basic concepts from group theory and formal languages. The empty word is denoted with ε. For a word $w = a_1 a_2 \cdots a_n$ let $|w| = n$ be the length of w, and for $1 \leq i \leq j \leq n$ let $w[i] = a_i$, $w[i : j] = a_i \cdots a_j$, $w[: i] = w[1 : i]$ and $w[i :] = w[i : n]$. Moreover, let $w[i : j] = \varepsilon$ for $i > j$.

We let $\mathbb{N} = \{0, 1, 2, \ldots\}$. A set of vectors $A \subseteq \mathbb{N}^k$ is *linear* if there exist vectors $v_0, \ldots, v_n \in \mathbb{N}^k$ such that $A = \{v_0 + \lambda_1 \cdot v_1 + \cdots + \lambda_n \cdot v_n \mid \lambda_1, \ldots, \lambda_n \in \mathbb{N}\}$. The tuple of vectors $(v_0, \ldots, v_n)$ is a *linear represention* of A. Its *magnitude* is the largest number appearing in one the vectors $v_0, \ldots, v_n$. A set $A \subseteq \mathbb{N}^k$ is *semilinear* if it is a finite union of linear sets $A_1, \ldots, A_m$. A *semilinear representation* of A is a list of linear representations for the linear sets $A_1, \ldots, A_m$. Its *magnitude* is the maximal magnitude of the linear representations for the sets $A_1, \ldots, A_m$. The magnitude of a semilinear set A is the smallest magnitude among all semilinear representations of A.

In the context of knapsack problems, we will consider semilinear subsets as mapping $f \colon \{x_1, \ldots, x_k\} \to \mathbb{N}$ for a finite set of variables $X = \{x_1, \ldots, x_k\}$. Such a mapping f can be identified with the vector $(f(x_1), \ldots, f(x_k))$. This allows to use all vector operations (e.g. addition and scalar multiplication) on the set $\mathbb{N}^X$ of all mappings from X to $\mathbb{N}$. The pointwise product $f \cdot g$ of two mappings $f, g \in \mathbb{N}^X$ is defined by $(f \cdot g)(x) = f(x) \cdot g(x)$ for all $x \in X$. Moreover, for mappings $f \in \mathbb{N}^X$, $g \in \mathbb{N}^Y$ with $X \cap Y = \emptyset$ we define $f \oplus g \colon X \cup Y \to \mathbb{N}$ by $(f \oplus g)(x) = f(x)$ for $x \in X$ and $(f \oplus g)(y) = g(y)$ for $y \in Y$. All operations on $\mathbb{N}^X$ will be extended to subsets of $\mathbb{N}^X$ in the standard pointwise way.

It is well-known that the semilinear subsets of $\mathbb{N}^k$ are exactly the sets definable in *Presburger arithmetic*. These are those sets that can be defined with a first-order formula $\varphi(x_1, \ldots, x_k)$ over the structure $(\mathbb{N}, 0, +, \leq)$ [7]. Moreover, the transformations between such a first-order formula and an equivalent semilinear representation are effective. In particular, the semilinear sets are effectively closed under Boolean operations.

3 Hyperbolic groups

Let G be a finitely generated group with the finite symmetric generating set Σ, i.e., $a \in \Sigma$ implies that $a^{-1} \in \Sigma$. The *Cayley-graph* of G (with respect to Σ) is the undirected graph $\Gamma = \Gamma(G)$ with node set G and all edges (g, ga) for $g \in G$ and $a \in \Sigma$. We view Γ as a geodesic metric space, where every edge (g, ga) is identified with a unit-length interval. It is convenient to label the directed edge from g to ga with the generator a. The distance between two points p, q is denoted with $d_\Gamma(p, q)$. For $g \in G$ let $|g| = d_\Gamma(1, g)$. For $r \geq 0$, let $\mathcal{B}_r(1) = \{g \in G \mid d_\Gamma(1, g) \leq r\}$.

Paths can be defined in a very general way for metric spaces, but we only need paths that are induced by words over Σ. Given a word $w \in \Sigma^*$ of length n, one obtains a unique path $P[w] \colon [0, n] \to \Gamma$, which is a continuous mapping from the real interval $[0, n]$ to Γ. It maps the subinterval $[i, i+1] \subseteq [0, n]$ isometrically onto the edge (g_i, g_{i+1}) of Γ, where g_i (resp., g_{i+1}) is the group element represented by the word $w[:i]$ (resp., $w[:i+1]$). The path $P[w]$ starts in $1 = g_0$ and ends in g_n (the group element represented by w). We also say that $P[w]$ is the unique path that starts in 1 and is labelled with the word w. More generally, for $g \in G$ we denote with $g \cdot P[w]$ the path that starts in g and is labelled with w. When writing $u \cdot P[w]$ for a word $u \in \Sigma^*$, we mean the path $g \cdot P[w]$, where g is the group element represented by u. A path $P \colon [0, n] \to \Gamma$ of the above form is geodesic if $d_\Gamma(P(0), P(n)) = n$; it is a (λ, ϵ)-*quasigeodesic* if $|a - b| \leq \lambda \cdot d_\Gamma(P(a), P(b)) + \varepsilon$ for all $a, b \in [0, n]$; and it is ζ-*local* (λ, ϵ)-*quasigeodesic* if $|a - b| \leq \lambda \cdot d_\Gamma(P(a), P(b)) + \varepsilon$ for all $a, b \in [0, n]$ with $|a - b| \leq \zeta$.

A word $w \in \Sigma^*$ is geodesic if the path $P[w]$ is geodesic, which means that there is no shorter word representing the same group element from G. Similarly, we define the notion of (ζ-local) (λ, ϵ)-quasigeodesic words. A word $w \in \Sigma^*$ is *shortlex reduced* if it is the length-lexicographically smallest word that represents the same group element as w. For this, we have to fix an arbitrary linear order on Σ. Note that if $u = xy$ is shortlex reduced then x and y are shortlex reduced too. For a word $u \in \Sigma^*$ we denote with $\mathsf{shlex}(u)$ the unique shortlex reduced word with $\mathsf{shlex}(u) = u$ in G.

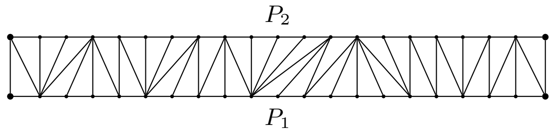


Fig. 1: Paths that asynchronously K-fellow travel

A *geodesic triangle* consists of three points $p, q, r \in G$ and geodesic paths $P_1 = P_{p,q}$, $P_2 = P_{p,r}$, $P_3 = P_{q,r}$ (the three sides of the triangle), where $P_{x,y}$ is a geodesic path from x to y. The geodesic triangle is *δ-slim* for $\delta \geq 0$, if for all $i \in \{1,2,3\}$, every point on P_i has distance at most δ from a point on $\bigcup_{j\in\{1,2,3\}\setminus\{i\}} P_j$. The group G is called *δ-hyperbolic*, if every geodesic triangle is δ-slim. Finally, G is hyperbolic, if it is δ-hyperbolic for some $\delta \geq 0$. Finitely generated free groups are for instance 0-hyperbolic. The property of being hyperbolic is independent of the chosen generating set Σ. The word problem for a hyperbolic group can be solved in real time [10].

Let us fix a δ-hyperbolic group G with the finite symmetric generating set Σ for the rest of the section, and let Γ be the corresponding geodesic metric space. We will apply a couple of well-known results for hyperbolic groups.

Lemma 1 ([6, 8.21]). *Let $g \in G$ be of infinite order and let $n \geq 0$. Let u be a geodesic word representing g. Then the word u^n is (λ, ϵ)-quasigeodesic, where $\lambda = N|g|$, $\epsilon = 2N^2|g|^2 + 2N|g|$ and $N = |\mathcal{B}_{2\delta}(1)|$.*

Consider paths $P_1 \colon [0, n_1] \to \Gamma$, $P_2 \colon [0, n_2] \to \Gamma$ and let K be a positive real number. We say that P_1 and P_2 *asynchronously K-fellow travel* if there exist two continuous non-decreasing mappings $\varphi_1 \colon [0,1] \to [0, n_1]$ and $\varphi_2 \colon [0,1] \to [0, n_2]$ such that $\varphi_1(0) = \varphi_2(0) = 0$, $\varphi_1(1) = n_1$, $\varphi_2(1) = n_2$ and for all $0 \leq t \leq 1$, $d_\Gamma(P_1(\varphi_1(t)), P_2(\varphi_2(t))) \leq K$. Intuitively, this means that one can travel along the paths P_1 and P_2 asynchronously with variable speeds such that at any time instant the current points have distance at most K. By slightly increasing K one obtains a ladder graph of the form shown in Figure 1, where the edges connecting the horizontal P_1- and P_2-labelled paths represent paths of length at most K that connect elements from G.

Lemma 2 ([21]). *Let P_1 and P_2 be (λ, ϵ)-quasigeodesic paths such that P_i starts in g_i and ends in h_i. Assume that $d_\Gamma(g_1, h_1), d_\Gamma(g_2, h_2) \leq h$. There exists a computable bound $K = K(\delta, \lambda, \epsilon, h) \geq h$ such that P_1 and P_2 asynchronously K-fellow travel.*

Finally we need the following lemma, see [17].

Lemma 3. *Fix constants λ, ϵ and let $\kappa = K(\delta, \lambda, \epsilon, 0)$ be taken from Lemma 2. Let $v_1, v_2 \in \Sigma^*$ be geodesic words and $u_1, u_2 \in \Sigma^*$ (λ, ϵ)-quasigeodesic words such that $v_1 u_1 = u_2 v_2$ in G. Consider a factorization $u_1 = x_1 y_1$ with $|x_1| \geq \lambda(|v_1| + 2\delta + \kappa) + \epsilon$ and $|y_1| \geq \lambda(|v_2| + 2\delta + \kappa) + \epsilon$ Then there exists a factorization $u_2 = x_2 y_2$ and $c \in \mathcal{B}_{2\delta+2\kappa}(1)$ such that $v_1 x_1 = x_2 c$ and $c y_1 = y_2 v_2$ in G.*

4 Knapsack problems

Let G be a finitely generated group with the finite symmetric generating set Σ. Moreover, let X be a set of variables that take values from $\mathbb{N}$. A *knapsack expression* over G is a formal expression of the form $E = u_1^{x_1} v_1 u_2^{x_2} v_2 \cdots u_k^{x_k} v_k$ with $k \geq 1$, $x_1, \ldots, x_k \in X$, $x_i \neq x_j$ for $i \neq j$, and $u_1, v_1, \ldots, u_k, v_k \in \Sigma^*$. Let $X_E = \{x_1, \ldots, x_k\}$ be the set of variables that occur in E. A solution for E is a mapping $\sigma \in \mathbb{N}^{X_E}$ such that the word $u_1^{\sigma(x_1)} v_1 u_2^{\sigma(x_2)} v_2 \cdots u_k^{\sigma(x_k)} v_k$ represents the identity element of G. With $\mathsf{sol}(E)$ we denote the set of all solutions of E. The *length* of E is defined as $|E| = \sum_{i=1}^{k} |u_i| + |v_i|$, whereas k is its *depth*. The *knapsack problem for* G is the following decision problem: Given a knapsack expression E over G, is $\mathsf{sol}(E)$ non-empty?

The group G is called *knapsack-semilinear* if for every knapsack expression E over G, the set $\mathsf{sol}(E)$ is semilinear and a semilinear representation can be effectively computed from E. The discrete Heisenberg group $H_3(\mathbb{Z})$ (which consists of all upper triangular (3×3)-matrices over the integers, where all diagonal entries are 1) is an example of a group which is not knapsack-semilinear, but for which the knapsack problem is decidable, see [12].

The group G is called *polynomially knapsack-bounded* if there is a fixed polynomial $p(n)$ such that for a given a knapsack expression E over G, one has $\mathsf{sol}(E) \neq \emptyset$ if and only if there exists $\nu \in \mathsf{sol}(E)$ with $\nu(x) \leq p(|E|)$ for all variables x in E. Finally, G is called *knapsack-tame* if there is a fixed polynomial $p(n)$ such that for a given a knapsack expression E over G one can compute a semilinear representation for $\mathsf{sol}(E)$ of magnitude at most $p(|E|)$. Thus, every knapsack-tame group is knapsack-semilinear as well as polynomially knapsack-bounded.

5 Complexity of knapsack in hyperbolic groups

In this section we show that for every hyperbolic group the knapsack problem belongs to the complexity class LogCFL. This class consists of all computational problems that are logspace reducible to a context-free language. The class LogCFL is included in the parallel complexity class NC^2 and has several alternative characterizations; see [24,26] for details. For our purposes, a characterization via AuxPDAs is most suitable. An AuxPDA (for auxiliary pushdown automaton) is a nondeterministic pushdown automaton with a two-way input tape and an additional work tape. Here we only consider AuxPDAs with the following two restrictions:

- The length of the work tape is restricted to $O(\log n)$ for an input of length n (logspace bounded).
- There is a polynomial $p(n)$, such that every computation path of the AuxPDA on an input of length n has length at most $p(n)$ (polynomially time bounded).

Whenever we speak of an AuxPDA in the following, we implicitly assume that the AuxPDA is logspace bounded and polynomially time bounded. The class of languages that are accepted by such AuxPDAs is exactly LogCFL [24]. A *one-way*

AuxPDA is an AuxPDA that never moves the input head to the left. Hence, in every step, the input head either does not move, or moves to the right.

In order to show that knapsack for a hyperbolic group belongs to LogCFL, we use the following important result from [22]:

Theorem 4 ([22]). *Every hyperbolic group is polynomially knapsack-bounded.*

This result is also a direct corollary of Theorem 7 from the next section.

In [15] it is shown that the word problem for a hyperbolic group belongs to LogCFL. Here, we extend the proof from [15] to the knapsack problem. First, we consider another problem of independent interest. An acyclic NFA is a nondeterministic finite automaton $\mathcal{A} = (Q, \Sigma, \Delta, q_0, F)$ (Q is a finite set of states, Σ is the input alphabet, $\Delta \subseteq Q \times \Sigma^* \times Q$ is the set of transition triples, $q_0 \in Q$ is the initial state, and $F \subseteq Q$ is the set of final states) such that the relation $\{(p, q) \in Q \times Q \mid \exists w \in \Sigma^* \colon (p, w, q) \in \Delta\}$ is acyclic. Note that we allow transitions labelled with words; this will be convenient in the proof of the next theorem. For a finitely generated group G with the finite generating set Σ (the concrete choice of Σ is not relevant), the *membership problem for acyclic NFAs over* G is the following computational problem: Given an acyclic NFA $\mathcal{A}$ with input alphabet Σ, does $\mathcal{A}$ accept a word $w \in \Sigma^*$ such that $w = 1$ in G?

Theorem 5. *Membership for acyclic NFAs over a hyperbolic group belongs to* LogCFL.

Proof. Let G be a hyperbolic group with the symmetric generating set Σ and let $\mathcal{A}$ be the input NFA. Let $W = \{w \in \Sigma^* \mid w = 1 \text{ in } G\}$ be the word problem for G. In [15] it is shown that W is a growing context-sensitive language, i.e., it can be generated by a grammar where all productions are strictly length-increasing (except for the start production $S \to \varepsilon$). Hence, by the main result of [1], W can be recognized by a one-way AuxPDA $\mathcal{P}$ in logarithmic space and polynomial time.

An AuxPDA for the membership problem for acyclic NFAs over G guesses a path in the NFA $\mathcal{A}$ and thereby simulates the AuxPDA $\mathcal{P}$ on the word spelled by the guessed path. If the final state of the input NFA $\mathcal{A}$ is reached and the AuxPDA $\mathcal{P}$ accepts at the same time, then the overall AuxPDA accepts. It is important that the AuxPDA $\mathcal{P}$ works one-way since the guessed path in $\mathcal{A}$ cannot be stored in logspace. This implies that the AuxPDA cannot re-access input symbols that have already been processed. The AuxPDA is clearly logspace bounded and polynomially time bounded since $\mathcal{A}$ is acyclic. $\square$

Theorem 6. *For every hyperbolic groups G, knapsack can be solved in* LogCFL. *Moreover, if G contains a copy of F_2 (the free group of rank 2) then knapsack for G is* LogCFL*-complete.*

Proof. Let G be a hyperbolic group. It is straightforward to present a logspace reduction from knapsack for G to the membership problem for acyclic NFAs. By Theorem 5, this proves the first statement of the theorem. Consider a knapsack expression $E = u_1^{x_1} v_1 u_2^{x_2} v_2 \cdots u_k^{x_k} v_k$ over G. By Theorem 4, there exists a polynomial $p(x)$ such that $\mathsf{sol}(E) \neq \emptyset$ if and only if there exists a solution $(n_1, \ldots, n_k) \in \mathsf{sol}(E)$ such that $n_i \leq p(|E|)$ for all $1 \leq i \leq k$. For our reduction it therefore suffices to construct an acyclic NFA for the finite language $\{u_1^{n_1} v_1 u_2^{n_2} v_2 \cdots u_k^{n_k} v_k \mid 0 \leq n_1, \ldots, n_k \leq p(|E|)\}$, which is easy (see also [19, Section 4.2.5]).

The second statement from the theorem follows from [19, Proposition 4.26], where it was shown that knapsack for F_2 is LogCFL-complete. $\square$

6 Hyperbolic groups are knapsack-semilinear

In this section, we prove the following strengthening of Theorem 4:

Theorem 7. *Every hyperbolic group is knapsack-tame.*

Let us remark that the total number of vectors in a semilinear representation can be exponential, even for the simplest case $G = \mathbb{Z}$. Take the (additively written) knapsack expression $E = x_1 + x_2 + \cdots + x_n - n$. Then $\mathsf{sol}(E)$ is finite and consists of $\binom{2n-1}{n} \geq 2^n$ vectors.

Let us fix a δ-hyperbolic group G for the rest of Section 6 and let Σ be a finite symmetric generating set for G.

6.1 Knapsack expressions of depth two

We first consider knapsack expressions of depth 2 where all powers are quasi-geodesic. It is well known that the semilinear sets are exactly the Parikh images of the regular languages. We need the following quantitative version of this result:

Theorem 8 ([25, Theorem 4.1], see also [13]). *Let k be a fixed constant. Given an NFA $\mathcal{A}$ over an alphabet of size k with n states, one can compute in polynomial time a semilinear representation of the Parikh image of $L(\mathcal{A})$. Moreover, all numbers appearing in the semilinear representation are polynomially bounded in n.*

Lemma 9. *Let λ and ϵ be fixed constants. For all geodesic words $u_1, v_1, u_2, v_2 \in \Sigma^*$ such that $u_1 \neq \varepsilon \neq u_2$ and u_1^n, u_2^n are (λ, ϵ)-quasigeodesic for all $n \geq 0$, the set $\{(x_1, x_2) \in \mathbb{N} \times \mathbb{N} \mid v_1 u_1^{x_1} = u_2^{x_2} v_2$ in $G\}$ is effectively semilinear with magnitude bounded by $p(|u_1| + |v_1| + |u_2| + |v_2|)$ for a fixed polynomial $p(n)$.*

Proof. Let $S := \{(x_1, x_2) \in \mathbb{N} \times \mathbb{N} \mid v_1 u_1^{x_1} = u_2^{x_2} v_2$ in $G\}$. We will define an NFA $\mathcal{A}$ over the alphabet $\{a_1, a_2\}$ such that the Parikh image of $L(A)$ is S. Moreover, the number of states of $\mathcal{A}$ is polynomial in $|u_1| + |u_2| + |v_1| + |v_2|$. This allows us to apply Theorem 8. We will allow transitions that are labelled with words (having length polynomial in $|u_1| + |u_2| + |v_1| + |v_2|$). Moreover, instead of writing in the transitions these words, we write their Parikh images (so, for instance, a transition $p \xrightarrow{a_1 a_2 a_1} q$ is written as $p \xrightarrow{(2,1)} q$.

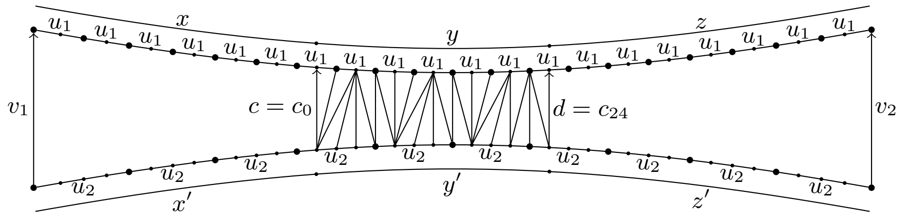


Fig. 2: Example for the construction from the proof of Lemma 9.

Let $\ell_i = |u_i|$ and $m_i = |v_i|$. Take the constant κ from Lemma 3 and define $N_1 = \lambda(m_1 + 2\delta + \kappa) + \epsilon$ and $N_2 = \lambda(m_2 + 2\delta + \kappa) + \epsilon$. We split the solution set S into $S_1 = S \cap \{(n_1, n_2) \in \mathbb{N} \times \mathbb{N} \mid n_1 < (N_1 + N_2)/\ell_1\}$ and $S_2 = S \setminus S_1$. For all $(n_1, n_2) \in S_1$ we have $|u_1^{n_1}| = n_1\ell_1 < N_1 + N_2$. Hence, $|\mathsf{shlex}(u_2^{n_2})| = |\mathsf{shlex}(v_1 u_1^{n_1} v_2^{-1})| < N_1 + N_2 + m_1 + m_2$. Since $u_2^{n_2}$ is (λ, ϵ)-quasigeodesic we get $|u_2^{n_2}| = n_2\ell_2 < \lambda(N_1 + N_2 + m_1 + m_2) + \epsilon$, i.e., $n_2 < (\lambda(N_1 + N_2 + m_1 + m_2) + \epsilon)/\ell_2$. Hence, S_1 is finite and its magnitude is bounded by $\mathcal{O}(m_1 + m_2)$.

We now deal with pairs $(n_1, n_2) \in S_2$. Consider such a pair (n_1, n_2) and the quasigeodesic rectangle consisting of the four paths $Q_1 = P[v_1]$, $P_1 = v_1 \cdot P[u_1^{n_1}]$, $P_2 = P[u_2^{n_2}]$, and $Q_2 = u_2^{n_2} \cdot P[v_2]$. Since $|u_1^{n_1}| \geq N_1 + N_2$, we factorize the word $u_1^{n_1}$ as $u_1^{n_1} = xyz$ with $|x| = N_1$ and $|z| = N_2$. By Lemma 3 we can factorize $u_2^{n_2}$ as $u_2^{n_2} = x'y'z'$ such that there exist $c, d \in \mathcal{B}_{2\delta+2\kappa}(1)$ with $v_1 x = x'c$ and $dz = z'v_2$ in G, see Figure 2 (where $n_1 = 20$, $n_2 = 10$, $\ell_1 = 2$ and $\ell_2 = 4$). Since $u_2^{n_2}$ is (λ, ϵ)-quasigeodesic, we have

$$|x'| \leq \lambda(m_1 + |x| + 2\delta + 2\kappa) + \epsilon = \lambda(m_1 + N_1 + 2\delta + 2\kappa) + \epsilon, \qquad (1)$$

$$|z'| \leq \lambda(m_2 + |z| + 2\delta + 2\kappa) + \epsilon = \lambda(m_2 + N_2 + 2\delta + 2\kappa) + \epsilon. \qquad (2)$$

Consider now the subpath P_1' of P_1 from $P_1(|x|)$ to $P_1(n_1\ell_1 - |z|)$ and the subpath P_2' of P_2 from $P_2(|x'|)$ to $P_2(n_2\ell_2 - |z'|)$. These are the paths labelled with y and y', respectively, in Figure 2. By Lemma 2 these paths asynchronously γ-fellow travel, where $\gamma := K(\delta, \lambda, \epsilon, 2\delta + 2\kappa)$ is a constant. In Figure 2 this is visualized by the part between the c-labelled edge and the d-labelled edge. W.l.o.g. we assume that $\gamma \geq 2\delta + 2\kappa$.

We now define the NFA $\mathcal{A}$ over the alphabet $\{a_1, a_2\}$ (recall the we replace edge labels from $\{a_1, a_2\}^*$ by their Parikh images). The state set of $\mathcal{A}$ is

$$Q = \{q_0, q_f\} \cup \{(i, b, j) \mid 0 \leq i < \ell_1, 0 \leq j < \ell_2, b \in \mathcal{B}_\gamma(1)\}.$$

The unique initial (resp., final) state is q_0 (resp., q_f). To define the transitions of $\mathcal{A}$ set $p = \lfloor N_1/\ell_1 \rfloor = \lfloor |x|/|u_1| \rfloor$, $r = N_1 \bmod \ell_1 = |x| \bmod |u_1|$, $s = \lceil N_2/\ell_1 \rceil = \lceil |z|/|u_1| \rceil$, $t = -N_2 \bmod \ell_1 = -|z| \bmod |u_1|$. Thus, we have $x = u_1^p u_1[: r]$ and $z = u_1^s[t + 1 :]$. There are the following types of transitions (transitions without a label are implicitly labelled by the zero vector $(0, 0)$), where $0 \leq i < \ell_1$, $0 \leq j < \ell_2$, $b, b' \in \mathcal{B}_\gamma(1)$.

1. $q_0 \xrightarrow{(p,p')} (r, c, r')$ if there exists a number $0 \leq k \leq \lambda(m_1 + N_1 + 2\delta + 2\kappa) + \epsilon$ (this is the possible range for the length of x' in (1)) such that $p' = \lfloor k/\ell_2 \rfloor$, $r' = k \bmod \ell_2$, and $v_1 u_1^p u_1[: r] = u_2^{p'} u_2[: r']c$ in G.
2. $(i, b, j) \to (i + 1, b', j)$ if $i + 1 < \ell_1$ and $bu_1[i + 1] = b'$ in G.
3. $(\ell_1 - 1, b, j) \xrightarrow{(1,0)} (0, b', j)$ if $bu_1[\ell_1] = b'$ in G.
4. $(i, b, j) \to (i, b', j + 1)$ if $j + 1 < \ell_2$ and $b = u_2[j + 1]b'$ in G.
5. $(i, b, \ell_2 - 1) \xrightarrow{(0,1)} (i, b', 0)$ if $b = u_2[\ell_2]b'$ in G.
6. $(t, d, t') \xrightarrow{(s,s')} q_f$ if there exists a number $0 \leq k \leq \lambda(m_2 + N_2 + 2\delta + 2\kappa) + \epsilon$ (this is the possible range for the length of z' in (2)) such that $s' = \lceil k/\ell_2 \rceil$, $t' = -k \bmod \ell_2$, and $du_1[t + 1 :]u_1^s = u_2[t' + 1 :]u_2^{s'} v_2$ in G.

The construction is best explained using the example in Figure 2. As mentioned above, the vertical lines between $c = c_0$ and $d = c_{24}$ represent the asynchronous γ-fellow travelling. The vertical lines are labelled with group elements $c_0, c_1, \ldots, c_{23}, c_{24} \in \mathcal{B}_\gamma(1)$ from left to right. In order to not overload the figure we only show c_0 and c_{24}. Note that $x = u_1^6 u_1[1]$, $x' = u_2^3 u_2[1]$, $z = u_1[2]u_1^7$, $z' = u_2[2 : 4]u_2^3$. Basically, the NFA $\mathcal{A}$ moves the vertical edges from left to right and thereby stores (i) the label c_i of the vertical edge, (ii) the position in the current u_2-factor where the vertical edge starts (position 0 means that we have just completed a u_2-factor), and (iii) the position in the current u_1-factor where the vertical edge ends. If a u_1-factor (resp., u_2-factor) is completed then the automaton makes a $(1, 0)$-labelled (resp., $(0, 1)$-labelled) transition. The complete run that corresponds to Figure 2 is:

$$\begin{aligned} q_0 \xrightarrow{(6,3)} & (1, c_0, 1) \xrightarrow{(1,0)} (0, c_1, 1) \to (1, c_2, 1) \to (2, c_3, 1) \to \\ & (3, c_4, 1) \xrightarrow{(0,1)} (0, c_5, 1) \xrightarrow{(1,0)} (0, c_6, 0) \to (0, c_7, 1) \to \\ & (1, c_8, 1) \xrightarrow{(1,0)} (0, c_9, 1) \to (1, c_{10}, 1) \to (1, c_{11}, 2) \to \\ & (1, c_{12}, 3) \xrightarrow{(0,1)} (1, c_{13}, 0) \xrightarrow{(1,0)} (0, c_{14}, 0) \to (0, c_{15}, 1) \to \\ & (1, c_{16}, 1) \xrightarrow{(1,0)} (0, c_{17}, 1) \to (1, c_{18}, 1) \to (1, c_{19}, 2) \to \\ & (1, c_{20}, 3) \xrightarrow{(1,0)} (0, c_{21}, 3) \xrightarrow{(0,1)} (0, c_{22}, 0) \to (0, c_{23}, 1) \to \\ & (1, c_{24}, 1) \xrightarrow{(8,4)} q_f \end{aligned}$$

With the above intuition it is straightforward to show that the Parikh image of $L(\mathcal{A})$ is indeed S_2. Also note that the number of states of $\mathcal{A}$ is bounded by $\mathcal{O}(\ell_1 \ell_2)$. The statement of the lemma then follows directly from Theorem 8. $\square$

6.2 Reduction to quasi-geodesic knapsack expressions

Let us call a knapsack expression $E = u_1^{x_1} v_1 u_2^{x_2} v_2 \cdots u_k^{x_k} v_k$ over G (λ, ϵ)-quasi-geodesic if all $u_1, \ldots, u_k, v_1, \ldots, v_k$ are geodesic and for all $1 \leq i \leq k$ and all

$n \geq 0$ the word u_i^n is (λ, ϵ)-quasigeodesic. We say that E has *infinite order*, if all u_i represent group elements of infinite order. The goal of this section is to reduce a knapsack expression to a finite number (in fact, exponentially many) of (λ, ϵ)-quasigeodesic knapsack expressions of infinite order for certain constants λ, ϵ:

Proposition 10. *There are fixed constants λ, ϵ such that from a given knapsack expression E over G one can compute a finite list of knapsack expressions E_i $(i \in I)$ over G such that*

- $\mathsf{sol}(E) = \bigcup_{i \in I} \big((m_i \cdot \mathsf{sol}(E_i) + d_i) \oplus \mathcal{F}_i\big)$,
- *every $\mathcal{F}_i$ is a semilinear subset of $\mathbb{N}^Y$ for a subset $Y \subseteq X_E$,*
- *the magnitude of every $\mathcal{F}_i$ is bounded by a constant that only depends on G,*
- *every E_i is a (λ, ϵ)-quasigeodesic knapsack expression of infinite order with variables from $Z := X_E \setminus Y$,*
- *the size of every E_i is bounded by $\mathcal{O}(|E|)$, and*
- *all m_i and d_i are vectors from $\mathbb{N}^Z$ where all entries are bounded by a constant that only depends on G (here, $m_i \cdot \mathsf{sol}(E_i) = \{m_i \cdot z \mid z \in \mathsf{sol}(E)\}$ and $m_i \cdot z$ is the pointwise multiplication of the vectors m_i and z).*

Once Proposition 10 is shown, we can conclude the proof of Theorem 7 by showing that all sets $\mathsf{sol}(E_i)$ are semilinear and that their magnitudes are bounded by $p(|E_i|)$ for a fixed polynomial $p(n)$. This will be achieved in the next section.

A detailed proof of Proposition 10 can be found in the long version [17]; here we only provide a sketch. Consider a knapsack expression $E = u_1^{x_1} v_1 u_2^{x_2} v_2 \cdots u_k^{x_k} v_k$. We can assume that every u_i is shortlex reduced. Let $g_i \in G$ be the group element represented by the word u_i. Reducing to the case, where all g_i have infinite order is relatively easy. In a hyperbolic group G the order of torsion elements is bounded by a fixed constant that only depends on G, see also the proof of [22, Theorem 6.7]). This allows to check for each g_i whether it has finite order, and to compute the order in the positive case. Let $Y \subseteq \{x_1, \ldots, x_k\}$ be those variables x_i such that g_i has finite order. For $x_i \in Y$ let $o_i < \infty$ be the order of g_i. Let $\mathcal{F}$ be the set of mappings $f \colon Y \to \mathbb{N}$ such that $0 \leq f(x_i) < o_i$ for all $x_i \in Y$. For every such mapping $f \in \mathcal{F}$ let E_f be the knapsack expression that is obtained from E by replacing for every $x_i \in Y$ the power $u_i^{x_i}$ by $u_i^{f(x_i)}$ (which is merged with the word v_i). Moreover, let $\mathcal{F}_f$ be the set of all mappings $g \colon Y \to \mathbb{N}$ such that $g(x_i) \equiv f(x_i) \bmod o_i$ for every $x_i \in Y$. Then the set $\mathsf{sol}(E)$ can be written as $\mathsf{sol}(E) = \bigcup_{f \in \mathcal{F}} \mathsf{sol}(E_f) \oplus \mathcal{F}_f$. Note that $\mathcal{F}_f$ is a semilinar set of magnitude $\mathcal{O}(1)$.

In a second step we reduce every E_f (which has infinite order) to (λ, ϵ)-quasigeodesic knapsack expressions for fixed constants λ and ϵ. Let us again write $E_f = u_1^{x_1} v_1 u_2^{x_2} v_2 \cdots u_k^{x_k} v_k$. We first use Lemma 1, which tells us that for every $n \geq 0$ and $1 \leq i \leq k$, the word u_i^n is (λ_i, ϵ_i)-quasigeodesic for $\lambda_i = N|u_i|$, $\epsilon_i = 2N^2|u_i|^2 + 2N|u_i|$. In order to reduce these λ_i, ϵ_i to fixed constants we mainly use the following two results from [3], where $L = 34\delta + 2$ and $K = |\mathcal{B}_{4\delta}(1)|^2$ (these are constants):

- Let $u = u_1 u_2$ be shortlex reduced, where $|u_1| \leq |u_2| \leq |u_1| + 1$, and $\tilde{u} = \mathsf{shlex}(u_2 u_1)$. If $|\tilde{u}| \geq 2L + 1$ then for every $n \geq 0$, the word $\tilde{u}^n$ is L-local $(1, 2\delta)$-quasigeodesic [3, Lemma 3.1].
- Let u be geodesic such that $|u| \geq 2L+1$ and for every $n \geq 0$, the word u^n is L-local $(1, 2\delta)$-quasigeodesic. Then one can compute $c \in \mathcal{B}_{4\delta}(1)$ and $1 \leq m \leq K$ such that $(\mathsf{shlex}(c^{-1} u^m c))^n$ is geodesic for all $n \geq 0$ [3, Section 3.2]. □

6.3 Proof of Theorem 7

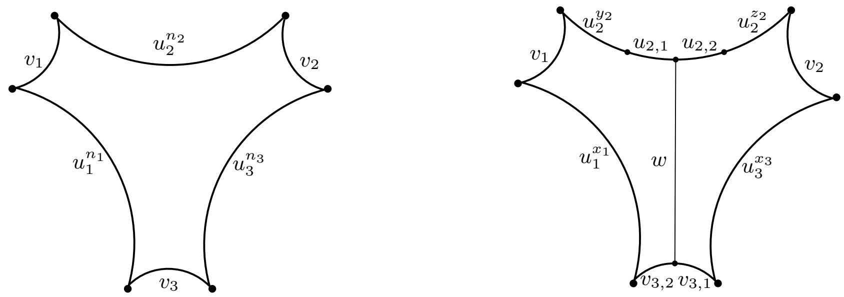
In this subsection we sketch the proof of Theorem 7; a detailed proof can be found in the full version [17]. Consider a knapsack expression $E = u_1^{x_1} v_1 u_2^{x_2} v_2 \cdots u_k^{x_k} v_k$. We can assume that all u_i, v_i are geodesic. By Proposition 10 we can moreover assume that for all $1 \leq i \leq k$, u_i represents a group element of infinite order and that u_i^n is (λ, ϵ)-quasigeodesics for all $n \geq 0$, where λ, ϵ are fixed constants. We want to show that $\mathsf{sol}(E)$ is semilinear and has a magnitude that is polynomially bounded by $|E|$.

For the case $k = 1$ we have to consider all $n \in \mathbb{N}$ with $u_1^n = v_1^{-1}$ in G. Since u_1 represents a group element of finite order there is at most one such n. Moreover, since u_i^n is (λ, ϵ)-quasigeodesic, such an n has to satisfy $|u_1| \cdot n \leq \lambda |v_1| + \epsilon$, which yields a linear bound on n. For the case $k = 2$ we can directly use Proposition 9. Now assume that $k \geq 3$. We want to show that the set $\mathsf{sol}(E)$ is a semilinear subset of $\mathbb{N}^k$ (later we will consider the magnitude of $\mathsf{sol}(E)$). For this we construct a Presburger formula with free variables $x_1, \ldots, x_k$ that is equivalent to $E = 1$. We do this by induction on the depth k. Therefore, we can use in our Presburger formula also knapsack equations of the form $F = 1$, where F has depth at most $k - 1$. One can also easily observe that it suffices to construct a Presburger formula for $\mathsf{sol}(E) \cap (\mathbb{N} \setminus \{0\})^k$.

Consider a tuple $(n_1, \ldots, n_k) \in \mathsf{sol}(E) \cap (\mathbb{N} \setminus \{0\})^k$ and the corresponding $2k$-gon that is defined by the (λ, ϵ)-quasigeodesic paths $P_i = (u_1^{n_1} v_1 \cdots u_{i-1}^{n_{i-1}} v_{i-1}) \cdot P[u_i^{n_i}]$ and the geodesic paths $Q_i = (u_1^{n_1} v_1 \cdots u_i^{n_i}) \cdot P[v_i]$, see Figure 3a for the case $k = 3$. Since all paths P_i and Q_i are (λ, ϵ)-quasigeodesic, we can apply [22, Lemma 6.4]: Every side of the $2k$-gon is contained in the h-neighborhoods of the other sides, where $h = \xi + \xi \log(2k)$ for a constant ξ that only depends on the constants $\delta, \lambda, \varepsilon$.

Let us now consider the side P_2 of the quasigeodesic $(2k)$-gon. It is labelled with $u_2^{x_2}$. Every point on P_2 must have distance at most h from one of the sides $P_1, Q_1, Q_2, P_3, \ldots, P_k, Q_k$. We distinguish several cases. In each case we cut the $2k$-gon into smaller pieces along paths of length $\leq 2h+1$ (in fact, length h except for one case), and these smaller pieces will correspond to knapsack expressions of depth $< k$. This is done until all knapsack expressions have depth at most two. Let us consider one typical case, the other cases are considered in the long version [17].

Assume that there is a point $p \in P_2$ that has distance at most h from a point $q \in Q_i$, where $3 \leq i \leq k$. The situation looks as shown in Figure 3b. For every tuple $t = (w, u_{2,1}, u_{2,2}, v_{i,1}, v_{i,2})$ such that $w \in \Sigma^*$ is of length at most h, $u_2 = u_{2,1} u_{2,2}$ and $v_i = v_{i,1} v_{i,2}$, we construct two new knapsack expressions $F_t =$



(a) The $2k$-gon for $k = 3$.

(b) Splitting the $2k$-gon into two parts.

Fig. 3: Planar diagrams from the proof of Theorem 7.

$u_1^{x_1} v_1 u_2^{y_2} (u_{2,1} w v_{i,2}) u_{i+1}^{x_{i+1}} v_{i+1} \cdots u_k^{x_k} v_k$, $G_t = u_{2,2} u_2^{z_2} v_2 u_3^{x_3} v_3 \cdots u_i^{x_i} (v_{i,1} w^{-1})$ and the formula

$$\bigvee_t \exists y_2, z_2 \colon x_2 = y_2 + 1 + z_2 \wedge F_t = 1 \wedge G_t = 1, \tag{3}$$

where t ranges over all tuples of the above form. Here y_2, z_2, y_i, z_i are new variables. Note that F_t and G_t have depth at most $k - 1$.

There are several other cases in which we can similarly split E into several (at most three) knapsack expressions of depth $< k$. In each case, we get a formula similar to (3), and we take the disjunction of all these formulas. This shows that $\mathsf{sol}(E)$ is semilinear.

It remains to argue that the magnitude of $\mathsf{sol}(E)$ is bounded polynomially in $|E|$. Iterating the splitting procedure results in a disjunction of formulas of the form

$$\exists y_1, \ldots, y_m \bigwedge_{i \in I} E_i = 1 \bigwedge_{j \in J} z_j = z'_j + z''_j + 1, \tag{4}$$

where every E_i is a knapsack expression of depth at most two. Moreover, for $i \neq j$, E_i and E_j have no common variables. The existentially quantified variables $y_1, \ldots, y_m$ are the new variables that were introduced when splitting factors $u_i^{x_i}$ (e.g., y_2, z_2 in (3)). The variables z_j, z'_j, z''_j in (4) are from $\{x_1, \ldots, x_k, y_1, \ldots, y_m\}$. The equations $z_j = z'_j + z''_j + 1$ in (4) result from the splitting of factors $u_i^{x_i}$. For instance, $x_2 = y_2 + 1 + z_2$ in (3) is one such equation.

In order to bound the magnitude of $\mathsf{sol}(E)$ it suffices to consider a single conjunctive formula of the form (4), since disjunction corresponds to union of semilinear sets, which does not increase the magnitude. We can also ignore the existential quantifiers in (4), because existential quantification corresponds to projection onto some of the coordinates, which cannot increase the magnitude. Hence, we have to consider the magnitude of the semilinear set A defined by the subformula $\bigwedge_{i \in I} E_i = 1 \bigwedge_{j \in J} z_j = z'_j + z''_j + 1$ of (4). To bound the magnitude of A, we show that (i) the size of every E_i in (4) is bounded by $\mathcal{O}(|E|^2)$ and (ii) that the size of the index set I is bounded by $\mathcal{O}(k^2)$. From (i) it follows that the magnitude of every set $\mathsf{sol}(E_i)$ is bounded polynomially in $|E|$. For the additional variables that are defined by the equations $z_j = z'_j + z''_j + 1$ in (4) one has to notice that these equations $z_j = z'_j + z''_j + 1$ result in a tree-shaped additive circuit whose input gates are the variables that appear in the E_i ($i \in I$). By (ii) this circuit has $\mathcal{O}(k^2)$ input gates. From this, one can finally deduce that the magnitude of the set A is indeed polynomially bounded in E. $\square$

7 More groups with knapsack in LogCFL

Let $\mathcal{C}$ be the smallest class of groups such that (i) every hyperbolic group belongs to $\mathcal{C}$, (ii) if $G \in \mathcal{C}$ then also $G \times \mathbb{Z} \in \mathcal{C}$, and (iii) if $G, H \in \mathcal{C}$ then also $G * H \in \mathcal{C}$ (where $G * H$ is the free product of G and H). From Theorem 7 and [19, Proposition 4.11 and 4.17] it follows that every group $G \in \mathcal{C}$ is knapsack-tame and hence polynomially knapsack-bounded. Hence, knapsack for a group $G \in \mathcal{C}$ is logspace reducible to membership for acyclic NFAs over G (the reduction in the proof of Theorem 6 works for any group). Finally, it was shown in the full version [17] that the word problem for every group in $\mathcal{C}$ can be accepted by a one-way AuxPDA in logarithmic space and polynomial time (the proof is essentially the same as in [19, Lemma 4.8]). This allows to generalize the proof of Theorem 5 to groups from $\mathcal{C}$. Hence, for every group $G \in \mathcal{C}$, membership for acyclic NFAs over G and knapsack for G can be solved in LogCFL.

8 Conclusion

In this paper, it is shown that every hyperbolic group is knapsack-tame and that the knapsack problem can be solved in LogCFL. Here is a list of open problems that one might consider for future work.

- For the following important groups, it is not known whether the knapsack problem is decidable: braid groups B_n (with $n \geq 3$), solvable Baumslag-Solitar groups $\mathsf{BS}_{1,p} = \langle a, t \mid t^{-1}at = a^p \rangle$ (with $p \geq 2$), and automatic groups which are not in any of the known classes with a decidable knapsack problem.
- In [12], it was shown that knapsack is decidable for every co-context-free group. The algorithm from [12] has an exponential running time. Is there a more efficient solution?
- Is there a polynomially knapsack-bounded group which is not knapsack-tame?

Acknowledgements. This work has been supported by the DFG research project LO 748/13-1.

References

1. Gerhard Buntrock and Friedrich Otto. Growing context-sensitive languages and Church-Rosser languages. *Information and Computation*, 141:1–36, 1998.

2. Michael Elberfeld, Andreas Jakoby, and Till Tantau. Algorithmic meta theorems for circuit classes of constant and logarithmic depth. *Electronic Colloquium on Computational Complexity (ECCC)*, 18:128, 2011.
3. David B. A. Epstein and Derek F. Holt. The linearity of the conjugacy problem in word-hyperbolic groups. *International Journal of Algebra and Computation*, 16(2):287–306, 2006.
4. Elizaveta Frenkel, Andrey Nikolaev, and Alexander Ushakov. Knapsack problems in products of groups. *Journal of Symbolic Computation*, 74:96–108, 2016.
5. Moses Ganardi, Daniel König, Markus Lohrey, and Georg Zetzsche. Knapsack problems for wreath products. In *Proceedings of STACS 2018*, volume 96 of *LIPIcs*, pages 32:1–32:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.
6. Etienne Ghys and Pierre de La Harpe. *Sur les groupes hyperboliques d'après Mikhael Gromov*. Progress in mathematics. Birkhäuser, 1990.
7. Seymour Ginsburg and Edwin H. Spanier. Semigroups, Presburger formulas, and languages. *Pacific Journal of Mathematics*, 16(2):285–296, 1966.
8. Mikhail Gromov. Hyperbolic groups. In S. M. Gersten, editor, *Essays in Group Theory*, number 8 in MSRI Publ., pages 75–263. Springer, 1987.
9. Christoph Haase. *On the complexity of model checking counter automata*. PhD thesis, University of Oxford, St Catherine's College, 2011.
10. Derek F. Holt. Word-hyperbolic groups have real-time word problem. *International Journal of Algebra and Computation*, 10:221–228, 2000.
11. Richard M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972.
12. Daniel König, Markus Lohrey, and Georg Zetzsche. Knapsack and subset sum problems in nilpotent, polycyclic, and co-context-free groups. In *Algebra and Computer Science*, volume 677 of *Contemporary Mathematics*, pages 138–153. American Mathematical Society, 2016.
13. Eryk Kopczynski and Anthony Widjaja To. Parikh images of grammars: Complexity and applications. In *Proceedings of LICS 2010*, pages 80–89. IEEE Computer Society, 2010.
14. Jörg Lehnert and Pascal Schweitzer. The co-word problem for the Higman-Thompson group is context-free. *Bulletin of the London Mathematical Society*, 39(2):235–241, 2007.
15. Markus Lohrey. Decidability and complexity in automatic monoids. *International Journal of Foundations of Computer Science*, 16(4):707–722, 2005.
16. Markus Lohrey and Georg Zetzsche. Knapsack in graph groups, HNN-extensions and amalgamated products. *CoRR*, abs/1509.05957, 2015.
17. Markus Lohrey. Knapsack in hyperbolic groups. *CoRR*, abs/1807.06774, 2018. URL: `https://arxiv.org/abs/1807.06774`.
18. Markus Lohrey and Georg Zetzsche. Knapsack in graph groups, HNN-extensions and amalgamated products. In *Proceedings of STACS 2016*, volume 47 of *LIPIcs*, pages 50:1–50:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.
19. Markus Lohrey and Georg Zetzsche. Knapsack in graph groups. *Theory of Computing Systems*, 62(1):192–246, 2018.
20. Alexei Mishchenko and Alexander Treier. Knapsack problem for nilpotent groups. *Groups Complexity Cryptology*, 9(1):87–98, 2017.
21. Alexei Myasnikov and Andrey Nikolaev. Verbal subgroups of hyperbolic groups have infinite width. *Journal of the London Mathematical Society*, 90(2):573–591, 2014.

22. Alexei Myasnikov, Andrey Nikolaev, and Alexander Ushakov. Knapsack problems in groups. *Mathematics of Computation*, 84:987–1016, 2015.
23. Alexander Yu. Ol'shanskii. Almost every group is hyperbolic. *International Journal of Algebra and Computation*, 2(1):1–17, 1992.
24. Ivan H. Sudborough. On the tape complexity of deterministic context–free languages. *Journal of the ACM*, 25(3):405–414, 1978.
25. Anthony Widjaja To. Parikh images of regular languages: Complexity and applications. *CoRR*, abs/1002.1464, 2010. URL: `http://arxiv.org/abs/1002.1464`.
26. Heribert Vollmer. *Introduction to Circuit Complexity*. Springer-Verlag, 1999.