Algorithmic and asymptotic properties of groups

Mark Sapir*

Abstract. This is a survey of the recent work in algorithmic and asymptotic properties of groups. I discuss Dehn functions of groups, complexity of the word problem, Higman embeddings, and constructions of finitely presented groups with extreme properties (monsters).

Mathematics Subject Classification (2000). Primary 20F65; Secondary 20F10.

Keywords. Turing machine, S-machine, Dehn function, Higman embedding, conjugacy problem, amenable group.

1. Introduction

Although the theory of infinite groups is very rich and full of powerful results, there are very few results having more influence on group theory and surrounding areas of mathematics (especially geometry and topology) as the following five.

- The Boone–Novikov theorem about existence of finitely presented groups with undecidable word problem [8], [35].
- The Higman theorem about embeddability of recursively presented groups into finitely presented groups [28];
- The Adian–Novikov solution of the Burnside problem [36].
- Gromov's theorem about groups with polynomial growth [23].
- Olshanskii and his students' theorems about existence of groups with all proper subgroups cyclic (Tarski monsters), and other finitely generated groups with extreme properties [38].

In this paper, I am going to survey the last ten years of my work on the topics related to these results.

Acknowledgement. Most of the work surveyed here is joint with J.-C. Birget, C. Druţu, V. Guba, A. Olshanskii and E. Rips. I am very grateful to them for cooperation.

*This work was supported in part by the NSF grants DMS 0245600 and DMS-0455881.

Proceedings of the International Congress
of Mathematicians, Madrid, Spain, 2006

2. *S*-machines

Recall that a Turing machine, say, with one tape is a triple (Y, Q, Θ) where Y is a tape alphabet, Q is the set of states, Θ is a set of commands (transitions) of the form $\theta = [U \to V]$ where U has the form vqu and V has the form $v'q'u'$. Here u, v, u' and v' are words in the tape alphabet, $q, q' \in Q$. A *configuration* of the Turing machine is a word wqw' where w, w' are words in the tape letters, q is a state letter. To apply the command $[U \to V]$ to a configuration, one has to replace U by V.

In order to specify a Turing machine with many tapes, one needs several disjoint sets of state letters. A *configuration* of the machine is a word of the form $u_1q_1u_2 \dots u_Nq_Nu_{N+1}$ where q_i are state letters, u_i are words in tape letters. Of course one needs to separate tapes. That can be done by using the special symbols (endmarkers) marking the beginning and the end of each tape. But these symbols can be treated as state letters as well. Every transition has the form $[U_1 \to V_1, \dots, U_N \to V_N]$, where $[U_i \to V_i]$ is a transition of a 1-tape machine.

Among all configurations of a machine M, one chooses one *accept* configuration W. Then a configuration W_1 is called *accepted* if there exists a *computation* $W_1 \to W_2 \to \cdots \to W_n = W$ where each step consists in application of a command of W.

Recall that the time function of a (non-deterministic) Turing machine is the smallest function $f(n)$ such that every accepted input w of size at most n requires at most $f(n)$ steps of the machine to be accepted.

The "common denominator" of the proofs of most of the results I am reviewing here is the notion of an S-machine that I introduced in [50]. Roughly speaking, S-machines make building groups with prescribed properties as easy as programming a Turing machine.

Essentially, an S-machine is simply an HNN-extension of a free group, although not every HNN-extension of a free group is an S-machine.

Let us start with an example that we shall call the *Miller machine*. It is the famous group of C. Miller [34]. Let $G = \langle X \mid R\rangle$ be a finitely presented group. The Miller machine is the group $M(G)$ generated by $X \cup \{q\} \cup \{\theta_x \mid x \in X\} \cup \{\theta_r \mid r \in R\}$ subject to the following relations

$$\theta x = x\theta, \quad \theta_x xq = qx\theta_x, \quad \theta_r q = qr\theta_r$$

where θ is any letter in $\Theta = \{\theta_x \mid x \in X\} \cup \{\theta_r \mid r \in R\}$. Clearly, this is an HNN-extension of the free group $\langle X, q\rangle$ with free letters $\theta \in \Theta$. The main feature of $M(G)$ discovered by Miller is that $M(G)$ *has undecidable conjugacy problem provided G has undecidable word problem.* In fact it is easy to see that qw is conjugated to q in $M(G)$ if and only if $w = 1$ in G.

To see that $M(G)$ can be viewed as a machine, consider any word uqv where u, v are words in $X \cup X^{-1}$. If we conjugate uqv by θ_r, we get the word $uqrv$ because $\theta_r q = qr\theta_r$ and θ_r commutes with u and v (here and below we do not distinguish words that are freely equal). Hence conjugation by θ_r amounts to executing a command $[q \to qr]$. Similarly, conjugation by θ_x amounts to executing a command $[q \to x^{-1}qx]$. If u ends with x, then executing this command means moving q one letter to the left. Thus conjugating words of the form uqv by θ's and their inverses, we can move the "head" q to the left and to the right, and insert relations from R.

The work of the Miller machine $M(G)$ can be drawn in the form of a diagram (see Figure 1) that we call a *trapezium*. It is a tessellation of a disc. Each cell corresponds to one of the relations of the group. The bottom layer of cells in Figure 1 corresponds to the conjugation by θ_x, the next layer corresponds to the conjugation by θ_r, etc. These layers are the so-called *θ-bands*. The bottom side of the boundary of the trapezium is labeled by the first word in the computation (uqv), the top side is labeled by the last word in the computation (q), the left and the right sides are labeled by the *history of computation*, the sequence of θ's and their inverses corresponding to the commands used in the *computation* $uqv \to \dots \to q$. The words written on the top and bottom sizes of the θ-bands are the intermediate words in the computation. We shall always assume that they are freely reduced.

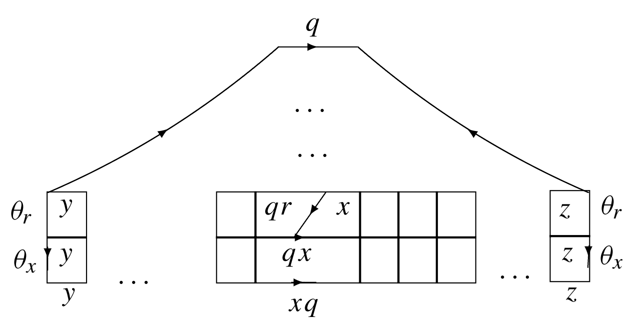

Figure 1. Trapezium of the Miller machine for a deduction $uqv \to \dots \to q$. Here $u = y\dots x$, $v = \dots z$.

The Miller machine has one tape and one state letter. General S-machines can have many tapes and many state letters. Here is a formal definition.

Let $F(Q, Y)$ be the free group generated by two sets of letters $Q = \bigcup_{i=1}^{N} Q_i$ and $Y = \bigcup_{i=1}^{N-1} Y_i$ where Q_i are disjoint and non-empty (below we always assume that $Q_{N+1} = Q_1$, and $Y_N = Y_0 = \emptyset$).

The set Q is called the set of q-letters, the set Y is called the set of a-letters.

In order to define an HNN-extension, we consider also a collection Θ of N-tuples of θ-letters. Elements of Θ are called *rules*. The components of θ are called *brothers* $\theta_1, \dots, \theta_N$. We always assume that all brothers are different. We set $\theta_{N+1} = \theta_1$, $Y_0 = Y_N = \emptyset$.

With every $\theta \in \Theta$, we associate two sequences of elements in $F(Q \cup Y)$: $B(\theta) = [U_1, \dots, U_N]$, $T(\theta) = [V_1, \dots, V_N]$, and a subset $Y(\theta) = \cup Y_i(\theta)$ of Y, where $Y_i(\theta) \subseteq Y_i$.

The words U_i, V_i satisfy the following restriction:

(*) For every $i = 1, \dots, N$, the words U_i and V_i have the form

$$U_i = v_{i-1}k_i u_i, \quad V_i = v'_{i-1}k'_i u'_i$$

where $k_i, k'_i \in Q_i$, u_i and u'_i are words in the alphabet $Y_i^{\pm 1}$, v_{i-1} and v'_{i-1} are words in the alphabet $Y_{i-1}^{\pm 1}$.

Now we are ready to define an S-machine $\mathcal{S}$ by generators and relations. The generating set X of the S-machine $\mathcal{S}$ consists of all q-, a- and θ-letters. The relations are:

$$U_i\theta_{i+1} = \theta_i V_i,\ i = 1, \dots, s, \quad \theta_j a = a\theta_j$$

for all $a \in Y_j(\theta)$. The first type of relations will be called (q, θ)-*relations*, the second type (a, θ)-*relations*.

Sometimes we will denote the rule θ by $[U_1 \to V_1, \dots, U_N \to V_N]$. This notation contains all the necessary information about the rule except for the sets $Y_i(\theta)$. In most cases it will be clear what these sets are: they are usually equal to either Y_i or $\emptyset$. By default $Y_i(\theta) = Y_i$.

Every S-rule $\theta = [U_1 \to V_1, \dots, U_s \to V_s]$ has an inverse $\theta^{-1} = [V_1 \to U_1, \dots, V_s \to U_s]$; we set $Y_i(\theta^{-1}) = Y_i(\theta)$.

Remark 2.1. Every S-machine is indeed an HNN-extension of the free group $F(Y, Q)$ with finitely generated associated subgroups. The free letters are θ_1 for every $\theta \in \Theta$. We leave it as an exercise to find the associated subgroups.

Every Turing machine T can be considered as an S-machine $S'(T)$ in the natural way: the generators of the free group are all tape letters and all state letters. The commands of the Turing machine are interpreted as rules of the S-machine. The main problem in that conversion is the following: there is a much bigger freedom in applying S-rules than in executing the corresponding commands of the Turing machine. Indeed, the Turing machine is in general not *symmetric* (i.e. if $[U \to V]$ is a command of the Turing machine then $[V \to U]$ is usually not) while every S-machine is symmetric. Another difference is that Turing machines work only with positive words, and S-machines work with arbitrary group words. Hence the language accepted by $S'(T)$ is usually much bigger than the language accepted by T.

Nevertheless, it can be proved that if T is symmetric, and a computation $w_1 \to w_2 \to \cdots$ of the S-machine $S'(T)$ involves only positive words, then that is a computation of T.

This leads to the following idea of converting any Turing machine T into an S-machine $S(T)$. First we construct a symmetric Turing machine T' that is equivalent to T (recognizes the same language). That is a fairly standard Computer Science trick (see [50]): the machine T' first guesses a computation of T, then executes it, then erases all the tapes. Note that the time function and the space function of T' are equivalent to the time function of T.

The second step is to compose the S-machine $S'(T')$ with a machine that checks positivity of a word. That machine starts working after every step of $S'(T')$. That is if an application of a rule of $S'(T')$ gives a non-positive (reduced) word then the checking machine does not allow the machine $S'(T')$ to proceed to the next step.

There are several checking machines. One of them – the *adding machine* – is very simple but its time function is exponential (see [42]). Another one is very complicated but it has a quadratic time function (see [50]).

Here is the definition of the adding machine. We present it here also in order to show an example of a program of an S-machine. It is not difficult to program an S-machine, but it does require some practice.

Let A be a finite set of letters. Let the set A_1 be a copy of A. It will be convenient to denote A by A_0. For every letter $a_0 \in A_0$, a_1 denotes its copy in A_1. The set of state letters of the adding machine $Z(A)$ is $P_1 \cup P_2 \cup P_3$ where $P_1 = \{L\}$, $P_2 = \{p(1), p(2), p(3)\}$, $P_3 = \{R\}$. The set of tape letters is $Y_1 \cup Y_2$ where $Y_1 = A_0 \cup A_1$ and $Y_2 = A_0$.

The adding machine $Z(A)$ has the following rules (there a is an arbitrary letter from A) and their inverses. The comments explain the meanings of these rules.

- $r_1(a) = [L \to L, p(1) \to a_1^{-1} p(1) a_0, R \to R]$.

  Comment. The state letter $p(1)$ moves left searching for a letter from A_0 and replacing letters from A_1 by their copies in A_0.

- $r_{12}(a) = [L \to L, p(1) \to a_0^{-1} a_1 p(2), R \to R]$.

  Comment. When the first letter a_0 of A_0 is found, it is replaced by a_1, and p turns into $p(2)$.

- $r_2(a) = [L \to L, p(2) \to a_0 p(2) a_0^{-1}, R \to R]$.

  Comment. The state letter $p(2)$ moves toward R.

- $r_{21} = [L \to L, p(2) \to p(1), R \to R]$, $Y_1(r_{21}) = Y_1$, $Y_2(r_{21}) = \emptyset$.

  Comment. $p(2)$ and R meet, the cycle starts again.

- $r_{13} = [L \to L, p(1) \to p(3), R \to R]$, $Y_1(r_{13}) = \emptyset$, $Y_2(r_{13}) = A_0$.

  Comment. If $p(1)$ never finds a letter from A_0, the cycle ends, $p(1)$ turns into $p(3)$; p and L must stay next to each other in order for this rule to be executable.

- $r_3(a) = [L \to L, p(3) \to a_0 p(3) a_0^{-1}, R \to R]$, $Y_1(r_3(a)) = Y_2(r_3(a)) = A_0$.

  Comment. The letter $p(3)$ returns to R.

The underlying algorithm of the adding machine is simple: the machine starts with a word $Lwp(1)R$, where w is a word in $A \cup A^{-1}$, L, $p(1)$, R are state letters. It considers the sequence of indexes of the letters in w as a binary number. The initial number is 0. The machine proceeds by adding 1 to this number until it produces $2^n - 1$ where n is the length of the word (each cycle of the machine adds a 1). After that, the machine returns the word to its initial state (all indexes are 0). If the initial word contained a negative letter, the state letter of the adding machine never becomes $p(3)$.

To compose a checking machine Z with an S-machine $\mathcal{S}$ means inserting state letters of Z between any two consecutive state letters of $\mathcal{S}$, and changing the rules of $\mathcal{S}$ in an appropriate way: every rule of $\mathcal{S}$ "turns on" the checking machines. After they finish their work, $\mathcal{S}$ can apply another rule (provided the word is still positive). If Z is a checking machine then the composition of $\mathcal{S}$ and Z is denoted by $\mathcal{S} \circ Z$.

The following results from [50] are very important for the applications. The equivalence of S-machines, their time functions, space functions, etc. are defined as for ordinary Turing machines.

We say that two increasing functions $f, g\colon \mathbb{Z}_+ \to \mathbb{Z}_+$ are equivalent if

$$\frac{1}{C} g\left(\frac{n}{C}\right) - Cn \leq f(n) \leq Cg(Cn) + Cn \tag{1}$$

for some constant C. We are not going to distinguish equivalent functions in this note. Thus $n^{3.2}$ is the same as $5n^{3.2}$ but different from $n^{3.2}\log n$.

Theorem 2.2 (Sapir, [50]). *Let T be a Turing machine. Then there exists an S-machine $\mathcal{S}$ that is polynomially equivalent to T. Moreover the time function of $\mathcal{S}$ is equivalent* (*in the sense of* (1)) *to the cube of the time function of T, the space function of $\mathcal{S}$ is equivalent to the time function of T.*

Moreover, one can use Miller's machines to simulate any Turing machine.

Theorem 2.3 (Sapir, [45]). *For every Turing machine T there exists a finitely presented group G such that the Miller machine $M(G)$ is polynomially equivalent to T.*

Thus any Turing machine can be effectively simulated by an S-machine with one tape and only one state letter.

3. Dehn functions and the word problem

3.1. The definition. Let $G = \langle X \mid R \rangle$ be a finitely presented group. We shall always assume that $X = X^{-1}$, R is a collection of words in the alphabet X closed under taking inverses and cyclic shifts, i.e. if $r \in R$, $r \equiv ab$ then $r^{-1} \in R$ and $ba \in R$.

The word problem in G asks, given a word w in X (i.e. a product of generators of G), whether w is equal to 1 in G. Clearly, the word problem in all "ordinary" groups is algorithmically decidable. For example, if the group is linear, then in order to check if $w = 1$, one can just multiply matrices representing the generators of G in the order of their appearance in w.

About 70 years ago, van Kampen noticed that $w = 1$ in G if and only if one can tessellate a disc with boundary labeled by w by tiles (cells) whose boundaries are labeled by words in R.

That tessellation is called a *van Kampen* diagram for w. It is also sometimes called *Dehn* or *disc* diagram of w. For example, the trapezium in Figure 1 is a van Kampen diagram over the S-machine $M(G)$ with boundary label $huqvh^{-1}q^{-1}$ where h is the history of the computation.

For every $w = 1$ in G, the area $\mathrm{a}(w)$ is the smallest number of cells in the van Kampen diagram for w, or the simplicial area of the (null-homotopic) loop labeled by w in the Cayley complex $\mathrm{Cayley}(G, X)$. Combinatorially, that is the smallest number of factors in any representation of w as a product of conjugates of the words from R. From the logic point of view, that is the length of the shortest "proof" that $w = 1$ in G (steps of the "proof" are insertions of relations from R into W).

It is easy to see that the word problem in G is decidable if and only if the area of a word w representing 1 in G is bounded from above by a recursive function in the length of w. Madlener and Otto [33] and, independently, Gersten [19] introduced a very basic characteristic of the algorithmic complexity of a group G, the Dehn function $\delta_G(n)$ of G: *it is the smallest function $d(n)$ such that the area of a word w of length $\le n$ representing 1 in G does not exceed $d(n)$.* Of course $\delta_G(n)$ depends on the choice of generating set X. But Dehn functions corresponding to different generating sets are *equivalent* in the sense of (1). Similarly, one can introduce the *isodiametric function* of G by looking at the diameter of van Kampen diagrams instead of their areas.

For example, the area of the trapezium in Figure 1 is approximately $|h|$ times the length of the longest θ-band in that trapezium, that can be interpreted as the product of the time of the computation by its space. That observation is the key to converting properties of the S-machine into the properties of the Dehn function.

3.2. The description. Dehn functions reflect in an easy and natural way both geometric and algorithmic properties of a group, so it is natural to ask which functions appear as Dehn functions of groups.

The first observation is not difficult.

Theorem 3.1 (See [50, Theorem 1.1]). *Every Dehn function of a finitely presented group G is* (*equivalent to*) *the time function of a Turing machine solving* (*non-deterministically*) *the word problem in G.*

The proof of this theorem in [50] is more complicated than it should have been. An easier proof can be obtained by using [45, Lemma 1].

Not every increasing function can be equivalent to a time function of a Turing machine. For example, if a time function $f(n)$ does not exceed a recursive function then it must be recursive. On the other hand, any "natural" function is the time function of a Turing machine. In particular, if $f(n)$ can be computed in time $\le f(n)$ then $f(n)$ is the time function of a deterministic Turing machine computing $f(n)$.

By a theorem proved by Gromov and Olshanskii among others, every finitely presented group with subquadratic Dehn function is in fact hyperbolic, so its Dehn function is linear. It is possible to deduce from a result of Kapovich and Kleiner [29] that if the Dehn function is subquadratic even on an infinite subset of natural numbers then the group is still hyperbolic. Dehn functions of nilpotent groups are bounded by a polynomial [5]. Another source of groups with polynomial Dehn function is the class of groups with simply connected asymptotic cones (Gromov, [24]). Moreover if the asymptotic cones are simply connected then the isodiametric function of the group is linear.

Recall that the asymptotic cone of a group G (see [24], [16] or [18]) is the ultra-limit (or Gromov–Hausdorff limit) of a sequence X/d_i where X is a Cayley graph of G, $\lim d_i = \infty$, X/d_i is the metric space X with distance function divided by d_i [23]. Asymptotic cones capture "global" geometric properties of the group G.

Of a particular interest are groups with quadratic Dehn function. That class of groups includes the classes of automatic groups and CAT(0)-groups. Higher dimensional Heisenberg groups [3], [43] and some solvable non-virtually nilpotent groups [17] also have quadratic Dehn functions. That class contains more complicated groups as well. The most striking example so far is the R. Thompson group

$$F = \langle x_0, x_1 \mid x_1^{x_0^2} = x_1^{x_0 x_1}, x_1^{x_0^3} = x_1^{x_0^2 x_1} \rangle$$

where $a^b = b^{-1}ab$. Recall that F is the group of all piecewise linear increasing self-homeomorphisms of the unit interval with finitely many dyadic singular points and all slopes powers of 2. Guba showed in [25] that F has a quadratic Dehn function. One of the most interesting unsolved problems about this class is whether $\mathrm{SL}_n(\mathbb{Z})$ belongs to it for $n \geq 4$.

A very non-trivial result of Bridson and Groves [11] shows that every cyclic extension of a finitely generated free group has quadratic Dehn function. On the other hand, Olshanskii and I proved [42] that HNN extensions of free groups having undecidable conjugacy problem must have Dehn function at least $n^2 \log n$. Together with the result of Bridson and Groves it gives another proof of decidability of the conjugacy problem for cyclic extensions of free groups [7].

It is still unknown whether every group with quadratic Dehn function has decidable conjugacy problem. Olshanskii and I gave a "quasi-proof" of that in [42].

I think that it is most probable that the class of Dehn functions $\geq n^2 \log n$ is as wide as the class of time functions of Turing machines. The next theorem confirms that conjecture in the case of Dehn functions $\geq n^4$.

Theorem 3.2 (See [50]). 1. *Let $\mathcal{D}_4$ be the set of all Dehn functions $d(n) \geq n^4$ of finitely presented groups. Let $\mathcal{T}_4$ be the set of time functions $t(n) \geq n^4$ of arbitrary Turing machines. Let $\mathcal{T}^4$ be the set of superadditive functions which are fourth powers of time functions. Then*

$$\mathcal{T}^4 \subseteq \mathcal{D}_4 \subseteq \mathcal{T}_4.$$

2. *For every time function $T(n)$ of a non-deterministic Turing machine with superadditive $T^4(n)$ there exists a finitely presented group G with Dehn function $T^4(n)$ and the isodiametric function $T^3(n)$.*

Recall that a function f is superadditive if $f(n+m) \geq f(m) + f(n)$ for any m, n. The question of whether all Dehn functions are superadditive is one of the unsolved mysteries of the subject. Together with Victor Guba [26], we proved that *the Dehn function of any non-trivial free product is superadditive.* Thus if there are non-superadditive Dehn functions then there are groups G such that G and $G * \mathbb{Z}$ have different Dehn functions!

Theorem 3.2 has many corollaries. For example, it implies that the *isoperimetric spectrum*, i.e. the set of α's such that $\lfloor n^\alpha \rfloor$ is a Dehn function, contains all numbers $\alpha \geq 4$ whose n-th digit can be computed by a deterministic Turing machine in time less than 2^{2^n}. All "constructible" numbers (rational numbers, algebraic numbers, values of elementary functions at rational points, etc.) satisfy this condition. On the other hand, Theorem 3.1 implies that if α is in the isoperimetric spectrum then the n-th digit of α can be computed in time $\leq 2^{2^{2^n}}$ (see [50] for details). The difference in the number of 2's in these expressions, is the difference between P and NP in Computer Science (if $P = NP$ then there should be two 2's in both expressions).

Note that before [50] has been submitted to Annals of Mathematics (in 1997), only a discrete set of non-integer numbers in the isoperimetric spectrum was known [10]. By the time the paper appeared in print (2002), that set increased by a dense subset in $[2, \infty)$ [9]. Groups in [10] with Dehn functions $\lfloor n^\alpha \rfloor$, $\alpha \notin \mathbb{N}$, have easier presentations than groups based on S-machines having the same Dehn functions, but the construction in [10], [9] is far from universal, and one cannot expect anything like Theorem 3.2 proved using their methods.

Other applications of Theorem 3.2 are:

- the first example of a finitely presented group with NP-complete word problem,
- examples of finitely presented groups with easy word problem (solvable in quadratic time) and arbitrary large (recursive) Dehn functions.

3.3. The proof. Here is how our construction from [50] works. Take any Turing machine M. Let M' be the symmetric Turing machine described above. Let $S(M')$ be the S-machine obtained as a composition of $S'(M')$ with a positivity checking S-machine from [50] working in quadratic time. The time function of $S(M')$ is T^3 and the space function is T where T is the time function of M. We can assume that the accepting configuration of $S(M')$ is some fixed word W of the form $k_1 w_1 k_2 w_2 \dots k_N$ where $N > 8$ (for some small cancellation reasons) and all w_i are copies of each other written in disjoint alphabets and containing no tape letters. That can be achieved by taking N copies of the initial Turing machine and making all of them work in parallel. Finally add one *hub* relation $W = 1$ to the S-machine $S(M')$. The resulting group $G(M)$ has Dehn function T^4 provided T^4 is superadditive, and isodiametric function T^3.

The main idea of the proof is the following. Take the standard trapezium corresponding to a computation $W_1 \to \cdots \to W_n = W$, identify its left and right sides (which have the same label). The resulting diagram has one hole with boundary label W. Insert the cell corresponding to the hub relation $W = 1$ into the hole. The result is a van Kampen diagram, called a *disc corresponding to the equality* $W_1 = 1$. The area of that diagram is equal to the area of the trapezium (plus 1). So it is equivalent to the product of the time of the computation by its space. The diameter of the disc with perimeter $\leq n$ is the time of the computation. Hence the worst area we can get is T^4, and the worst diameter is T^3. That gives the lower bound of the Dehn function and the isodiametric function. The upper bound is obtained by using certain surgeries on van Kampen diagrams. It turns out that every van Kampen diagram over the group $G(M)$ can be decomposed into a few discs and a diagram whose area is at most cubic (with respect to the perimeter of the original diagram). Thus if the area of a van Kampen diagram is large then most of the area is concentrated in the discs. It turns out also that the sum of the perimeters of the discs does not exceed a constant multiple of the perimeter of the diagram. This gives the desired upper bound of T^4 for the Dehn function (it is in this part of the proof where the superadditivity of T^4 is used) and T^3 for the isodiametric function.

3.4. The Dehn functions of S-machines and chord diagrams. It is easy to see that the Dehn function of an S-machine is at most cubic. Indeed, every van Kampen diagram with perimeter of length n over the presentation of an S-machine is covered by θ-bands that start and end on the boundary. There are also q-bands composed of (q, θ)-cells, and a-bands composed of the commutativity (a, θ)-cells. It can be proved that every θ-band intersects a q-band (an a-band) at most once. Hence the total number of (q, θ)-cells is at most n^2. Every other cell is an (a, θ)-cell. Each a-band starts on the boundary of the diagram or on the boundary of a (q, θ)-cell, so the total number of such bands is at most n^2 and the length of each of them is at most n. Hence the total area is at most n^3.

It was conjectured by Rips and myself that the Dehn function of an S-machine should in fact depend on the program of the S-machine. We thought that S-machines should provide examples of groups with Dehn functions strictly between n^2 and n^3. It turned out to be the case. In particular, Olshanskii and I proved in [42] that if $\mathcal{S}$ is any S-machine accepting language L, then the composition $\mathcal{S} \circ Z(A)$ of $\mathcal{S}$ and the adding machine has Dehn function at most $n^2 \log n$ and accepts the same language L. Hence we get the following result.

Theorem 3.3 ([42]). *There exists an S-machine with undecidable conjugacy problem and Dehn function $n^2 \log n$.*

The idea of analyzing van Kampen diagrams over S-machines is to show that if the area of a diagram is large then "most" of the area is inside large subtrapezia, and then analyze trapezia (i.e. computations of the S-machines). Note that in a trapezium, every θ-band intersects every q-band, thus the band structure of a trapezium is somewhat regular. In order to analyze irregular diagrams, Olshanskii introduced a measure of irregularity, the *dispersion*. In fact, the dispersion is an invariant of the *cord diagram* associated with every van Kampen diagram over an S-machine: the role of chords is played by the θ-bands (T-chords) and the q-bands (Q-chords). It is similar to a Vassiliev invariant of knots.

As I mentioned before, $n^2 \log n$ is the smallest Dehn function of an HNN extension of a free group with undecidable conjugacy problem. If the undecidability condition is dropped, one can construct Dehn functions strictly between n^2 and $n^2 \log n$. In particular, Olshanskii [40] constructed an *S-machine with non-quadratic Dehn function bounded from above by a quadratic function on arbitrary long intervals.* This gives the first example of a finitely presented group with two non-homeomorphic asymptotic cones [41]: *one of the asymptotic cones of this group is simply connected, and another one is not.*

Non-finitely presented groups with "very many" asymptotic cones are constructed in [18] using completely different methods.

Theorem 3.4 (Druţu, Sapir [18]). *There exist finitely generated groups with continuously many* (*maximal theoretically possible if the Continuum Hypothesis is true* [32]) *non-homeomorphic asymptotic cones.*

It is very interesting whether one can replace "finitely generated" by "finitely presented" in Theorem 3.4. One can try to use Higman embeddings from Section 4 to construct such examples.

3.5. Non-simply connected asymptotic cones. Note that Theorem 3.2 gave some of the first examples of groups with polynomial Dehn function and non-simply connected asymptotic cones because their Dehn functions can be polynomial (if the original Turing machine had polynomial time function) while their isodiametric functions are not linear. The first examples of groups with polynomial (cubic) Dehn functions, linear isodiametric functions and non-simply connected asymptotic cones were given in [44]. That answered a question of Druţu from [16].

The groups in [44] are S-machines. The easiest example is this:

$$G = \langle \theta_1, \theta_2, a, k \mid a^{\theta_i} = a, k^{\theta_i} = ka, i = 1, 2\rangle.$$

The S-machine has one tape letter, one state letter and two rules $[k \to ka]$ (and their inverses).

There is also an S-machine with Dehn function $n^2 \log n$ satisfying the same asymptotic properties. Note that $n^2 \log n$ cannot be lowered to n^2 because of a result of Papasoglu [49]: *all groups with quadratic Dehn functions have simply connected asymptotic cones.*

If a group has non-simply connected asymptotic cone, it is natural to ask what is its fundamental group. We do not know what are the fundamental groups of asymptotic cones of S-machines. These groups may provide some interesting invariants of S-machines and Turing machines, so it is worthwhile studying them.

The following theorem gives a partial answer to the question of what kind of groups can be fundamental groups of asymptotic cones of finitely generated groups.

Theorem 3.5 ([18]). *For every countable group C there exists an asymptotic cone of a finitely generated group G whose fundamental group is isomorphic to the free product of continuously many copies of C.*

The proof does not use S-machines but uses some small cancellation arguments. It would be interesting to find finitely presented groups with similar "arbitrary" fundamental groups of asymptotic cones. Perhaps the Higman embeddings discussed in the next section will help solving that problem. Another very interesting problem (due to Gromov [24]) is whether there exists an asymptotic cone of a finitely generated group with non-trivial but at most countable fundamental group.

4. Higman embeddings

The flexibility of S-machines allowed us to construct several versions of Higman embeddings (embeddings of recursively presented groups into finitely presented ones) preserving certain properties of the group.

4.1. An easy construction. The easiest known construction of a Higman embedding is the following. Let H be a recursively presented group $\langle X \mid R\rangle$. Then the set of all words in $X \cup X^{-1}$ that are equal to 1 in H is recursively enumerable. Hence we can assume that R consists of all these words. Then there exists an S-machine recognizing R. More precisely, for every word w in X, it starts with a word $q_1wq_2q_3\dots q_m$ and ends with a word $\bar{q}_1\bar{q}_2\dots\bar{q}_m$ if and only if $w \in R$.

Again, as in the proof of Theorem 3.2, we consider $N > 8$ copies of this S-machine and assume that the input of the S-machine $\mathcal{S}$ has the form

$$K(w) = k_1q_1wq_2\dots q_mk_2q_1'w'q_2'\dots q_m'k_3\dots k_{N+1}$$

and the accepting configuration

$$W = k_1\bar{q}_1\bar{q}_2\dots\bar{q}_mk_2\bar{q}_1'\dots k_{N+1}.$$

Here w', w'', ... are copies of w written in disjoint alphabets.

Let G be the group constructed as in the proof of Theorem 3.2 (see Section 3.3) by imposing the hub relation $W = 1$ on $\mathcal{S}$. Then the word $K(w)$ is equal to 1 in G if and only if $w \in R$.

Consider now another S-machine $\mathcal{S}'$ with input configuration

$$K'(w) = k_1q_1q_2\dots q_mk_2q_1'w'q_2'\dots q_m'k_3\dots k_{N+1}.$$

That machine works exactly like $\mathcal{S}$ in the part of the word between k_2 and k_{N+1}, and does nothing in the part between k_1 and k_2. Let G' be the group obtained by imposing the relation $W = 1$ on S'. Then $K'(w) = 1$ in G' if and only if $w \in R$.

Finally consider the amalgamated product $\mathcal{G} = G *_A G'$ where A is generated by all state and tape letters that appear in $K'(W)$. In that group, for every $w \in R$, both $K(w) = 1$ and $K'(w) = 1$. Hence $w = 1$. Thus there exists a natural homomorphism from H into $\mathcal{G}$. It is possible (and not too hard) to prove that this homomorphism is injective. Hence H is inside a finitely presented group G.

Another version of embedding used in [6] employs the so called *Aanderaa trick* [1]: instead of the amalgamated product, we used an HNN extension (see also the survey [45]).

Ones the embedding is established, it is important to understand which properties of a group H can be preserved.

4.2. Dehn functions and quasi-isometric Higman embeddings. First results have been obtained by Clapham [12] and Valiev [53] (see [46] for the history of these results): they proved that the solvability (even recursively enumerable degree) of the word problem and the level in the polynomial hierarchy of the word problem is preserved under some versions of Higman embedding.

In [6], Birget, Olshanskii, Rips and the author of this paper obtained a much stronger result.

Theorem 4.1 ([6]). *Let H be a finitely generated group with word problem solvable by a non-deterministic Turing machine with time function $\le T(n)$ such that $T(n)^4$ is superadditive. Then H can be embedded into a finitely presented group G with Dehn function $\le n^2T(n^2)^4$ in such a way that H has bounded distortion in G.*

This theorem immediately implies the following characterization of groups with word problem in *NP*.

Theorem 4.2 ([6]). *A finitely generated group H has word problem in NP if and only if H is embedded quasi-isometrically into a finitely presented group with polynomial Dehn function.*

Note that the "if" part of this theorem is trivial: if a finitely generated group is a (not necessarily quasi-isometric) subgroup of a group with polynomial Dehn function, its word problem is in *NP*. The converse part is highly non-trivial, although one can prove that the embedding described in Section 4.1 satisfies the desired properties (in [6], we used the Aanderaa trick).

From the logic point of view, Theorem 4.2 means that for every (arbitrary clever) algorithm solving the word problem in a finitely generated group, there exists a finitely presented group $G > H$ such that the word problem in H (and, moreover, in G) can be solved by the Miller machine $M(G)$ in approximately the same time as the initial algorithm.

4.3. Preserving the solvability of the conjugacy problem. The conjugacy problem turned out to be much harder to preserve under embeddings. Collins and Miller [14] and Gorjaga and Kirkinskiĭ[20] proved that even subgroups of index 2 of finitely presented groups do not inherit solvability or unsolvability of the conjugacy problem.

In 1976 D. Collins [31] posed the following question (Problem 5.22): *Does there exist a version of the Higman embedding theorem in which the degree of unsolvability of the conjugacy problem is preserved?* In [46], [47] we solved this problem affirmatively. In particular, we proved the following results.

Theorem 4.3 ([46]). *A finitely generated group H has solvable conjugacy problem if and only if it is Frattini embedded into a finitely presented group G with solvable conjugacy problem.*

Theorem 4.4 ([47]). *Every countable recursively presented group with solvable word and power problems is embeddable into a finitely presented group with solvable conjugacy and power problem.*

Recall that a subgroup H of a group G is Frattini embedded in G if every two elements of H that are conjugate in G are also conjugate inside H. We say that G has solvable *power problem* if there exists an algorithm which, given u, v in G says if $v = u^n$ for some $n \neq 0$.

Theorem 4.4 is a relatively easy application of Theorem 4.3.

The construction in [46] is much more complicated than in [6]. First we embed H into a finitely presented group H_1 preserving the solvability of the word problem. Then we use the Miller S-machine $M(H_1)$ to solve the word problem in H. In order to overcome technical difficulties, we needed certain parts of words appearing the computation to be always positive. The standard positivity checkers do not work because they are S-machines as well, and can insert negative letters! So we used some ideas from the original Boone–Novikov proofs. That required introducing new generators, x-letters (in addition to the a-, q-, and θ-letters in S-machines) and Baumslag–Solitar relations. In addition, to analyze the conjugacy problem in G, we had to consider annular diagrams which are more complicated than van Kampen disc diagrams. Different types of annular diagrams (spirals, roles, etc.) required different treatment.

We do not have any reduction of the complexity of the conjugacy problem in H to the complexity of the conjugacy problem in G. In particular, solving the conjugacy problem in G, in some cases required solving systems of equations in free groups (i.e. the Makanin–Razborov algorithm).

5. Non-amenable finitely presented groups

One of the most important applications of S-machines and Higman embeddings so far was the construction of a finitely presented counterexample to the von Neumann problem, i.e. a finitely presented non-amenable group without non-Abelian free subgroups [48].

5.1. Short history of the problem. Hausdorff [27] proved in 1914 that one can subdivide the 2-sphere minus a countable set of points into 3 parts A, B, C, such that each of these three parts can be obtained from each of the other two parts by a rotation, and the union of two of these parts can be obtained by rotating the third part. This implied that one cannot define a finitely additive measure on the 2-sphere which is invariant under the group SO(3). In 1924 Banach and Tarski [4] generalized Hausdorff's result by proving, in particular, that in $\mathbb{R}^3$, every two bounded sets A, B with non-empty interiors can be decomposed $A = \bigcup_{i=1}^n A_i$, $B = \bigcup_{i=1}^n B_i$ such that A_i can be rotated to B_i, $i = 1, \dots, n$ (the so called Banach–Tarski paradox). Von Neumann [54] was first who noticed that the cause of the Banach–Tarski paradox is not the geometry of $\mathbb{R}^3$ but an algebraic property of the group SO(3). He introduced the concept of an amenable group (he called such groups "measurable") as a group G which has a left invariant finitely additive measure μ, $\mu(G) = 1$, noticed that if a group is amenable then any set it acts upon freely also has an invariant measure and proved that a group is not amenable provided it contains a free non-Abelian subgroup. He also showed that groups like $\mathrm{PSL}(2, \mathbb{Z})$, $\mathrm{SL}(2, \mathbb{Z})$ contain free non-Abelian subgroups. So analogs of Banach–Tarski paradox can be found in $\mathbb{R}^2$ and even $\mathbb{R}$ (for a suitable group of "symmetries"). Von Neumann showed that the class of amenable groups contains Abelian groups, finite groups and is closed under taking subgroups, extensions, and infinite unions of increasing sequences of groups. Day [15] and Specht [51] showed that this class is closed under homomorphic images. The class of groups without free non-Abelian subgroups is also closed under these operations and contains Abelian and finite groups.

The problem of existence of non-amenable groups without non-Abelian free subgroups probably goes back to von Neumann and became known as the "von Neumann problem" in the fifties. Probably the first paper where this problem was formulated was the paper by Day [15]. It is also mentioned in the monograph by Greenleaf [21] based on his lectures given in Berkeley in 1967. Tits [52] proved that every non-amenable matrix group over a field of characteristic 0 contains a non-Abelian free subgroup. In particular every semisimple Lie group over a field of characteristic 0 contains such a subgroup.

First counterexamples to the von Neumann problem were constructed by Olshanskii [37]. He proved that the Tarsky monsters, both torsion-free and torsion (see [38]), are not amenable. Later Adian [2] showed that the non-cyclic free Burnside group of odd exponent $n \geq 665$ with at least two generators (that is the group given by the presentation $\langle a_1, \dots, a_m \mid u^n = 1$, where u runs over all words in the alphabet $\{a_1, \dots, a_m\}\rangle$) is not amenable.

Both Olshanskii's and Adian's examples are not finitely presented: in the modern terminology these groups are inductive limits of word hyperbolic groups, but they are not hyperbolic themselves. Since many mathematicians are mostly interested in groups acting "nicely" on manifolds, it is natural to ask if there exists a finitely presented non-amenable group without non-Abelian free subgroups. This question was explicitly formulated, for example, by Grigorchuk in [31] and by Cohen in [13]. This question is one of a series of similar questions about finding finitely presented "monsters", i.e. groups with unusual properties. Probably the most famous problem in that series is the (still open) problem about finding a finitely presented infinite torsion group. Other similar problems ask for finitely presented divisible group (group where every element has roots of every degree), finitely presented Tarski monster, etc. In each case a finitely generated example can be constructed as a limit of hyperbolic groups (see [38]), and there is no hope to construct finitely presented examples as such limits.

One difficulty in constructing a finitely presented non-amenable group without free non-Abelian subgroups is that there are "very few" known finitely presented groups without free non-Abelian subgroups. Most non-trivial examples are solvable or "almost" solvable (see [30]), and so they are amenable. The only previously known example of a finitely presented group without free non-Abelian subgroups for which the problem of amenability is non-trivial, is R. Thompson's group F (for the definition of F look in Section 3.2). The question of whether F is not amenable was formulated by R. Geoghegan in 1979. A considerable amount of work has been done to answer this question but it is still open.

5.2. The result. Together with A. Olshanskii, we proved the following theorem.

Theorem 5.1 ([48]). *For every sufficiently large odd n, there exists a finitely presented group $\mathcal{G}$ which satisfies the following conditions.*

1. *$\mathcal{G}$ is an ascending HNN extension of a finitely generated infinite group of exponent n.*

2. *$\mathcal{G}$ is an extension of a non-locally finite group of exponent n by an infinite cyclic group.*

3. *$\mathcal{G}$ contains a subgroup isomorphic to a free Burnside group of exponent n with 2 generators.*

4. *$\mathcal{G}$ is a non-amenable finitely presented group without free non-cyclic subgroups.*

Notice that parts 1 and 3 of Theorem 5.1 immediately imply part 2. By a theorem of Adian [2], part 3 implies that $\mathcal{G}$ is not amenable. Thus parts 1 and 3 imply part 4.

Note that the first example of a finitely presented group which is a cyclic extension of an infinite torsion group was constructed by Grigorchuk [22]. But the torsion subgroup in Grigorchuk's group does not have a bounded exponent and his group is amenable (it was the first example of a finitely presented amenable but not elementary amenable group).

5.3. The proof. Let us present the main ideas of our construction. We first embed the free Burnside group $B(m,n) = \langle \mathcal{B} \rangle$ of odd exponent $n \gg 1$ with $m > 1$ generators $\{b_1, \dots, b_m\} = \mathcal{B}$ into a finitely presented group $\mathcal{G}' = \langle \mathcal{C} \mid \mathcal{R} \rangle$ where $\mathcal{B} \subset \mathcal{C}$. This is done as in Section 4.1 using an S-machine recognizing all words of the form u^n. The advantage of S-machines is that such an S-machine can be easily and explicitly constructed (see [45]). Then we take a copy $\mathcal{A} = \{a_1, \dots, a_m\}$ of the set $\mathcal{B}$, and a new generator $\boldsymbol{t}$, and consider the group given by generators $\mathcal{C} \cup \mathcal{A}$ and the following three sets of relations:

(1) the set $\mathcal{R}$ of the relations of the finitely presented group $\mathcal{G}'$ containing $B(m,n)$;

(2) (u-relations) $y = u_y$, where u_y, $y \in \mathcal{C}$, is a certain word in $\mathcal{A}$ these words satisfy a very strong small cancellation condition; these relations make $\mathcal{G}'$ (and $B(m,n)$) embedded into a finitely presented group generated by $\mathcal{A}$;

(3) ($\boldsymbol{t}$-relations) $\boldsymbol{t}^{-1} a_i \boldsymbol{t} = b_i$, $i = 1, \dots, m$; these relations make $\langle \mathcal{A} \rangle$ a conjugate of its subgroup of exponent n (of course, the group $\langle \mathcal{A} \rangle$ gets factorized).

The resulting group $\mathcal{G}$ is obviously generated by the set $\mathcal{A} \cup \{\boldsymbol{t}\}$ and is an ascending HNN extension of its subgroup $\langle \mathcal{A} \rangle$ with the stable letter $\boldsymbol{t}$. Every element in $\langle \mathcal{A} \rangle$ is a conjugate of an element of $\langle \mathcal{B} \rangle$, so $\langle \mathcal{A} \rangle$ is an m-generated group of exponent n. This immediately implies that $\mathcal{G}$ is an extension of a group of exponent n (the union of increasing sequence of subgroups $\boldsymbol{t}^s \langle \mathcal{A} \rangle \boldsymbol{t}^{-s}$, $s = 1, 2, \dots$) by a cyclic group.

Hence it remains to prove that $\langle \mathcal{A} \rangle$ contains a copy of the free Burnside group $B(2,n)$.

In order to prove that, we construct a list of defining relations of the subgroup $\langle \mathcal{A} \rangle$. As we have pointed out, the subgroup $\langle \mathcal{A} \cup \mathcal{C} \rangle = \langle \mathcal{A} \rangle$ of $\mathcal{G}$ clearly satisfies all *Burnside relations* of the form $v^n = 1$. Thus we can add all Burnside relations

(4) $v^n = 1$ where v is a word in $\mathcal{A} \cup \mathcal{C}$

to the presentation of group $\mathcal{G}$ without changing the group.

If Burnside relations were the only relations in $\mathcal{G}$ among letters from $\mathcal{B}$, the subgroup of $\mathcal{G}$ generated by $\mathcal{B}$ would be isomorphic to the free Burnside group $B(m,n)$ and that would be the end of the story. Unfortunately there are many more relations in the subgroup $\langle \mathcal{B} \rangle$ of $\mathcal{G}$. Indeed, take any relation $r(y_1, \dots, y_s)$, $y_i \in \mathcal{C}$, of $\mathcal{G}$. Using u-relations (2), we can rewrite it as $r(u_1, \dots, u_s) = 1$ where $u_i \equiv u_{y_i}$. Then using $\boldsymbol{t}$-relations, we can substitute each letter a_j in each u_i by the corresponding letter $b_j \in \mathcal{B}$. This gives us a relation $r' = 1$ which will be called a relation *derived* from the relation $r = 1$, the operator producing derived relations will be called the $\boldsymbol{t}$-*operator*. We can apply the $\boldsymbol{t}$-operator again and again producing the second, third, …, derivatives $r'' = 1, r''' = 1, \dots$ or $r = 1$. We can add all *derived relations*

(5) $r' = 1, r'' = 1, \dots$ for all relations $r \in \mathcal{R}$

to the presentation of $\mathcal{G}$ without changing $\mathcal{G}$.

Now consider the group H generated by $\mathcal{C}$ subject to the relations (1) from $\mathcal{R}$, the Burnside relations (4) and the derived relations (5). The structure of the relations of H immediately implies that H contains subgroups isomorphic to $B(2, n)$. Thus it is enough to show that the natural map from H to $\mathcal{G}$ is an embedding.

The idea is to consider two auxiliary groups. The group $\mathcal{G}_1$ generated by $\mathcal{A} \cup \mathcal{C}$ subject to the relations (1) from $\mathcal{R}$, u-relations (2), the Burnside relations (4), and the derived relations (5). It is clear that $\mathcal{G}_1$ is generated by $\mathcal{A}$ and is given by relations (1) and (5) where every letter $y \in \mathcal{C}$ is replaced by the corresponding word u_y in the alphabet $\mathcal{A}$ plus all Burnside relations (4) in the alphabet $\mathcal{A}$. Let L be the normal subgroup of the free Burnside group $B(\mathcal{A}, n)$ (freely generated by $\mathcal{A}$) generated as a normal subgroup by all relators (1) from $\mathcal{R}$ and all derived relators (5) where letters from $\mathcal{C}$ are replaced by the corresponding words u_y. Then $\mathcal{G}_1$ is isomorphic to $B(\mathcal{A}, n)/L$.

Consider the subgroup U of $B(\mathcal{A}, n)$ generated (as a subgroup) by $\{u_y \mid y \in \mathcal{C}\}$. The words u_y, $y \in \mathcal{C}$, are chosen in such a way that the subgroup U is a free Burnside group freely generated by u_y, $y \in \mathcal{C}$, and it satisfies the *congruence extension* property, namely every normal subgroup of U is the intersection of a normal subgroup of $B(\mathcal{A}, n)$ with U.

All defining relators of $\mathcal{G}_1$ are inside U. Since U satisfies the congruence extension property, the normal subgroup $\bar{L}$ of U generated by these relators is equal to $L \cap U$. Hence $U/\bar{L}$ is a subgroup of $B(\mathcal{A}, n)/L = \mathcal{G}_1$. But by the choice of U, there exists a (natural) isomorphism between U and the free Burnside group $B(\mathcal{C}, n)$ generated by $\mathcal{C}$, and this isomorphism takes $\bar{L}$ to the normal subgroup generated by relators from $\mathcal{R}$ and the derived relations (5). Therefore $U/\bar{L}$ is isomorphic to H (since, by construction, H is generated by $\mathcal{C}$ subject to the Burnside relations, relations from $\mathcal{R}$ and derived relations)! Hence H is a subgroup of $\mathcal{G}_1$. Let $\mathcal{G}_2$ be the subgroup of H generated by $\mathcal{B}$.

Therefore we have

$$\mathcal{G}_1 \geq H \geq \mathcal{G}_2.$$

Notice that the map $a_i \to b_i$, $i = 1, \ldots, m$, can be extended to a homomorphism $\phi_{1,2}\colon \mathcal{G}_1 \to \mathcal{G}_2$. Indeed, as we mentioned above $\mathcal{G}_1$ is generated by $\mathcal{A}$ subject to Burnside relations, all relators from $\mathcal{R}$ and all derived relators (5) where letters from $\mathcal{C}$ are replaced by the corresponding words u_y. If we apply $\phi_{1,2}$ to these relations, we get Burnside relations and derived relations which hold in $\mathcal{G}_2 \leq H$.

The main technical statement of the paper shows that $\phi_{1,2}$ is an isomorphism, that is for every relation $w(b_1, \ldots, b_m)$ of $\mathcal{G}_2$ the relation $w(a_1, \ldots, a_m)$ holds in $\mathcal{G}_1$. This implies that the HNN extension $\langle \mathcal{G}_1, \boldsymbol{t} \mid \boldsymbol{t}^{-1}\mathcal{G}_1\boldsymbol{t} = \mathcal{G}_2\rangle$ is isomorphic to $\mathcal{G}$. Indeed, this HNN extension is generated by $\mathcal{G}_1$ and $\boldsymbol{t}$, subject to relations (1), (2), (4), (5) of $\mathcal{G}_1$ plus relations (3). So this HNN extension is presented by relations (1)–(5) which is the presentation of $\mathcal{G}$. Therefore $\mathcal{G}_1$ is a subgroup of $\mathcal{G}$, hence H is a subgroup of $\mathcal{G}$ as well.

The proof of the fact that $\phi_{1,2}$ is an isomorphism requires a detailed analysis of the group H. This group can be considered as a factor-group of the group H' generated by $\mathcal{C}$ subject to the relations (1) from $\mathcal{R}$ and derived relations (5) over the normal subgroup generated by Burnside relations (4). In other words, H is the *Burnside factor* of H'.

Burnside factors of free groups have been studied extensively starting with the celebrated paper by Adian and Novikov [36]. Later Olshanskii developed a geometric method of studying these factors in [38]. These methods were extended to arbitrary hyperbolic groups in [39]

The main problem we face in this paper is that H' is "very" non-hyperbolic. In particular, the set of relations $\mathcal{R}$ contains many commutativity relations, so H' contains non-cyclic torsion-free Abelian subgroups which cannot happen in a hyperbolic group.

We use a weak form of relative hyperbolicity that does hold in H'. In order to roughly explain this form of relative hyperbolicity used in the proof, consider the following example. Let $P = F_A \times F_B$ be the direct product of two free groups of rank m. Then the Burnside factor of P is simply $B(m,n) \times B(m,n)$. Nevertheless the theory of [38] cannot be formally applied to P. Indeed, there are arbitrarily thick rectangles corresponding to relations $u^{-1}v^{-1}uv = 1$ in the Cayley graph of P so diagrams over P are not A-maps in the terminology of [38] (i.e. they do not look like hyperbolic spaces). But one can obtain the Burnside factor of P in two steps. First we factorize F_A to obtain $Q = B(m,n) \times F_B$. Since F_A is free, we can simply use [38] to study this factor.

Now we consider all edges labeled by letters from A in the Cayley graph of Q as 0-edges, i.e. edges of length 0. As a result the Cayley graph of Q becomes a hyperbolic space (a tree). This allows us to apply the theory of A-maps from [38] to obtain the Burnside factor of Q. In fact Q is weakly relatively hyperbolic in the sense of our paper [48], i.e. it satisfies conditions (Z1), (Z2), (Z3) from the paper. The class of groups satisfying these conditions is very large and includes groups corresponding to S-machines considered in [48].

Recall that set $\mathcal{C}$ consists of tape letters, state letters, and command letters. In different stages of the proof some of these letters become 0-letters.

Trapezia corresponding to computations of the S-machine play central role in our study of the Burnside factor H of H'. As in [38], the main idea is to construct a graded presentation $\mathcal{R}'$ of the Burnside factor H of H' where longer relations have higher ranks and such that every van Kampen diagram over the presentation of H' has the so called property A from [38]. In all diagrams over the graded presentation of H, cells corresponding to the relations from $\mathcal{R}$ and derived relations are considered as 0-cells or cells of rank 1/2, and cells corresponding to Burnside relations from the graded presentation are considered as cells of ranks 1, 2,…. So in these van Kampen diagrams "big" Burnside cells are surrounded by "invisible" 0-cells and "small" cells.

The main part of property A from [38] is the property that if a diagram over $\mathcal{R}'$ contains two Burnside cells Π_1, Π_2 connected by a rectangular *contiguity* subdiagram Γ of rank 0 where the sides contained in the contours of the two Burnside cells are "long enough" then these two cells cancel, that is the union of Γ, Π, Π' can be replaced by a smaller subdiagram. This is a "graded substitute" to the classic property of small cancellation diagrams (where contiguity subdiagrams contain no cells).

In our case, contiguity subdiagrams of rank 0 turn out to be trapezia (after we clean them of Burnside 0-cells), so properties of contiguity subdiagrams can be translated into properties of the machine $\mathcal{S}$.

References

[1] Aanderaa, S., A proof of Higman's embedding theorem using Britton extensions of groups. In *Word Problems, Decision Problems and the Burnside problem in group theory*, North-Holland, Amsterdam, London 1973, 1–18.

[2] Adian, S. I., Random walks on free periodic groups. *Izv. Akad. Nauk SSSR Ser. Mat.* **46** (1982), 1139–1149.

[3] Allcock, D., An isoperimetric inequality for the Heisenberg groups. *Geom. Funct. Anal.* **8** (2) (1998), 219–233.

[4] Banach, S., Tarski, A., Sur la décomposition de ensembles de points an parties respectivement congruentes. *Fund. Math.* **6** (1924), 244–277.

[5] Baumslag, G., Miller III, C. F., Short, H., Isoperimetric inequalities and the homology of groups. *Invent. Math.* **113** (1993), 531–560.

[6] Birget, J.-C., Olshanskii, A. Y., Rips, E., Sapir, M. V., Isoperimetric functions of groups and computational complexity of the word problem. *Ann. of Math.* (2) **156** (2002), 467–518.

[7] Bogopolski, O., Martino, A., Maslakova, O., Ventura, E., Free-by-cyclic groups have solvable conjugacy problem. arXive, math.GR/0405178, 2004.

[8] Boone, W. W., Certain simple unsolvable problems in group theory I–VI. *Nederl. Akad. Wetensch. Proc. Ser. A* **57** (1954), 231–237; **57** (1954), 492–497; **58** (1955), 252–256; **58** (1955), 571–577; **60** (1957), 22–27; **60** (1957), 222–232.

[9] Brady, N., Bridson, M. R., There is only one gap in the isoperimetric spectrum. *Geom. Funct. Anal.* **10** (5) (2000), 1053–1070.

[10] Bridson, M. R., Fractional isoperimetric inequalities and subgroup distortion. *J. Amer. Math. Soc.* **12** (4) (1999), 1103–1118.

[11] Bridson, M. R., Groves, D., Free-group automorphisms, train tracks and the beaded decomposition. arXiv math.GR/0507589, 2005.

[12] Clapham, C. R. J., An embedding theorem for finitely generated groups. *Proc. London. Math. Soc.* (3) **17** (1967), 419–430.

[13] Cohen, J. M., Cogrowth and amenability of discrete groups. *J. Funct. Anal.* **48** (3) (1982), 301–309.

[14] Collins, D. J., Miller III, C. F., The conjugacy problem and subgroups of finite index. *Proc. London Math. Soc.* (3) **34** (3) (1977), 535–556.

[15] Day, Mahlon M., Amenable semigroups. *Illinois J. Math.* **1** (1957), 509–544.

[16] Druţu, C., Quasi-isometry invariants and asymptotic cones. International Conference on Geometric and Combinatorial Methods in Group Theory and Semigroup Theory (Lincoln, NE, 2000), *Internat. J. Algebra Comput.* **12** (1–2) (2002), 99–135.

[17] Druţu, C., Filling in solvable groups and in lattices in semisimple groups. *Topology* **43** (5) (2004), 983–1033.

[18] Druţu, C., Sapir, M. V., Tree-graded spaces and asymptotic cones of groups. With an appendix by Denis Osin and Mark Sapir. *Topology* **44** (5) (2005), 959–1058.

[19] Gersten, S. M., Dehn functions and l_1-norms for finite presentations. In *Algorithms and Classification in Combinatorial Group Theory* (ed. by G. Baumslag, C. F. Miller), Math. Sci. Res. Inst. Publ. 23, Springer-Verlag, New York 1992.

[20] Gorjaga, A. V., Kirkinskiĭ, A. S., The decidability of the conjugacy problem cannot be transferred to finite extensions of groups. *Algebra i Logika* **14** (4) (1975), 393–406 (in Russian).

[21] Greenleaf, F. P., *Invariant means on topological groups and their applications.* Van Nostrand Reinhold, New York 1969.

[22] Grigorchuk, R. I., An example of a finitely presented amenable group that does not belong to the class EG. *Mat. Sb.* **189** (1) (1998), 79–100.

[23] Gromov, M., Groups of polynomial growth and expanding maps. *Inst. Hautes Études Sci. Publ. Math.* **53** (1981), 53–73.

[24] Gromov, M., Asymptotic invariants of infinite groups. In *Geometric Group Theory* (ed. by G. A. Niblo and M. A. Roller), Volume 2, London Math. Soc. Lecture Note Ser. 182, Cambridge University Press, Cambridge 1993, 1–295.

[25] Guba, V. S., The Dehn Function of Richard Thompson's Group F is Quadratic. *Invent. Math.* **163** (2006), 313–342.

[26] Guba, V. S., Sapir, M. V., On Dehn functions of free products of groups. *Proc. Amer. Math. Soc.* **127** (7) (1999), 1885–1891.

[27] Hausdorff, F., *Grundzüge der Mengenlehre*. Veit & Company, Leipzig 1914.

[28] Higman, G., Subgroups of finitely presented groups. *Proc. Roy. Soc. Ser. A* **262** (1961), 455–475.

[29] Kapovich, M., Kleiner, B., Geometry of quasi-planes. Preprint, 2004.

[30] Kharlampovich, O. G., Sapir, M. V., Algorithmic problems in varieties. *Internat. J. Algebra Comput.* **5** (4–5) (1995), 379–602.

[31] *Kourovka Notebook*. *Unsolved Problems in Group Theory*. 5th edition, Novosibirsk 1976.

[32] Kramer, L., Shelah, S., Tent, K., Thomas, S., Asymptotic cones of finitely presented groups. *Adv. Math.* **193** (1) (2005), 142–173.

[33] Madlener, K., Otto, F., Pseudo-natural algorithms for the word problem for finitely presented monoids and groups. *J. Symbolic Comput.* **1** (1989), 383–418.

[34] Miller III, C. F., *On group-theoretic decision problems and their classification*. Ann. of Math. Stud. 68, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo 1971.

[35] Novikov, P. S., *On the algorithmic insolvability of the word problem in group theory*. Trudy Mat. Inst. im. Steklov. no. 44, Izdat. Akad. Nauk SSSR, Moscow 1955, 143 pp; English transl. Amer. Math. Soc. Transl. Ser. 9, Amer. Math. Soc., Providence, R. I., 1958, pp. 1–122.

[36] Novikov, P. S., Adian, S. I., Infinite periodic groups. I, II, III. *Izv. Akad. Nauk SSSR Ser. Mat.* **32** (1968) 212–244, 251–524, 709–731 (in Russian).

[37] Olshanskii, A. Yu., On the question of the existence of an invariant mean on a group. *Uspekhi Mat. Nauk* **35** (4) (1980), 199–200; English transl. *Russian Math. Surveys* **35** (4) (1980), 180–181.

[38] Olshanskii, A. Yu., *The geometry of defining relations in groups*. Nauka, Moscow 1989.

[39] Olshanskii, A. Yu., The SQ-universality of hyperbolic groups. *Mat. Sb.* **186** (8) (1995), 119–132; English transl. *Sb. Math.* **186** (8) (1995), 1199–1211.

[40] Olshanskii, A. Yu., Groups with quadratic-non-quadratic Dehn functions. arXiv, math.GR/0504349, 2005.

[41] Olshanskii, A. Yu., Sapir, M. V., A finitely presented group with two non-homeomorphic asymptotic cones. arXiv math.GR/0504350, 2005.

[42] Olshanskii, A. Yu., Sapir, M. V., Groups with small Dehn functions and bipartite chord diagrams. *Geom. Funct. Anal.*, to appear; arXiv, math.GR 0411174, 2004.

[43] Olshanskii, A. Yu., Sapir, M. V., Quadratic isometric functions of the Heisenberg groups. A combinatorial proof. *J. Math. Sci.* (*New York*) **93** (6) (1999), 921–927.

[44] Olshanskii, A. Yu., Sapir, M. V., Groups with non-simply connected asymptotic cones. arXiv, math.GR/0501542, 2005.

[45] Olshanskii, A. Yu., Sapir, M. V., Length and area functions on groups and quasi-isometric Higman embeddings. *Internat. J. Algebra Comput.* **11** (2) (2001), 137–170.

[46] Olshanskii, A. Yu., Sapir, M. V., The conjugacy problem and Higman embeddings. *Mem. Amer. Math. Soc.* **170** (804) (2004).

[47] Olshanskii, A. Yu., Sapir, M. V., Subgroups of finitely presented groups with solvable conjugacy problem. arXiv, math.GR/0405337, 2004.

[48] Olshanskii, A. Yu., Sapir, M. V., Non-amenable finitely presented torsion-by-cyclic groups. *Inst. Hautes Études Sci. Publ. Math.* **96** (2002), 43–169.

[49] Papasoglu, P., On the asymptotic cone of groups satisfying a quadratic isoperimetric inequality. *J. Differential Geom.* **44** (1996), 789–806.

[50] Sapir, M. V., Birget, J. C., Rips, E., Isoperimetric and isodiametric functions of groups. *Ann. of Math.* (2) **157** (2002), 345–466.

[51] Specht, W., Zur Theorie der messbaren Gruppen. *Math. Z.* **74** (1960), 325–366.

[52] Tits, J., Free subgroups of linear groups. *J. Algebra* **20** (1972), 250–270.

[53] Valiev, M. K., On polynomial reducibility of the word problem under embedding of recursively presented groups in finitely presented groups. In *Mathematical Foundations of Computer Science* 1975 (ed. by J. Bečvář), Lecture Notes in Comput. Sci. 32, Springer-Verlag, Berlin 1975, 432–438.

[54] von Neumann, J., Zur allgemeinen Theorie des Maßes. *Fund. Math.* **13** (1929), 73–116.

Department of Mathematics, Vanderbilt University, Nashville, TN 37240, U.S.A.
E-mail: m.sapir@vanderbilt.edu