TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 362, Number 9, September 2010, Pages 4655–4682
S 0002-9947(10)04959-7
Article electronically published on April 14, 2010

THE WORD AND GEODESIC PROBLEMS IN FREE SOLVABLE GROUPS

A. MYASNIKOV, V. ROMAN'KOV, A. USHAKOV, AND A. VERSHIK

Abstract. We study the computational complexity of the Word Problem (WP) in free solvable groups $S_{r,d}$, where $r \geq 2$ is the rank and $d \geq 2$ is the solvability class of the group. It is known that the Magnus embedding of $S_{r,d}$ into matrices provides a polynomial time decision algorithm for WP in a fixed group $S_{r,d}$. Unfortunately, the degree of the polynomial grows together with d, so the uniform algorithm is not polynomial in d. In this paper we show that WP has time complexity $O(rn \log_2 n)$ in $S_{r,2}$, and $O(n^3 rd)$ in $S_{r,d}$ for $d \geq 3$. However, it turns out, that a seemingly close problem of computing the geodesic length of elements in $S_{r,2}$ is NP-complete. We prove also that one can compute Fox derivatives of elements from $S_{r,d}$ in time $O(n^3 rd)$; in particular, one can use efficiently the Magnus embedding in computations with free solvable groups. Our approach is based on such classical tools as the Magnus embedding and Fox calculus, as well as on relatively new geometric ideas; in particular, we establish a direct link between Fox derivatives and geometric flows on Cayley graphs.

Contents

Received by the editors July 11, 2008.
2000 *Mathematics Subject Classification.* Primary 20F16; Secondary 68W30.
Key words and phrases. Free solvable groups, word problem, geodesic problem, Fox derivatives, Steiner tree problem, theoretical computer science.

1. Introduction

In this paper we study the computational complexity of several algorithmic problems related to the Word Problem (WP) in free solvable groups. Let $S_{r,d}$ be a free solvable group of rank $r \geq 2$ and of solvability class $d \geq 2$. We present here a uniform decision algorithm that solves WP in time $O(rn \log_2 n)$ in the free metabelian group $S_{r,2}$ (also denoted by M_r), and $O(n^3 rd)$ in the free solvable group $S_{r,d}$ for $d \geq 3$, where n is the length of the input word. In particular, this algorithm is at most cubic in n and linear in r and d for all free solvable groups $S_{r,d}$. Notice, that in all previously known polynomial time decision algorithms for WP in $S_{r,d}$, the degree of the polynomial grows together with d. In fact, we prove more; we show that one can compute Fox derivatives of elements from $S_{r,d}$ in time $O(n^3 rd)$. This allows one to use efficiently the Magnus embedding in computations with free solvable groups. On the other hand, we describe geodesics in $S_{r,d}$ and show that a seemingly close problem of finding the geodesic length of a given element from $S_{r,2}$ is surprisingly hard; it is NP-complete. Our approach is based on such classical tools as the Magnus embedding and Fox calculus, as well as, on relatively new (in group theory) geometric ideas from [13] and [54]. In particular, we establish a direct link between Fox derivatives and geometric flows on Cayley graphs.

The study of algorithmic problems in free solvable groups can be traced to the work [37] of Magnus, who in 1939 introduced an embedding (now called the Magnus embedding) of an arbitrary group of the type F/N' into a matrix group of a particular type with coefficients in the group ring of F/N (see Section 2.2 below). Since WP in free abelian groups is decidable in polynomial time, by induction, this embedding immediately gives a polynomial time decision algorithm for a fixed free solvable group $S_{r,d}$. However the degree of the polynomial here grows together with d.

In the 1950s, R. Fox introduced his free differential calculus and made the Magnus embedding much more transparent [21, 22, 23, 24] (see also Section 2.3). Namely, besides other things, he showed that an element $w \in F$ belongs to $N' = [N, N]$ if and only if all partial derivatives of w are equal to 0 in the integer group ring of F/N. This reduces WP in F/N' directly to the word problem in F/N. In particular, it solves, by induction, WP in $S_{r,d}$. Again, the decision algorithm is polynomial in a fixed group $S_{r,d}$, but the degree of the polynomial grows with d, which is not a surprise since the partial derivatives of w describe precisely the image of w under the Magnus embedding.

A few years later P. Hall proved the finiteness of all subdirect indecomposable finitely generated abelian-by-nilpotent groups. This implies that all finitely generated abelian-by-nilpotent, in particular, metabelian, groups are residually finite. About the same time, Gruenberg extended this result to arbitrary free solvable groups [28]. Now one can solve WP in $S_{r,d}$ in the following way. Given $w \in S_{r,d}$, as a word in the fixed set of generators, one can start two processes in parallel. The first one enumerates effectively all consequences of the defining relations of $S_{r,d}$ in

F_r (which is possible since the group is recursively presented) until the word w occurs, and the second one enumerates all homomorphisms from $S_{r,d}$ into all finite symmetric groups S_n (checking if a given r-tuple of elements in S_n generates a solvable group of class d) until it finds one where the image of w is nontrivial. However, computer experiments show that the algorithm described above is extremely inefficient (though its complexity is unknown).

Another shot at WP in metabelian groups comes from their linear representations. V. Remeslennikov proved in [46] that a finitely generated metabelian group (under some restrictions) is embeddable into $GL(n, R)$ for a suitable n and a suitable ring $R = K_1 \times \ldots \times K_n$ which is a finite direct product of fields K_i. In [55], see also [56], B. Wehrfritz generalized this result to arbitrary finitely generated metabelian groups G. It follows that G is embeddable into a finite direct product of linear groups. Since WP in linear groups is polynomial time decidable, this implies that WP in G is polynomial time decidable. Notice that it is unclear if there is a uniform polynomial time decision algorithm for WP in arbitrary finitely generated metabelian groups.

In comparison, observe that there are finitely presented solvable groups of class 3 with undecidable WP. In [35] O. Kharlampovich constructed the first example of such a group by reducing the halting problem for the universal Minski machine to WP of the group. There are several results which clarify the boundary between decidability and undecidability of the word problems in solvable groups; we refer to a survey [36] for details.

Our approach to WP in free solvable groups is based on the Fox Theorem mentioned above. Using binary tree search techniques and associative arrays we are able to compute Fox's derivatives of elements w of a free solvable group $S_{r,d}$ in time $O(n^3 d)$, where $n = |w|$. The significance of this result goes beyond WP for these groups; it gives a fast algorithm to compute images of elements under the Magnus embedding. This opens up an opportunity to solve effectively other algorithmic problems in groups $S_{r,d}$ using the classical techniques developed for wreath products of groups.

In the second half of the paper, Section 4, we study algorithmic problems on geodesics in free metabelian groups. Let G be a group with a finite set of generators $X = \{x_1, \ldots, x_r\}$ and $\mu : F(X) \to G$ the canonical epimorphism. For a word w in the alphabet $X^{\pm 1}$ by $|w|$ we denote the *length* of w. The *geodesic length* $l_X(g)$ of an element $g \in G$ relative to X is defined by

$$l_X(g) = \min\{|w| \mid w \in F(X), w^\mu = g\}.$$

We write, sometimes, $l_X(w)$ instead of $l_X(w^\mu)$. A word $w \in F(X)$ is called *geodesic* in G relative to X if $|w| = l_X(w)$. We are interested here in the following two algorithmic *search* problems in G.

The Geodesic Problem (GP): Given a word $w \in F(X)$, find a word $u \in F(X)$ which is geodesic in G such that $w^\mu = u^\mu$.

The Geodesic Length Problem (GLP): Given a word $w \in F(X)$, find $l_X(w)$.

Though GLP seems easier than GP, in practice, to solve GLP one usually solves GP first, and only then computes the geodesic length. It is an interesting question if there exists a group G and a finite set X of generators for G relative to which GP is strictly harder than GLP.

As is customary in complexity theory, one can modify the search problem GLP to get the corresponding bounded *decision* problem (that requires only answers "yes" or "no"):

The Bounded Geodesic Length Problem (BGLP): Let G be a group with a finite generating set X. Given a word $w \in F(X)$ and a natural number k, determine if $l_X(w) \leq k$.

In Section 4.1 we compare in detail the algorithmic "hardness" of the problems WP, BGLP, GLP, and GP in a given group G. Here we would like only to mention that in the list of the problems above each one is Turing reducible in polynomial time to the next one in the list, and GP is Turing reducible to WP in exponential time (see definitions in Section 4.1).

Among general facts on computational complexity of geodesics notice that if G has a polynomial *growth*, i.e., there is a polynomial $p(n)$ such that for each $n \in \mathbb{N}$ cardinality of the ball B_n of radius n in the Cayley graph $\Gamma(G, X)$ is at most $p(n)$, then one can easily construct this ball B_n in polynomial time with an oracle for WP in G. If, in addition, such a group G has WP decidable in polynomial time, then all the problems above have polynomial time complexity with respect to any finite generating set of G (since the growth and WP stay polynomial for any finite set of generators). Now, by Gromov's theorem [26], groups of polynomial growth are virtually nilpotent, hence linear, so they have WP decidable in polynomial time. It follows that all Geodesic Problems are polynomial time decidable in groups of polynomial growth (finitely generated virtually nilpotent groups). On the other hand, there are many groups of exponential growth where GP is decidable in polynomial time, for example, hyperbolic groups [16] or metabelian Baumslag-Solitar groups $BS(1, n) = \langle a, t \mid t^{-1}at = a^n \rangle$, $n \geq 2$ (see [15] and Section 4.1 for comments).

In general, if WP in G is decidable in polynomial time, then BGLP is in the class NP; i.e., it is decidable in polynomial time by a nondeterministic Turing machine.

It might happen though that BGLP in a group G is as hard as any in the class NP; i.e., it is NP-complete. The simplest example of this type is due to Parry, who showed in [45] that BGLP is NP-complete in the metabelian group $\mathbb{Z}_2 wr (\mathbb{Z} \times \mathbb{Z})$ (the wreath product of $\mathbb{Z}_2$ and $\mathbb{Z} \times \mathbb{Z}$). Correspondingly, the search problems GP and GLP are *NP-hard*, which means, precisely, that some (any) NP-complete problem is Turing reducible to them in polynomial time.

Our viewpoint on geodesics in free solvable groups is based on geometric ideas from the following two papers. In 1993, Droms, Lewin, and Servatius introduced a new geometric approach to study WP and GLP in groups of type F/N' via paths in the Cayley graph of F/N [13].

In 2004, Vershik and Dobrynin studied the algebraic structure of solvable groups, using the homology of related Cayley graphs [54]. This approach was outlined earlier in the papers [52, 53], where possible applications to random walks on metabelian groups have been discussed. In the papers [52, 53] (see also [54]) a new robust presentation of a free metabelian group $S_{r,2}$ was introduced as an extension of $\mathbb{Z}^r$ by the integer first homology group of the lattice $\mathbb{Z}^r$ (viewed as a one-dimensional complex) with a distinguished 2-cocycle. Similar presentations of other metabelian and solvable groups laid out foundations of a new approach to algorithmic problems in solvable groups.

It seems that these ideas are still underdeveloped in the group-theoretic context, despite their obvious potential. Meanwhile, in semigroup theory, similar geometric

techniques have been widely used to deal with free objects in semidirect products of varieties. One can find an explicit exposition of these techniques in the papers due to Almeida [1] and Almeida and Weil [2], while in [49, 4, 5, 3], Auinger, Rhodes and Steinberg use similar machinery on a regular basis. Earlier, similar methods, though sometimes implicitly, were used in inverse semigroup theory; we refer here to papers [43, 38, 39, 9].

In group theory most of the results in this area relied on various forms of the Magnus embedding and Fox derivatives. The role that the Magnus embeddings play in varieties of groups was clarified by Shmel′kin [50]. In [40] Matthews proved that the conjugacy problem (CP) in free metabelian groups is decidable, and Kargapolov and Remeslennikov generalized this to free solvable groups $S_{r,d}$ [33]. A few years later Remeslennikov and Sokolov described precisely the image of F/N' under the Magnus embedding and showed that CP is residually finite in $S_{r,d}$ [48]. We refer to a survey [47] on algorithmic problems in solvable groups.

In Sections 2.4 and 2.5 we study elements of groups of the type F/N' via flows on the Cayley graph Γ of F/N. It turns out that the flow generated by a word $w \in F$ on the graph Γ directly corresponds to the Fox derivatives of w in the group ring $\mathbb{Z}F/N$. This simple observation links together the techniques developed in group theory for the Magnus embeddings with the extensive geometric and the graph-theoretic machinery for flows. Indeed, the set of geometric circulations (flows where the Kirchhoff law holds for all vertices, including the source and the sink) forms a group which is naturally isomorphic to the first homology group $H_1(\Gamma, \mathbb{Z})$ of Γ. In this context the geometric circulations on Γ represent precisely the 1-cycles of Γ (viewed as a 1-complex). The classical result in homology theory describes $H_1(\Gamma, \mathbb{Z})$ as the abelianization of the fundamental group $\pi_1(\Gamma)$, which, in this case, is isomorphic to the free group N. Putting all these together one has another geometric proof of the Fox theorem, as well as the description of the kernel of the Magnus embedding.

In Section 2.7 we describe geodesics in groups F/N' as Euler tours in some finite subgraphs of Γ generated by the supports of the flows of the elements of F/N' on Γ. The description is geometric, explicit, and it gives a natural way to compute the geodesic length of elements. In this part geometric ideas seem unavoidable. However, this simplicity becomes treacherous when one is concerned with the efficiency of computations.

We prove that BGLP (relative to the standard basis) is NP-complete even in $S_{r,2}$. Consequently, the problems GP and GLP are NP-hard in $S_{r,2}$. To show this we construct a polynomial-time reduction of the Rectilinear Steiner Tree Problem (RSTP), which is NP-complete, to BGLP in $S_{r,2}$. The necessary information on RSTP is outlined in Section 4.3, and the proof of the main theorem is in Section 4.4. Notice that in [13], GLP was claimed to be polynomial time decidable in arbitrary finitely generated free solvable groups, but the argument turned out to be fallacious.

In the second half of the 20th century, free solvable groups, as well as solvable wreath products of groups and finitely generated metabelian groups, were intensely studied, but mostly from the viewpoint of combinatorial group theory.

Now they stand at the heart of research in various areas of algebra. On the one hand, the rejuvenated interest in these groups stems from random walks on groups and cohomology theory. For example, wreath products of abelian groups

give exciting examples and counterexamples to several conjectures on the numerical characteristics of random walks. It seems that the main reasons that facilitate research here come from some paradoxical properties of the groups themselves: all these groups are amenable (as solvable groups), but they have exponential growth and may have nontrivial Poisson boundary [32], etc. These groups, contrary to, say, free nilpotent groups, may have irreducible unitary representations with nontrivial cohomology. Some numerical characteristics of these groups are very intriguing, giving new exciting examples in quantitative group theory. For example, metabelian "lamplighter" groups have intermediate growth of the drift, positive entropy, etc. These groups were intensively studied recently (see papers [32, 17] and the bibliography in the latter).

On the other hand, metabelian groups are currently at the focus of very active research in geometric group theory. In 1983, Gromov proposed a program for studying finitely generated groups as geometric objects [27]. One of the principal directions of this program is the classification of finitely generated groups up to quasi-isometry. It follows from Gromov's result on groups with polynomial growth [26] that a group quasi-isometric to a nilpotent group is virtually nilpotent. In the case of solvable groups the situation is much less known. Erschler showed in [17] that a group quasi-isometric to a solvable group may be not virtually solvable. Thus, the class of virtually solvable groups is not closed under quasi-isometry. On the other hand there are interesting classes of solvable nonpolycyclic groups that are quasi-isometrically rigid, for example, solvable Baumslag-Solitar groups (Farb and Mosher [19, 20]). We refer to the papers [42] and [18] for some recent results in this area.

It seems timely to try to extend the results of this paper to the classes of solvable groups mentioned above. There are many interesting open questions concerning computational complexity of algorithmic problems in these classes of solvable groups; we discuss some of them in Section 5.

All polynomial time algorithms presented in this work are implemented and available at [11].

2. Preliminaries

2.1. **The Word Problem.** Let $F = F_r = F(X)$ be a free group with a basis $X = \{x_1, \dots, x_r\}$. A subset $R \subseteq F$ defines a *presentation* $P = \langle X \mid R \rangle$ of a group $G = F/N$, where $N = ncl(R)$ is the *normal closure* of R in F. If R is finite (recursively enumerable), then the presentation is called finite (recursively enumerable).

We say that the Word Problem WP for P is *decidable* if the normal closure N is a decidable subset of $F(X)$, i.e., if there exists an algorithm $\mathcal{A}$ to decide whether a given word $w \in F(X)$ belongs to N or not. The *time function* $T_{\mathcal{A}} : F(X) \to \mathbb{N}$ of the algorithm $\mathcal{A}$ is defined as the number of steps required for $\mathcal{A}$ to halt on an input $w \in F(X)$. We say that the Word Problem for P is decidable in polynomial time if there exists a decision algorithm $\mathcal{A}$, as above, and constants $c, k \in \mathbb{N}$ such that

$$T_{\mathcal{A}}(w) \leq c|w|^k$$

for every $w \in F(X)$ (here $|w|$ is the length of the word w). In this case we say that the time complexity of WP for P is $O(n^k)$.

2.2. **Free solvable groups and the Magnus embedding.** For a free group $F = F(X)$ of rank r, denote by $F^{(1)} = F' = [F, F]$ the *derived* subgroup of F, and by $F^{(d)} = [F^{(d-1)}, F^{(d-1)}]$ the *d-th derived subgroup* of F, $d \geq 2$. The quotient group $A_r = F_r/F_r'$ is a *free abelian group* of rank r, $M_r = F_r/F_r^{(2)}$ is a *free metabelian group* of rank r, and $S_{r,d} = F_r/F_r^{(d)}$ is a *free solvable group* of rank r and class d. In the sequel we usually identify the set X with its canonical images in A_r, M_r and $S_{r,d}$.

One of the most powerful approaches to the study of free solvable groups is via the Magnus embedding. To explain we need to introduce some notation. Let $G = F/N$ and let $\mathbb{Z}G$ be the group ring of G with integer coefficients. By $\mu : F \to G$ we denote the canonical factorization epimorphism, as well as its linear extension to $\mu : \mathbb{Z}F \to \mathbb{Z}G$. Let T be a free (left) $\mathbb{Z}G$-module of rank r with a basis $\{t_1, \ldots, t_r\}$. Then the set of matrices

$$M(G) = \begin{pmatrix} G & T \\ 0 & 1 \end{pmatrix} = \left\{ \begin{pmatrix} g & t \\ 0 & 1 \end{pmatrix} \mid g \in G, t \in T \right\}$$

forms a group with respect to the matrix multiplication. It is easy to see that the group $M(G)$ is a discrete wreath product $M(G) = A_r wr G$ of the free abelian group A_r and the group G.

Theorem 2.1 ([37]). *The homomorphism $\phi : F \to M(G)$ defined by*

$$x_i \overset{\phi}{\mapsto} \begin{pmatrix} x_i^{\mu} & t_i \\ 0 & 1 \end{pmatrix}, \quad i = 1, \ldots, r,$$

satisfies $\ker \phi = N'$. *Therefore, ϕ induces a monomorphism*

$$\phi : F/N' \hookrightarrow M(F/N).$$

The monomorphism ϕ is now called the *Magnus embedding*. The Magnus embedding allows one to solve WP in the group F/N' if WP in $G = F/N$ is decidable. Indeed, observe that

$$x_i^{-1} \overset{\phi}{\mapsto} \begin{pmatrix} (x_i^{-1})^{\mu} & (-x_i^{-1})^{\mu} t_i \\ 0 & 1 \end{pmatrix}, \quad i = 1, \ldots, r.$$

Now, if for $i = 1, \ldots, r$ and $\varepsilon = \pm 1$ we define the value

$$\delta(x_i^{\varepsilon}) = \begin{cases} 1, & \text{if } \varepsilon = 1; \\ (-x_i^{-1})^{\mu}, & \text{if } \varepsilon = -1; \end{cases}$$

then given a word $w = x_{i_1}^{\varepsilon_1} \ldots x_{i_n}^{\varepsilon_n} \in F(X)$ one can compute its image $\phi(w)$ in $M(G)$ as follows:

$$\begin{aligned} \phi(w) &= \phi(x_{i_1}^{\varepsilon_1}) \ldots \phi(x_{i_n}^{\varepsilon_n}) \\ &= \begin{pmatrix} \mu(x_{i_1}^{\varepsilon_1}) & \delta(x_{i_1}^{\varepsilon_1}) t_{i_1} \\ 0 & 1 \end{pmatrix} \cdot \ldots \cdot \begin{pmatrix} \mu(x_{i_n}^{\varepsilon_n}) & \delta(x_{i_n}^{\varepsilon_n}) t_{i_n} \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \mu(x_{i_1}^{\varepsilon_1} \ldots x_{i_n}^{\varepsilon_n}) & \displaystyle\sum_{j=1}^{n} (x_{i_1}^{\varepsilon_1} \ldots x_{i_{j-1}}^{\varepsilon_{j-1}})^{\mu} \delta(x_{i_j}^{\varepsilon_j}) t_{i_j} \\ 0 & 1 \end{pmatrix} \end{aligned}$$

and then, using a decision algorithm for WP in G, check if the resulting matrix $\phi(w)$ is the identity matrix or not. To estimate the complexity of such an algorithm, notice first that the coefficients from $\mathbb{Z}G$ that occur in the upper-right corner of the matrix $\phi(w)$ have $O(|w|)$ summands. Secondly, to check whether or not an element $h = m_1v_1 + \ldots + m_kv_k \in \mathbb{Z}G$, where $m_i \in \mathbb{Z}$ and $v_i \in G$ are given as words in the generators X from G, is trivial in $\mathbb{Z}G$ it requires $O(k^2)$ comparisons of the type $v_i = v_j$? in G. This gives an estimate for the time function T' of WP in F/N' via the time function T for WP in F/N:

$$T'(n) = O(rn^2T(n)),$$

where $n = |w|$. Since WP in A_r can be decided in linear time, the estimate above shows that the complexity of WP in M_r is $O(rn^3)$. Moreover, induction on the solvability class d gives a polynomial estimate $O(r^{d-1}n^{2d-1})$ for WP in the free solvable group $S_{r,d}$. Thus, the Magnus embedding gives a straightforward polynomial time (in r and n) decision algorithm for WP in $S_{r,d}$, but the degree of the polynomial grows with d. In particular, this algorithm is not polynomial as a uniform algorithm on the whole class of free solvable groups.

2.3. **Free Fox derivatives.** Let $F = F_r(X)$ be a free group of rank r with a basis $X = \{x_1, ..., x_r\}$. The trivial group homomorphism $F \to 1$ extends to a ring homomorphism $\varepsilon : \mathbb{Z}F \to \mathbb{Z}1 \simeq \mathbb{Z}$. The kernel of ε is called *the fundamental ideal* Δ_F of $\mathbb{Z}F$; it is a free (left) $\mathbb{Z}F$-module freely generated by elements $x_1 - 1, \ldots, x_r - 1$.

In [21, 22, 23, 24], R. Fox introduced and gave a thorough account of the free differential calculus in the group ring $\mathbb{Z}F$. Here we recall some notions and results, referring to the books [10, 6, 29] for details.

A map $D : \mathbb{Z}F \to \mathbb{Z}F$ is called a *derivation* if it satisfies the following conditions:

(D1) $D(u + v) = D(u) + D(v)$;
(D2) $D(uv) = D(u)v^{\varepsilon} + uD(v)$, where ε is the ring homomorphism defined above.

For every $x_i \in X$ there is a unique derivation, the so-called free partial derivative $\partial/\partial x_i$, such that $\partial x_j/\partial x_i = \delta_{ij}$, where δ_{ij} is the Kronecker delta. It turns out that for every $u \in \mathbb{Z}F$,

$$u - u^{\varepsilon} = \frac{\partial u}{\partial x_1}(x_1 - 1) + ... + \frac{\partial u}{\partial x_r}(x_r - 1). \tag{1}$$

Since Δ_F is a free $\mathbb{Z}F$-module, the equality (1) gives another definition of the partial derivatives.

Condition (D2) implies the following useful formulas, that allow one to compute easily partial derivatives of elements of $\mathbb{Z}F$:

$$\frac{\partial}{\partial x_i}(uv) = \frac{\partial}{\partial x_i}u + u\frac{\partial}{\partial x_i}v, \quad \text{for any } u, v \in F \tag{2}$$

and

$$\frac{\partial}{\partial x_i}(u^{-1}) = -u^{-1}\frac{\partial}{\partial x_i}u, \quad \text{for any } u \in F. \tag{3}$$

Therefore,

$$\frac{\partial}{\partial x_i}(x_j^{-1}) = -\delta_{i,j}x_j^{-1}, \tag{4}$$

and, hence, for a word $w = x_{i_1}^{\varepsilon_1}\dots x_{i_n}^{\varepsilon_n} \in F(X)$ one has

$$\begin{aligned}\frac{\partial w}{\partial x_i} &= \sum_{j=1}^{n} x_{i_1}^{\varepsilon_1}\dots x_{i_{j-1}}^{\varepsilon_{j-1}}(\partial x_{i_j}^{\varepsilon_j}/\partial x_i)\\ &= \sum_{1\le j\le n,\ i_j=i,\ \varepsilon_j=1} x_{i_1}^{\varepsilon_1}\dots x_{i_{j-1}}^{\varepsilon_{j-1}} - \sum_{1\le j\le n,\ i_j=i,\ \varepsilon_j=-1} x_{i_1}^{\varepsilon_1}\dots x_{i_j}^{\varepsilon_j}.\end{aligned} \tag{5}$$

The following result is one of the principal technical tools in this area; it follows easily from the Magnus embedding theorem, but in the current form it is due to Fox [21, 22, 23, 24].

Theorem (Fox)**.** *Let N be a normal subgroup of F and $\mu : F \to F/N$ the canonical epimorphism. Then for every $u \in F$ the following equivalence holds:*

$$\forall i\ \ (\partial u/\partial x_i)^{\mu} = 0 \Longleftrightarrow u \in [N, N].$$

In particular, for $N = F^{(d)}$ the standard epimorphism $\mu : F \to S_d = F/F^{(d)}$ gives rise to a ring homomorphism $\mu : \mathbf{Z}F \to \mathbf{Z}S_d$ such that

$$F^{(d+1)} = \{u \in F \mid (\partial u/\partial x_i)^{\mu} = 0 \ \textit{for}\ i = 1, \dots, r\}. \tag{6}$$

Composition of $\partial/\partial x_i$ with μ gives an *induced* partial derivative $\partial^{\mu}/\partial x_i : \mathbf{Z}F \to \mathbf{Z}S_d$, which we often denote again by $\partial/\partial x_i$ omitting μ (when it is clear from the context).

The partial derivatives $\partial^{\mu}/\partial x_i$ are useful when computing images under the Magnus embedding. Indeed, by induction on the length of $w \in F$ it is easy to show that the image of w under the Magnus embedding $\phi : F/N' \to M(F/N)$ can be written as follows:

$$w^{\phi} = \begin{pmatrix} w^{\mu} & \sum_{i=1}^{r} \partial^{\mu} w/\partial x_i \cdot t_i \\ 0 & 1 \end{pmatrix}.$$

This shows that the faithfulness of the Magnus embedding is, in fact, equivalent to the Fox theorem above.

2.4. **Flows on F/N.** In this section we relate flow networks on the Cayley graph of F/N to the elements of F/N'.

Let $X = \{x_1, \dots, x_r\}$ be a finite alphabet. An *X-labeled directed graph* Γ (or *X-digraph*) is a pair of sets (V, E), where the set V is called the *vertex set* and the set $E \subseteq V \times V \times X$ is called the *edge set*. An element $e = (v_1, v_2, x) \in E$ designates an edge with the *origin* v_1 (also denoted by $o(e)$), the *terminus* v_2 (also denoted by $t(e)$), labeled by x. If for $e \in E$ we have $o(e) = t(e)$, then we say that e is a *loop*. The graph Γ can be finite or infinite.

Example 2.2. The Cayley graph $\Gamma(G, X)$ of the group $G = F/N$ is an X-digraph.

Given an X-digraph Γ, we can make Γ into a directed graph labeled by the alphabet $X^{\pm 1} = X \cup X^{-1}$. Namely, for each edge $e = (v_1, v_2, x)$ of Γ we introduce a formal inverse $e^{-1} = (v_2, v_1, x^{-1})$. For the new edges e^{-1} we set $(e^{-1})^{-1} = e$. The new graph, endowed with this additional structure, is denoted by $\hat{\Gamma}$. In fact in many instances we abuse notation by disregarding the difference between Γ and $\hat{\Gamma}$.

Remark 2.3. If X is a generating set of G such that $X \cap X^{-1} = \emptyset$, then $\hat{\Gamma}(G, X)$ is the Cayley graph $\Gamma(G, X^{\pm 1})$ of G relative to the generating set $X^{\pm 1}$.

The edges of $\hat{\Gamma}$ inherited from Γ are called *positively oriented* or *positive*. The formal inverses of positive edges in $\hat{\Gamma}$ are called *negatively oriented* or *negative*. The edge set of $\hat{\Gamma}$ splits in a disjoint union $E(\hat{\Gamma}) = E^+(\Gamma) \sqcup E^-(\Gamma)$ of the sets of positive and negative edges.

The use of $\hat{\Gamma}$ allows us to define the notion of a *path* in Γ. Namely, a *path* p in Γ is a sequence of edges $p = e_1, \dots, e_k$ where each e_i is an edge of $\hat{\Gamma}$ and the origin of each e_i (for $i > 1$) is the terminus of e_{i-1}. In this situation we say that the *origin* $o(p)$ of p is $o(e_1)$ and the *terminus* $t(p)$ is $t(e_k)$. The *length* $|p|$ of this path is set to be k. Also, such a path p has a naturally defined label $\nu(p) = \nu(e_1) \dots \nu(e_k)$. Thus $\nu(p)$ is a word in the alphabet $\Sigma = X \cup X^{-1}$. Note that it is possible that $\nu(p)$ contains subwords of the form aa^{-1} or $a^{-1}a$ for some $a \in X$. If v is a vertex of Γ, we will consider the sequence $p = v$ to be a path with $o(p) = t(p) = v$, $|p| = 0$ and $\nu(p) = 1$ (the empty word).

In general, one can consider labels in an arbitrary inverse semigroup; the construction above applies to this case as well. In particular, we will consider directed graphs with labels in $\mathbb{Z}$. We also consider digraphs with no labels at all (to unify terminology, we view them sometimes as labeled in the trivial semigroup $\{1\}$); the construction above still applies.

Let $\Gamma = (V, E)$ be an X-digraph with two distinguished vertices s (called the *source*) and t (called the *sink*) from V. Recall that a *flow* (more precisely $\mathbb{Z}$-flow) on Γ is a function $f : E \to \mathbb{Z}$ such that

(F) for all $v \in V - \{s, t\}$ the equality $\sum_{o(e)=v} f(e) - \sum_{t(e)=v} f(e) = 0$ holds.

The number $f^*(v) = \sum_{o(e)=v} f(e) - \sum_{t(e)=v} f(e)$ is called the *net flow* at $v \in V$. The condition (F) is often referred to as the *Kirchhoff law* (see, for example, [7, 12]) or the *conservation law* [8].

For the digraph $\hat{\Gamma}$ the definition above can be formulated in the following equivalent way, which is the standard one in flow networks:

(F1) $f(e) = -f(e^{-1})$ for any $e \in E$;
(F2) $\sum_{o(e)=v} f(e) = 0$ for all $v \in V - \{s, t\}$.

Here the net flow at v is equal to $f^*(v) = \sum_{o(e)=v} f(e)$.

Usually a flow network comes equipped with a *capacity* function $c : E \to \mathbb{N}$, in which case a flow f has to satisfy the *capacity restriction*

(F3) $f(e) \leq c(e)$ for all $e \in E$.

In the sequel we do not make much use of the capacity function (it occurs in an obvious way), so in most cases we consider flows on graphs Γ satisfying the Kirchhoff law (F) (or, equivalently, on graphs $\hat{\Gamma}$ satisfying (F1) and (F2)).

A flow f is called a *circulation* if (F) holds for all vertices from V (including the source s and the sink t).

Example 2.4. Let $\Gamma = \Gamma(G, X)$ be the Cayley graph of $G = F/N$ relative to the generating set X. The constant function $f : E(\Gamma) \to \{1\}$ defines a circulation on Γ, since for every vertex $g \in V(\Gamma)$ and every label $x \in X$ there is precisely one edge (gx^{-1}, g) with label x incoming into g and precisely one edge (g, gx) with the label x leaving g.

An important class of flows on $\Gamma = \Gamma(G, X)$ comes from paths in Γ. A path p in Γ defines an integer-valued function $\pi_p : E(\Gamma) \to \mathbb{Z}$, such that for an edge e, $\pi(e)$ is the algebraic sum (with respect to the orientation) of the number of times the path p traverses e, i.e., each traversal of e in the positive direction adds $+1$, and in the negative direction adds -1. It is obvious that π_P is a flow on Γ with the source $o(p)$ (the initial vertex of p) and the sink $t(p)$ (the terminal vertex of p). Notice that π_p also satisfies the following conditions:

(F4) either π_p is a circulation (iff p is a closed path), or $f^*(s) = 1, f^*(t) = -1$;
(F5) π_p has finite support, i.e, the set $supp(\pi) = \{e \in E \mid \pi(e) \neq 0\}$ is finite.

We say that a flow π on Γ is *geometric* if it satisfies conditions (F4) and (F5).

It is easy to see that the set $\mathcal{C}(\Gamma)$ of all circulations on Γ forms an abelian group with respect to the operations (here $f, g \in \mathcal{C}(\Gamma)$):

- $(f + g)(e) = f(e) + g(e)$,
- $(-f)(e) = -f(e)$.

Meanwhile, the set $\mathcal{GC}(\Gamma)$ of all geometric circulations is a subgroup of $\mathcal{C}(\Gamma)$. On the other hand, the sum $f + g$ of two geometric flows gives a geometric flow only if the sink of f is equal to the source of g, or either f or g (or both) is a circulation. In fact, the set $\mathcal{GF}(\Gamma)$ of all geometric flows is a groupoid.

Let $\Pi(\Gamma)$ be the fundamental groupoid of paths in Γ. Then the map $\sigma : \Pi(\Gamma) \to \mathcal{GF}(\Gamma)$ defined for $p \in \Pi(\Gamma)$ by $\sigma(p) = \pi_p$ is a morphism in the category of groupoids; i.e., the following holds (here $p, q \in \Pi(\Gamma)$):

- $\pi_{pq} = \pi_p + \pi_q$, if pq is defined in $\Pi(\Gamma)$,
- $\pi_{p^{-1}} = -\pi_p$.

Now we will show that every geometric flow π on Γ can be realized as a path flow π_p for a suitable path p.

Lemma 2.5. *Let π be a geometric flow on Γ. Then there exists a path p in Γ such that $\pi = \pi_p$.*

Proof. Let π be a geometric flow on Γ with the source s and the sink t. Denote by Γ_π the subgraph of Γ generated by $supp(\pi) \cup \{s, t\}$. Suppose Q is a subgraph of Γ such that $\Delta = \Gamma_\pi \cup Q$ is a connected graph (every two vertices are connected by a path in $\hat{\Gamma}$). Clearly, π induces a flow on Δ. Now we construct another X-digraph Δ^* by adding new edges to Δ in the following manner. For every edge $e \in E(\Delta)$ with $|\pi(e)| > 1$ we add extra $|\pi(e)| - 1$ new edges $e^{(1)}, \dots, e^{|\pi(e)|-1}$ from $o(e)$ to $t(e)$ if $\pi(e) > 0$ and from $t(e)$ to $o(e)$ if $\pi(e) < 0$. We label the new edges by the same label if $\pi(e) > 0$, and by its inverse, otherwise. If $\pi(e) = 0$, then we add a new edge e^{-1} from $t(e)$ to $o(e)$ with the inverse label. In the case $|\pi(e)| = 1$ we do not add any new edges. Notice that every vertex in $V(\Delta^*) - \{s, t\}$ has even directed degree (the number of incoming edges is equal to the number of outgoing edges). There are two cases to consider.

Case 1. Suppose π is a circulation. Then every vertex in Δ^* has even directed degree. Therefore, the digraph Δ^* has an Euler tour p^*, i.e., a closed path at s that traverses every edge in Δ^* precisely once. Let $\phi : \Delta^* \to \hat{\Delta}$ be the morphism of X-digraphs that maps all the new edges $e^{(1)}, \dots, e^{|\pi(e)|-1}$ to their original edge e. Clearly, the image $p = \phi(p^*)$ is a path in $\hat{\Delta}$ such that $\pi_p = \pi$.

Case 2. Suppose π is not a circulation. Let q be a path in Δ^* from s to t. Then $\pi' = \pi - \pi_q$ is a circulation. Hence by Case 1 there exists a path p in Γ such that $\pi' = \pi_p$. Therefore, $\pi_{pq} = \pi_p + \pi_q = \pi' + \pi_q = \pi$, as required. □

2.5. **Geometric interpretation of Fox derivatives.** In this section we give a geometric interpretation of Fox derivatives.

Let $G = F/N$, $\mu : F \to F/N$ be the canonical epimorphism, and let $\Gamma = \Gamma(G, X)$ be the Cayley graph of G with respect to the generating set X. A word $w \in F(X)$ determines a unique path p_w in Γ labeled by w which starts at 1 (the vertex corresponding to the identity of G). As we mentioned in Section 2.4 the path p_w defines a geometric flow π_{p_w} on Γ, which we denote by π_w.

Lemma 2.6. *Let $w \in F = F(X)$. Then for any $g \in F/N$ and $x \in X$ the value of π_w on the edge $e = (g, gx)$ is equal to the coefficient in front of g in the Fox derivative $(\partial w/\partial x)^\mu \in \mathbb{Z}G$, i.e.,*

$$(\partial w/\partial x)^\mu = \sum_{g \in G, x \in X} \pi_w(g, gx) g.$$

Proof. The proof follows by induction on the length of w from formulas (5). □

Figure 1 is an example for $F = F(\{x_1, x_2\})$ and $N = F'$. The nonzero values of π_w are shown as weights on the edges (zero weights are omitted).

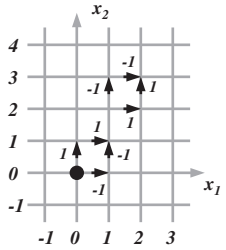

FIGURE 1. The values of π_w for $w = x_2x_1x_2x_1x_2x_1^{-1}x_2^{-3}x_1^{-1}$ on the (x_1, x_2)-grid. In this case $\partial w/\partial x_1 = -1 + x_2 - x_1x_2^3 + x_1x_2^2$ and $\partial w/\partial x_2 = 1 - x_1 + x_2^2x_2^2 - x_1x_2^2$.

The following theorem has been proven in [13, 54] using a homological argument similar to the one in Proposition 2.9. Here we give a short independent proof based on the Fox theorem.

Theorem 2.7 ([13, 54])**.** *Let N be a normal subgroup of F and $\sim_N$ an equivalence relation on F defined by*

$$u \sim_N v \iff \pi_u = \pi_v.$$

Then $F/N' = F/\sim_N$.

Proof. Let $u, v \in F$. Suppose $u = v$ in F/N'. Then $uv^{-1} \in N'$; hence by the Fox theorem $\partial^\mu(uv^{-1})/\partial x = 0$ for every $x \in X$ (here by $\partial^\mu/\partial x$ we denote the canonical image of $\partial/\partial x$ in the group ring $\mathbb{Z}(F/N)$). It follows from (3) and (2) that

$$\frac{\partial}{\partial x}(uv^{-1}) = \frac{\partial u}{\partial x} - uv^{-1}\frac{\partial v}{\partial x}. \tag{7}$$

Hence in $\mathbb{Z}(F/N)$,

$$0 = \frac{\partial^\mu}{\partial x}(uv^{-1}) = \frac{\partial^\mu u}{\partial x} - \frac{\partial^\mu v}{\partial x},$$

so, by Lemma 2.6 $\pi_u = \pi_v$, as claimed.

To show the converse, notice first that $\pi_u = \pi_v$ implies that $u = v$ in F/N. Indeed, it can be seen from the definition of π but also follows from (1) and Lemma 2.6 since in this case

$$(u^\mu - v^\mu) - (u - v)^\varepsilon = \sum_{x \in X} \frac{\partial^\mu}{\partial x}(u - v) \cdot (x - 1) = 0,$$

which implies $u^\mu = v^\mu$. Now by (7),

$$\frac{\partial^\mu}{\partial x}(uv^{-1}) = \frac{\partial^\mu u}{\partial x} - \frac{\partial^\mu v}{\partial x} = 0,$$

and, hence, by the Fox theorem, $uv^{-1} \in N'$. □

Remark 2.8. Theorem 2.7 relates the algebraic and geometric points of view on derivatives. One can prove this theorem (see Section 2.6) using a pure topological argument. Then the Fox theorem, as well as the Magnus embedding, come along as easy corollaries.

2.6. **Geometric circulations and the first homology group of Γ.** We describe here geometric circulations on $\Gamma = \Gamma(G, X)$ in a pure topological manner. For all required notions and results on homology of simplicial complexes we refer to [31] or [51].

In this section we view Γ as an infinite 1-complex.

Proposition 2.9. *Let $G = F/N$, $\Gamma = \Gamma(G, X)$, and $\sigma : \pi_1(\Gamma) \to \mathcal{GC}(\Gamma)$ be a map defined by $\sigma(p) = \pi_p$ for $p \in \pi_1(\Gamma)$. Then:*

- *σ is an epimorphism of groups;*
- *every geometric circulation on Γ defines a 1-cycle on Γ;*
- *$H_1(\Gamma, \mathbb{Z}) \simeq \mathcal{GC}(\Gamma)$;*
- *$\pi_1(\Gamma) \simeq N$ and $\ker \sigma = N'$.*

Proof. It was mentioned already in Section 2.4 that the map $p \to \pi_p$ is a morphism from the fundamental groupoid $\Pi(\Gamma)$ of paths in Γ into the groupoid of geometric flows $\mathcal{GF}(\Gamma)$. Hence, the restriction of this map onto the fundamental group $\pi_1(\Gamma)$ of Γ gives a homomorphism of groups. We have seen in Lemma 2.5 that σ is onto. This proves the first statement.

To see 2), observe first that a geometric circulation $f : E(\Gamma) \to \mathbb{Z}$, viewed as a formal sum $\sum_{e \in E(\Gamma)} f(e)e$, gives precisely a 1-chain in Γ (see, for example, [31]). Moreover, by definition, the net flow $f^*(v)$ at the vertex $v \in V(\Gamma)$ is the coefficient in front of v in the boundary ∂f of f. Therefore, $\partial f = 0$, so f is a 1-cycle.

3) follows easily from 2). Indeed, Γ is 1-complex, so there are no non-trivial 1-boundaries in Γ. In this event, $H_1(\Gamma, \mathbb{Z})$ is isomorphic to the group $\mathcal{GC}(\Gamma)$ of 1-cycles, as claimed.

4) It is a classical result that the kernel of $\sigma : \pi_1(\Gamma) \to H_1(\Gamma, \mathbb{Z})$ is equal to the derived subgroup of $\pi_1(\Gamma)$ (see [31]). To prove 4) it suffices to notice that $\pi_1(\Gamma) \simeq N$, which is easy. □

Remark 2.10. Proposition 2.9 gives a simple geometric proof of Theorem 2.7, that relates the algebraic and geometric points of view on derivatives. Now one can derive the Fox theorem from the geometric argument above and then obtain the description of the kernel of the Magnus embedding as an easy corollary.

2.7. **Geodesics in F/N'.** Let $G = F/N$ and $\mu : F(X) \to G$ be the canonical epimorphism. In this section we describe geodesics of elements of the group $H = F/N'$ relative to the set of generators X^μ.

It is convenient to view the free group $F = F(X)$ as the set of all freely reduced words in the alphabet $X^{\pm 1} = X \cup X^{-1}$ with the obvious multiplication.

To describe geodesics in H (relative to X) of a given word $w \in F$ we need a construction from Lemma 2.5. Recall that p_w is the path in the Cayley graph $\Gamma = \Gamma(G, X)$ from 1 to w^μ with the label w, and π_w is the induced geometric flow on Γ with the source 1 and the sink w^μ. By Γ_w we denote the subgraph of Γ generated by $supp(\pi_w) \cup \{1, w^\mu\}$. Suppose Q is a finite subgraph of Γ such that $\Delta = \Gamma_w \cup Q$ is a connected graph and Q has the least number of edges among all such subgraphs. It follows from minimality of Q that every connected component of Q is a tree. Moreover, if in the graph $\Delta = \Gamma_w \cup Q$ one collapses every connected component Γ_w to a point, then the resulting graph is a tree. We refer to Q as a *minimal forest* for w. In general, there could be several minimal forests for w.

Similarly as in the proof of Lemma 2.5, we construct a finite X-digraph Δ^* by replicating every edge $e \in E(\Delta)$ with $||\pi_w(e)| - 1|$ new edges in such a way that every vertex in $V(\Delta^*) - \{1, w^\mu\}$ has even directed degree, and the map that sends every replica of an edge e back to e (or e^{-1} depending on the orientation) is a morphism of X-labeled digraphs $\phi : \Delta^* \to \hat{\Delta}$. There are two cases.

Case I. Suppose that p_w is a closed path in Γ, i.e., $w \in N$. In this case every vertex in Δ^* has even degree, so Δ^* has an Euler tour, say p^*_Q at 1. Denote by p_Q the image $\phi(p^*_Q)$ of p^*_Q under ϕ. It follows from the construction (see Lemma 2.5) that p_Q is a closed path at 1 in Γ such that $\pi_w = \pi_{p_Q}$. Therefore, if $w_{p_Q} \in F$ is the label of p_Q, then by Theorem 2.7, $w = w_{p_Q}$ in H. Moreover, since p_Q is an Euler tour in Δ^*, its length, hence the length of w_{p_Q}, is equal to

$$|p_Q| = \sum_{e \in supp(p_w)} |\pi_w(e)| + 2|E(Q)|. \tag{8}$$

Case II. Suppose that p_w is not a closed path in Γ, i.e., $w \notin N$. By induction on $|w|$ it is easy to show that the vertices 1 and w^μ belong to the same connected component of Γ_w. Again, there exists an Euler tour p^*_Q in the graph Δ^* which starts at the source and ends at the sink. Clearly, π_{p_Q} satisfies the equality (8). If u is the label of the path π_{p_Q}, then $\pi_u = \pi_{p_Q} = \pi_w$ and u is a geodesic word for w.

Now, with the construction in place, we are ready to characterize geodesics in H of elements from N.

Theorem 2.11. *Let $H = F/N'$ and $w \in F$. Then the following hold:*

- *if Q is a minimal forest for w, then w_{p_Q} is a geodesic for w and*

$$l_X(w) = \sum_{e \in supp(p_w)} \pi_w(e) + 2|E(Q)|;$$

- *every geodesic word for w is equal (in F) to a word w_{p_Q} for a suitable minimal forest Q and an Euler path p^*_Q.*

Proof. Let $u \in F$ be a geodesic word for w^μ in H. Observe that $\Delta = supp(\pi_u) \cup p_u$ is a connected subgraph of Γ and

$$|p_u| \geq \sum_{e \in supp(p_u)} \pi_u(e) + 2|E(\Delta - supp(\pi_u))|.$$

Now, by Theorem 2.7, the equality $u = w$ in H implies $\pi_u = \pi_w$. In particular, $supp(\pi_u) = supp(\pi_w)$. Hence

$$|p_u| \geq \sum_{e \in supp(p_w)} \pi_w(e) + 2|E(\Delta - supp(\pi_w))|. \tag{9}$$

Since u is geodesic for w the number $|\Delta - supp(\pi_w)|$ is the minimal possible, so $Q = \Delta - supp(\pi_w)$ is a minimal forest for w. In fact, the equation (9) shows that the converse is also true. This proves the theorem. □

The discussion above shows that GP is easy in F/N' provided one can solve the following problem efficiently.

Minimal Forest Problem (MFP): Given a finite set of connected finite subgraphs $\Gamma_1, \ldots, \Gamma_s$ in Γ find a finite subgraph Q of Γ such that $\Gamma_1 \cup \ldots \cup \Gamma_s \cup Q$ is connected and Q has a minimal possible number of edges.

Proposition 2.12. *GP in F/N' (relative to X) is linear time reducible to MFP for $\Gamma(F/N, X)$.*

Proof. This follows from the discussion above. Indeed, given a word $w \in F$ one can in linear time compute the flow π_w and find the connected components $\Gamma_1, \ldots, \Gamma_s$ of $supp(\pi_w) \cup \{o(p_w), t(p_w)\}$ in $\Gamma = \Gamma(F/N, X)$. Then, solving MFP for these components, one gets the subgraph Q which makes the graph $\Gamma_1 \cup \ldots \cup \Gamma_s \cup Q$ connected. Obviously, it takes linear time to find an Euler path in the graph Δ, hence to find a geodesic for w. □

3. The Word Problem in free solvable groups

In this section we present fast algorithms to compute Fox derivatives of elements of a free group F in the group ring $\mathbb{Z}S_{r,d}$ of a free solvable group $S_{r,d}$. As an immediate application, we obtain a decision algorithm for WP in a free metabelian group M_r with time complexity $O(rn \log_2 n)$ and a decision algorithm for WP in $S_{r,d}, d \geq 3$, with time complexity $O(rdn^3)$. These are significant improvements in comparison with the known decision algorithms discussed in the Introduction and Section 2.2. As another application we get a fast algorithm to compute images of elements from $S_{r,d}$ under the Magnus embedding, which opens up an opportunity to efficiently use the classical techniques developed for wreath products.

3.1. **The Word Problem in free metabelian groups.** In this section we compute Fox derivatives of elements of F in the group ring $\mathbb{Z}(F/F')$. Then we apply this to WP in free metabelian groups.

Let $X = \{x_1, \dots, x_r\}$, $F = F_r = F(X)$, $M = M_r = F/F^{(2)}$, $A = A_r = F/F'$, and $\mu : F \to A$ be the canonical epimorphism. All Fox derivatives in this section are computed in the ring $\mathbb{Z}A$.

Let $w \in F$. Then

$$\frac{\partial^\mu w}{\partial x_i} = \sum_{a \in A} m_{a,i} a, \quad m_{a,i} \in \mathbb{Z}.$$

One can encode all the derivatives $\partial^\mu w/\partial x_i$ in one mapping $M_w : A \times \{1, \dots, r\} \to \mathbb{Z}$, where $M_w(a, i) = m_{a,i}$. Let

$$supp(M_w) = \{(a, i) \mid M_w(a, i) \neq 0\},$$

and let S_w be the restriction of M_w onto $supp(M_w)$. To compute Fox derivatives of w we construct a sequence of finite maps $S_0 = \emptyset, S_1, \dots, S_n = S_w$, as we read w. On each step k we either extend the domain $Dom(S_k)$ of S_k or change the value of S_k on some element from $Dom(S_k)$. To do this we need a data structure which allows one to do the following operations efficiently:

- for a given (a, i) determine if $(a, i) \in Dom(S_k)$ or not;
- add (a, i) to $Dom(S_k)$ if $(a, i) \notin Dom(S_k)$ and define $S_k(a, i) = q$ for some $q \in \mathbb{Z}$;
- change the value of S_k on (a, i) if $(a, i) \in Dom(S_k)$.

Every element $a \in A$ can be written uniquely in the form $a = x_1^{\delta_1(a)} \dots x_r^{\delta_r(a)}$, where $\delta_i(a) \in \mathbb{Z}$, so one may use the r-tuple of coordinates $\delta(a) = (\delta_1(a), \dots, \delta_r(a))$ to represent a. It follows from the formula (5) for partial derivatives that for every $(a, i) \in Dom(S_w)$ the components $\delta_j(a)$ of $\delta(a)$ satisfy the inequality $|\delta_j(a)| \leq |w|$, as well as the values of S_w and, hence, $|S_w(a, i)| \leq |w|$. Therefore, it takes $\lceil \log_2(|w|+1) \rceil$ bits to encode one coordinate $\delta_j(a)$ (one extra bit to encode the sign $+$ or $-$), $r\lceil \log_2(|w|+1) \rceil$ bits to encode $\delta(a)$, and at most $r\lceil \log_2(|w|+1) \rceil + \lceil \log_2 r \rceil$ bits to encode (a, i). We denote the binary word encoding (a, i) by $(a, i)^*$.

Thus every function S_k can be uniquely represented by a directed $\{0, 1\}$-labeled binary tree T_k with the root ε and with leaves labeled by integers such that $(a, i) \in Dom(S_k)$ if and only if there exists a path in T_k from the root ε to a leaf labeled by the code of (a, i) and such that the leaf of this path is labeled precisely by the integer $S_k(a, i)$. Notice that the height of T_k is equal to $r\lceil \log_2(|w| + 1) \rceil + \lceil \log_2 r \rceil$. Such a tree is visualized schematically in Figure 2.

Remark 3.1. It is clear that one can perform every operation mentioned above on this data structure in at most $r\lceil \log_2(|w| + 1) \rceil + \lceil \log_2 r \rceil$ elementary steps.

We use this data structure to design the following algorithm for computing S_w.

Algorithm 3.2 (Computing Fox derivatives in F/F').
INPUT. $r \in \mathbb{N}$ *and* $w = x_{i_1}^{\varepsilon_1} \dots x_{i_n}^{\varepsilon_n} \in F(X)$, *where* $i_j \in \{1, \dots, r\}$ *and* $\varepsilon_j = \pm 1$.

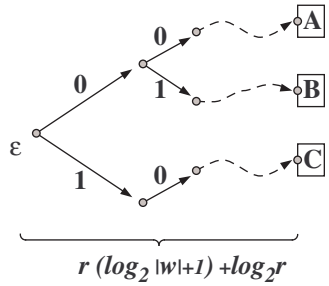

FIGURE 2. The tree T_k representing the function S_k.

OUTPUT. S_w.
COMPUTATIONS.

A. *Set* $S = \emptyset$ *and* $\delta(a) = (0, \dots, 0) \in \mathbb{Z}^r$.
B. *For* $j = 1, \dots, n$ *do:*
   (1) *if* $\varepsilon_j = 1$, *then*
      * *check if there is a path (from the root to a leaf) labeled by* $(\delta(a), i_j)^*$ *in* S;
      * *if such a path does not exist in* S, *then create it, add it to* S, *and put* 1 *as the corresponding value at the new leaf;*
      * *if such a path exists in* S, *then add* 1 *to the value at its leaf;*
      * *add* ε_j *to the* i_j*th coordinate of* $\delta(a)$;
   (2) *if* $\varepsilon_j = -1$, *then*
      * *add* ε_j *to the* i_j*th coordinate of* $\delta(a)$;
      * *check if there is a path (from the root to a leaf) labeled by* $(\delta(a), i_j)^*$ *in* S;
      * *if such a path does not exist in* S, *then create it, add it to* S, *and put* -1 *at the corresponding leaf;*
      * *if such a path exists in* S, *then subtract* 1 *from the value at its leaf.*
C. *Output* S_n.

Theorem 3.3. *Given* $r \in \mathbb{N}$ *and* $w \in F$, *Algorithm* 3.2 *computes all partial derivatives of* w *in* $\mathbb{Z}A$ *(the mapping* S_w*) in time* $O(r|w| \log_2 |w|)$.

Proof. Using the formula (5) for partial derivatives it is easy to check that given $w \in F$, Algorithm 3.2, indeed, computes the mapping S_w. To verify the complexity estimates observe first that Algorithm 3.2 performs exactly $|w|$ iterations at step $[B]$. Each iteration requires $O(r\lceil \log_2(|w|+1)\rceil)$ elementary steps (see Remark 3.1), so altogether one has $O(r|w|log_2|w|)$ as the time complexity estimate for Algorithm 3.2, as claimed. □

Algorithm 3.4 (Word Problem in free metabelian groups)**.**
INPUT. $r \in \mathbb{N}$ *and* $w \in F$.
OUTPUT. *True if* w *represents the identity in* M_r *and False otherwise.*

COMPUTATIONS.

- *Apply Algorithm* 3.2 *to compute* S_w.
- *Check, looking at the values assigned to leaves of* S_w*, if all Fox derivatives* $\partial w/\partial x_i$, $i = 1, \dots, r$ *are equal to* 0 *or not.*
- *If all the derivatives are* 0*, then output* $True$*. If there is a nonzero derivative, then output* $False$.

Theorem 3.5. *Algorithm* 3.4 *solves the Word Problem in a free metabelian group* M_r *in time* $O(r|w| \log_2 |w|)$.

Proof. This follows from Theorem 3.3 and the Fox theorem (see Section 2.3). □

3.2. **The Word Problem in free solvable groups.** In this section we present an algorithm to compute all Fox derivatives of a word $w \in F$ in the group ring $\mathbb{Z}S_{r,d-1}$, $d \geq 2$ in time $O(rd|w|^3)$. This gives a decision algorithm for WP in $S_{r,d}$ within time complexity $O(rd|n|^3)$.

Let $X = \{x_1, \dots, x_r\}$, $F = F_r = F(X)$, $S = S_{r,d} = F/F^{(d)}$, $d \geq 3$, and $\mu : F \to S_{r,d-1}$ be the canonical epimorphism. All Fox derivatives in this section are computed in the ring $\mathbb{Z}S_{r,d-1}$.

In Section 3.1 we used a unique representation of elements $a \in A_r$ by their coordinate vectors $\delta(a)$ to compute Fox derivatives in nearly linear time. Now we do not have such normal forms of elements of $S_{r,d}$, $d \geq 2$, so our computations are slightly different; however, the general strategy is quite similar. To speed up computations we use some data structures based on efficient partitioning techniques.

In general, let G be a group generated by X and let D be a finite subset of $F(X)$. A *G-partition* of D is a partition of D into a union of disjoint nonempty subsets D_i such that for any $u, v \in D$, $u = v$ in G if and only if they belong to some subset D_i. Clearly, the G-partition of D is unique. Observe that if a group H is a quotient of G, then the G-partition of D is the same or *finer* than the H-partition of D.

If the set D is ordered, say $D = \{w_0, \dots, w_n\}$, then the G-partition of D can be represented by a function $P : \{0, \dots, n\} \to \{0, \dots, n\}$ where $P(j) = i$ if and only if $w_i = w_j$ in G and i is the smallest with such a property. Given the G-partition P of D one can arrange this data in such a way that it takes linear time (in the size of j) to compute $P(j)$. In particular, given i, j one can check in linear time if $w_i = w_j$ in G. Also, for a given word $w \in D$ one can find in linear time an index i such that $w = w_i$. These are the two main subroutines concerning partitions of D.

Let $w = x_{i_1}^{\varepsilon_1} \dots x_{i_n}^{\varepsilon_n}$, where $i_j \in \{1, \dots, r\}$ and $\varepsilon_j = \pm 1$. Put

$$D_w = \{\varepsilon,\ x_{i_1}^{\varepsilon_1},\ x_{i_1}^{\varepsilon_1} x_{i_2}^{\varepsilon_2},\ \dots,\ x_{i_1}^{\varepsilon_1} \dots x_{i_n}^{\varepsilon_n}\} \subset F(X_r). \tag{10}$$

We order the set D_w as follows: $w_0 = \varepsilon, \dots, w_n = w$. Now, to check whether or not the derivative $\partial w/\partial x_i$ is trivial in $\mathbb{Z}S_{r,d-1}$ one has to determine which pairs (w_i, w_j) of elements from D_w represent the same element in $S_{r,d-1}$ and then cancel out w_i with w_j in $\partial w/\partial x_i$ if they have opposite signs.

The goal of Algorithm 3.10 below is to compute the $S_{r,d}$-partition for D_w. This is performed in a sequence of iterations. The algorithm starts out by computing the A_r-partition of D_w. On the second iteration the algorithm computes the M_r-partition of D_w. On the third step it computes the $S_{r,3}$-partition of D_w. It continues this way until it computes the $S_{r,d}$-partition of D_w.

To explain how the algorithm works, assume that the $S_{r,d-1}$-partition of D_w is given by the partition function P_{d-1} described above. Notice that the $S_{r,d}$-partition

of D_w is the same or finer than the $S_{r,d-1}$-partition of D_w, since $S_{r,d-1}$ is a quotient of $S_{r,d}$. This shows that to construct the $S_{r,d}$-partition P_d of D_w one has only to compare elements from D_w which are equal in $S_{r,d-1}$. Suppose that $w_s, w_t \in D_w$, $s < t$ and $w_s = w_t$ in $S_{r,d-1}$. To check if $w_s = w_t$ in $S_{r,d}$ we compare all their Fox derivatives in $\mathbb{Z}S_{r,d-1}$, so for every $k = 1, \dots, r$ we compute the following differences:

$$\begin{aligned}
&\partial w_s/\partial x_k - \partial w_t/\partial x_k \\
&= \sum_{1\le j\le s,\ i_j=k,\ \varepsilon_j=1} x_{i_1}^{\varepsilon_1}\dots x_{i_{j-1}}^{\varepsilon_{j-1}} - \sum_{1\le j\le s,\ i_j=k,\ \varepsilon_j=-1} x_{i_1}^{\varepsilon_1}\dots x_{i_j}^{\varepsilon_j} \\
&\quad - \sum_{1\le j\le t,\ i_j=k,\ \varepsilon_j=1} x_{i_1}^{\varepsilon_1}\dots x_{i_{j-1}}^{\varepsilon_{j-1}} + \sum_{1\le j\le t,\ i_j=k,\ \varepsilon_j=-1} x_{i_1}^{\varepsilon_1}\dots x_{i_j}^{\varepsilon_j} \\
&= - \sum_{s+1\le j\le t,\ i_j=k,\ \varepsilon_j=1} x_{i_1}^{\varepsilon_1}\dots x_{i_{j-1}}^{\varepsilon_{j-1}} + \sum_{s+1\le j\le t,\ i_j=k,\ \varepsilon_j=-1} x_{i_1}^{\varepsilon_1}\dots x_{i_j}^{\varepsilon_j} \\
&= - \sum_{s+1\le j\le t,\ i_j=k,\ \varepsilon_j=1} w_{j-1} + \sum_{s+1\le j\le t,\ i_j=k,\ \varepsilon_j=-1} w_j. \qquad (11)
\end{aligned}$$

Clearly, given w and s, as above one can compute the formal expression (11) in time $O(|w|)$. To check if (11), viewed as an element in $\mathbb{Z}S_{r,d-1}$, is equal to 0 it suffices to represent it in the standard group ring form $\sum_{g\in S_{r,d-1}} mg$ (where $m \in \mathbb{Z}$) and verify if all coefficients in this representation are zeros. Now we describe a particular procedure, termed the Collecting Similar Terms Algorithm, which gives the standard group ring form for (11). Given (11) one can compute in time $O(|w|)$ the following sum:

$$- \sum_{s+1\le j\le t,\ i_j=k,\ \varepsilon_j=1} w_{P(j-1)} + \sum_{s+1\le j\le t,\ i_j=k,\ \varepsilon_j=-1} w_{P(j)}. \qquad (12)$$

Observe now that two summands w_p and w_q in (12) are equal in $S_{r,d-1}$ if and only if $p = q$. It is easy to see that it takes time $O(|w|)$ to collect similar terms in (12), i.e., to compute the coefficients in the standard group ring presentation of (12).

It follows that $\partial w_s/\partial x_k = \partial w_t/\partial x_k$ in $\mathbb{Z}S_{r,d-1}$ if and only if all the coefficients in the standard group ring form of (12) are equal to 0. The argument above shows that one can check whether or not $\partial w_s/\partial x_k = \partial w_t/\partial x_k$ in $\mathbb{Z}S_{r,d-1}$ in time $O(|w|)$. Since we need to compare all partial derivatives $\partial/\partial x_k, k = 1, \dots, r$, of the elements w_s and w_t, it takes altogether $O(r|w|)$ time to verify if $w_s = w_t$ in $S_{r,d-1}$.

The routine above allows one to construct effectively the $S_{r,d}$-partition P_d of D_w when given the $S_{r,d-1}$-partition P_{d-1}. A more formal description of the algorithm is given below.

Algorithm 3.6 (Computing the $S_{r,d}$-partition of D_w)**.**
INPUT. *Positive integers $r \ge 2$, $d \ge 2$, and a word $w \in F(X)$.*
OUTPUT. *The $S_{r,d}$-partition function P_d for the set D_w.*
INITIALIZATION. *Compute the set D_w and form the initial (trivial) F-partition P_0 of D_w, so $P_0(i) = i$ for every $i \in \{0, \dots, n\}$.*
COMPUTATIONS.

A. *Compute the A-partition P_1 of D_w.*
B. *For $c = 2, \dots, d$ do:*

1) *For each* $0 \leq s < t \leq n$ *such that* $w_s = w_t$ *in* $S_{r,c-1}$ *check whether or not* $w_s = w_t$ *in* $S_{r,c}$.
2) *Form the* $S_{r,c}$*-partition* P_d *of* D_w.

C. *Output* P_d.

Lemma 3.7. *Given integers* $r, d \geq 2$ *and* $w \in F$, *Algorithm* 3.6 *computes the* $S_{r,d}$*-partition (the function* P_d*) of* D_w *in time* $O(dr|w|^3)$.

Proof. Algorithm 3.6 makes precisely d iterations $c = 1, \ldots, d$ by consequently computing the $S_{r,c}$-partitions of D_w. After the $S_{r,c-1}$-partition of D_w is computed, the algorithm computes the $S_{r,c}$-partition of D_w by comparing elements $w_s, w_t \in D_w$ in $S_{r,c}$. It requires at most $|w|(|w|+1)/2$ such checks, and, as was explained above, every such check can be done in time $O(r|w|)$. Altogether one needs $O(r|w|^3)$ steps to construct the function P_c on the iteration c. Since the algorithm makes altogether d iterations it takes it $O(dr|w|^3)$ time to produce P_d, as claimed. □

Now we are in a position to show two applications of Algorithm 3.6. The first one is concerned with computing Fox derivatives in $\mathbb{Z}S_{r,d}$.

Algorithm 3.8 (Computing Fox derivatives)**.**
INPUT. *Positive integers* $r \geq 2$, $d \geq 2$, *a word* $w \in F(X)$, *and a number* $k \in \{1, \ldots, r\}$.
OUTPUT. *The standard group ring presentation of the Fox derivative* $\partial w/\partial x_k$ *in* $\mathbb{Z}S_{r,d}$.
COMPUTATIONS.

A. *Compute, using formals* (5), *the Fox derivative* $\partial w/\partial x_k$ *in* $\mathbb{Z}F$.
B. *Compute, using Algorithm* 3.6, *the* $S_{r,d}$*-partition of* D_w.
C. *Compute, using the Collecting Similar Terms Algorithm, the standard group ring form of* $\partial w/\partial x_k$ *in* $\mathbb{Z}S_{r,d}$
D. *Output* $\partial w/\partial x_k$ *computed in [C].*

Lemma 3.9. *Given integers* $r, d \geq 2$, *a word* $w \in F$, *and a number* $k \in \{1, \ldots, r\}$, *Algorithm* 3.8 *computes the standard group ring presentation of the Fox derivative* $\partial w/\partial x_k$ *in* $\mathbb{Z}S_{r,d}$ *in time* $O(dr|w|^3)$.

Proof. The proof follows from Lemma 3.7. □

The second application of Algorithm 3.6 is to WP in $S_{r,d}$.

Algorithm 3.10 (WP in $S_{r,d}$)**.**
INPUT. *Positive integers* $r \geq 2$, $d \geq 2$, *and a word* $w \in F(X)$.
OUTPUT. *True if* $w = 1$ *in* $S_{r,d}$ *and False otherwise.*
COMPUTATIONS.

A. *Compute the set* D_w.
B. *Using Algorithm* 3.6, *compute the* $S_{r,d}$*-partition* P_d *of* D_w.
C. *If* $P_d(0) = P_d(n)$, *i.e.,* $1 = w$ *in* $S_{r,d}$, *then output True. Otherwise output False.*

Theorem 3.11. *Algorithm* 3.10 *solves the Word Problem in a free solvable group* $S_{r,d}$ *in time* $O(dr|w|^3)$.

Proof. This follows immediately from Lemma 3.7. □

4. Geodesics in free metabelian groups

In this section we discuss the computational hardness of different variations of geodesic problems and prove the main result about NP-completeness of BGLP in free metabelian groups.

4.1. **Algorithmic problems with geodesics in groups.** Let G be a group with a finite set of generators $X = \{x_1, \dots, x_r\}$ and let $\mu : F(X) \to G$ be the canonical epimorphism. In this section we view the free group $F(X)$ as the set of all freely reduced words in the alphabet $X^{\pm 1} = X \cup X^{-1}$ with the obvious multiplication.

For a word w in the alphabet $X^{\pm 1}$ by $|w|$ we denote the length of w. The *geodesic length* of an element $g \in G$ relative to X, denoted by $l_X(g)$, is the length of a shortest word $w \in F(X)$ representing g, i.e.,

$$l_X(g) = \min\{|w| \mid w \in F(X), w^\mu = g\}.$$

To simplify notation we write, sometimes, $l_X(w)$ instead of $l_X(w^\mu)$. A word $w \in F(X)$ is called *geodesic* in G relative to X if $|w| = l_X(w)$.

We are interested here in the following algorithmic *search* problem in a given group G described as above.

The Geodesic Problem (GP): Given a word $w \in F(X)$, find a geodesic (in G) word $\tilde{w} \in F(X)$ such that $w^\mu = \tilde{w}^\mu$.

One can consider the following variation of GP.

The Geodesic Length Problem (GLP): Given a word $w \in F(X)$, find $l_X(w)$.

Though GLP seems easier than GP (since a solution to GP gives, in linear time, a solution to GLP), in practice, to solve GP one usually solves GP first, and only then computes the geodesic length.

As is customary in complexity theory, one can modify the search problem GLP to get the corresponding bounded decision problem:

The Bounded Geodesic Length Problem (BGLP): Let G be a group with a finite generating set X. Given a word $w \in F(X)$ and a natural number k, determine if $l_X(w) \le k$.

It is instructive to compare the algorithmic "hardness" of the problems above and the Word Problem (WP). Clearly, if one of them is decidable, then all of them are decidable. To see the difference we need to recall a few definitions. Let A and B be algorithmic problems with input sets I_A and I_B. Then A is termed *Turing reducible* to B in polynomial time if there exists an algorithm $\mathcal{A}$ with an oracle for B (which can be viewed as a "subroutine" of $\mathcal{A}$ that for a given input $e \in I_B$ in one step returns the answer for B on e) that solves A in polynomial time. Similarly, one can define Turing reducibility in *exponential* time. In these cases we write $A \preceq_{T,p} B$ or, correspondingly, $A \preceq_{T,exp} B$.

Again, it is not hard to see that $WP \preceq_{T,p} BGLP \preceq_{T,p} GLP \preceq_{T,p} GP$. Moreover, since (by brute force algorithm) $GP \preceq_{T,exp} WP$, it follows that all these problems are Turing reducible to each other in exponential time. Moreover, if G has polynomial *growth*, i.e., if there is a polynomial $p(n)$ such that for each $n \in \mathbb{N}$ the cardinality of the ball B_n of radius n in the Cayley graph $\Gamma(G, X)$ is at most $p(n)$, then one can easily construct this ball B_n in polynomial time with an oracle for the WP in G (see, for example, [13] for details). It follows that if a group with

polynomial growth has WP decidable in polynomial time, then all the problems above have polynomial time complexity with respect to any finite generating set (since the growth and WP stay polynomial for any finite set of generators). Observe now, that by Gromov's theorem [26], finitely generated groups of polynomial growth are virtually nilpotent. It is also known that the latter have WP decidable in polynomial time (nilpotent finitely generated groups are linear). These two facts together imply that the Geodesic Problem is polynomial time decidable in finitely generated virtually nilpotent groups.

On the other hand, there are many groups of exponential growth where GP is decidable in polynomial time, for example, hyperbolic groups [16]. Among metabelian groups, the Baumslag-Solitar group $BS(1,2) = \langle a, t \mid t^{-1}at = a^2 \rangle$ has exponential growth (it is solvable but not polycyclic, and the claim follows from the Milnor theorem [41]) and GP in $BS(1,2)$ is decidable in polynomial time (see [15]).

In general, if WP in G is polynomially decidable, then BGLP is in the class NP, i.e., it is decidable in polynomial time by a nondeterministic Turing machine. Indeed, if $l_X(w) \leq k$, then there is a word $u \in F(X)$ of length at most k which is equal to w in G; this u is a "witness" of polynomial size which allows one to verify in polynomial time that $l_X(w) \leq k$ (just checking that $u = w$ in G). In this case GLP is Turing reducible in polynomial time to an NP problem, but we cannot claim the same for GP. Observe that BGLP is in NP for any finitely generated metabelian group, since they have WP decidable in polynomial time (see the Introduction).

It might happen though that WP in a group G is polynomial time decidable, but BGLP in G is as hard as any problem in the class NP, i.e., it is NP-complete. Recall (in the notation above) that a decision problem B is NP-complete if it is in NP and for any decision problem A from NP there is a computable in polynomial time function $f : I_A \to I_B$ (Karp reduction, or a polynomial reduction), such that A is true on $x \in I_A$ if and only if B is true on $f(x)$. The simplest example of this type is due to Parry, who showed in [45] that BGLP is NP-complete in the metabelian group $\mathbb{Z}_2 wr (\mathbb{Z} \times \mathbb{Z})$ (the wreath product of $\mathbb{Z}_2$ and $\mathbb{Z} \times \mathbb{Z}$). In this event, the search problems GP and GLP are called *NP-hard*; this means precisely that some (any) NP-complete problem is Turing reducible to them in polynomial time.

It would be very interesting to classify finitely generated metabelian groups with respect to computational hardness of their GP or GLP problems. In the next section we clarify the situation with free metabelian groups. Some remaining open problems are discussed in Section 5.

It was claimed in [13] that in free solvable groups of finite rank, GLP is decidable in polynomial time. Unfortunately, in this particular case their argument is fallacious. Our main result of this section is the following theorem.

Theorem 4.1 (Main Theorem)**.** *Let M_r be a free metabelian group M_r of finite rank $r \geq 2$. Then BGLP in M_r (relative to the standard basis) is NP-complete.*

Proof. The proof of this result consists of two parts. Firstly, in Section 4.2 (Corollary 4.5) we show that it suffices to prove that BGLP is NP-complete in M_2. Secondly, in Section 4.4 (Theorem 4.11) we give a proof that BGLP is, indeed, NP-complete in M_2. □

This immediately implies the following results.

Corollary 4.2. *The search problems GP and GLP are NP-hard in nonabelian M_r (relative to the standard basis).*

To prove the Main Theorem we reduce the problem to the case $r = 2$ and then show that BGLP in M_2 is NP-complete. To see the latter we construct a polynomial reduction of the Rectilinear Steiner Tree Problem to BGLP in M_r.

4.2. **Reduction to M_2.** Let $\mathcal{V}$ be a variety of groups. For groups $A, B \in \mathcal{V}$ we denote by $A *_{\mathcal{V}} B$ the free product of A and B relative to $\mathcal{V}$. In particular, if $A = \langle X \mid R \rangle$ and $B = \langle Y \mid S \rangle$ are presentations of A and B in $\mathcal{V}$, then $A *_{\mathcal{V}} B = \langle X \cup Y \mid R \cup S \rangle$ is a presentation of $A *_{\mathcal{V}} B$ in $\mathcal{V}$. As usual, $A *_{\mathcal{V}} B$ satisfies the canonical universal property: any two homomorphisms $A \to C, B \to C$ into a group $C \in \mathcal{V}$ extend to a unique homomorphism $A *_{\mathcal{V}} B \to C$ (we refer to [44] for details). It follows that if $F_{\mathcal{V}}(X \cup Y)$ is a free group in $\mathcal{V}$ with basis $X \cup Y$, then $F_{\mathcal{V}}(X \cup Y) = F_{\mathcal{V}}(X) *_{\mathcal{V}} F_{\mathcal{V}}(Y)$.

The following lemma claims that free $\mathcal{V}$-factors of a group G are isometrically embedded into G.

Lemma 4.3. *Let $A, B \in \mathcal{V}$ with finite generating sets X and Y. Then in the group $A *_{\mathcal{V}} B$ no geodesic word (relative to $X \cup Y$) for an element from A contains a letter from Y. In particular, for any word $w \in F(X)$ its geodesic length in A (relative to X) is equal to the geodesic length in $A *_{\mathcal{V}} B$ (relative to $X \cup Y$).*

Proof. Let $w \in F(X)$ be a geodesic word in A relative to X. Suppose that $u \in F(X \cup Y)$ is a geodesic word in $G = A *_{\mathcal{V}} B$ (relative to $X \cup Y$) that defines the same element as w. The identical map $A \to A$ and the trivial map $B \to 1$ give rise to a homomorphism $\phi : G \to A$. This ϕ, when applied to u, just "erases" all letters from Y. It follows that if u contains a letter from Y, then $|u^\phi| < |u| \le |w|$, a contradiction with the assumption that w is geodesic in A relative to X (since $w = u^\phi$ in A). □

Corollary 4.4. *In the notation above, each of the problems GP, GLP, BGLP in A (relative to X) is polynomial time reducible to the problem of the same type in $A *_{\mathcal{V}} B$ (relative to $X \cup Y$).*

Notice that $M_{n+m} = M_n *_{\mathcal{A}_2} M_m$, where $\mathcal{A}_2$ is the variety of all metabelian groups. Since WP is in P for groups from $\mathcal{M}$, Corollary 4.4 implies the following result.

Corollary 4.5. *If BGLP is NP-complete in M_2, then it is NP-complete in $M_r, r \ge 2$, relative to the standard bases.*

Remark 4.6. Corollary 4.5 easily generalizes to free groups in an arbitrary variety, provided they have WP decidable in polynomial time.

4.3. **Rectilinear Steiner Tree Problem.** The Steiner Tree Problem (STP), which was originally introduced by Gauss, is one of the initial twenty-one NP-complete problems that appeared in Karp's list [34]. We need the following *rectilinear* variation of STP.

Let $\mathbb{R}^2$ be the Euclidean plane and Γ the integer grid canonically embedded into $\mathbb{R}^2$ (all vertices from $\mathbb{Z}^2$ together with all the horizontal and vertical lines connecting them). If A is a finite subset of $\mathbb{Z}^2$, then a *rectilinear Steiner tree* (RST) for A is a subgraph T of Γ such that $A \cup T$ is connected in Γ. By $s(T)$ (size of T) we denote the number of edges in T. An RST for A is *optimal* if it has the smallest possible size among all RST for A; we denote such RST by T_A. Observe that a given A

may have several different optimal RST, but their size is the same; we denote it by $s(A)$.

Notice that, in general, T_A for A is not a spanning tree for A in Γ; it may contain some vertices from $\mathbb{Z}^2$ which are not in A (so-called, *Steiner* points). The Rectilinear Steiner Tree Problem (RSTP) asks, for a given finite $A \subseteq \mathbb{Z}^2$ and $k \in \mathbb{N}$, to decide if there exists some T_A for A with $s(T_A) < k$. It is known that RSTP is NP-complete [25].

4.4. **NP-completeness of BGLP in M_2.** Now we construct a polynomial reduction of RSTP to GLDP in M_2 relative to the standard basis $X = \{x, y\}$.

With each point $(s,t) \in \mathbb{Z}^2$ we associate a word

$$w_{s,t} = x_1^s x_2^t \cdot (x_2 x_1 x_2^{-1} x_1^{-1}) \cdot x_2^{-t} x_1^{-s}$$

in $F(x,y)$. Similarly, with a set of points $A = \{(s_1,t_1), \dots, (s_n,t_n)\} \subset \mathbb{Z}^2$, ordered in an arbitrary way, we associate a word

$$w_A = \prod_{i=1}^{n} w_{s_i,t_i}.$$

Observe that the word $w_{s,t}$, as well as w_A, belongs to $F' = [F,F]$, so in M_2 they define elements from M_2'. In particular, the path p_{w_A} is a closed path in the grid $\Gamma = \mathbb{Z}^2$, which is viewed as the Cayley graph of the abelianization F/F'.

For $A \subset \mathbb{Z}^2$, $(p,q) \in \mathbb{Z}^2$, and $m \in \mathbb{Z}$ we put

- $A + (p,q) = \{(s+p, t+q) \mid (s,t) \in A\}$,
- $mA = \{(ms, mt) \mid (s,t) \in A\}$.

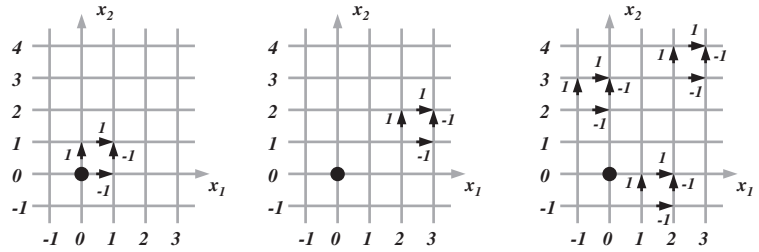

FIGURE 3. Flows on $\mathbb{Z}^2$ defined by words $w_{0,0}$, $w_{2,1}$, and $w_{\{(-1,2),(1,-1),(2,3)\}}$.

The following result is due to Hanan [30].

Theorem 4.7 ([30, Theorem 4])**.** *Let $A = \{(x_1,y_1), \dots, (x_n,y_n)\}$ be a finite subset of $\mathbb{Z} \times \mathbb{Z}$. There exists some T_A for A with the set of Steiner points $Q = \{(a_1,b_1), \dots, (a_q,b_q)\}$ such that $\{a_1, \dots, a_q\} \subseteq \{x_1, \dots, x_n\}$ and $\{b_1, \dots, b_q\} \subseteq \{y_1, \dots, y_n\}$.*

Corollary 4.8. *Let A be a finite subset of $\mathbb{Z} \times \mathbb{Z}$. Then*

(1) *$s(A + (b,c)) = s(A)$ for any $(b,c) \in \mathbb{Z}^2$;*
(2) *$ms(A) = s(mA)$ for any $m \in \mathbb{N}$.*

Proof. The first statement is obvious because the parallel shift $(x,y) \to (x,y)+(b,c)$ is an isomorphism of the Cayley graph Γ (in particular, an isometry).

To prove the second statement, notice first that $s(mA) \leq ms(A)$. Indeed, stretching T_A by the factor of m along all horizontal and vertical lines gives some RST for mA, hence the claim.

On the other hand, $s(A) \leq s(mA)/m$. To see this, observe that by Theorem 4.7 there exists $R = T_{mA}$ which lies inside the grid $m\mathbb{Z} \times m\mathbb{Z}$. Since the coordinates of all vertices in R are multiples of k, one can shrink R by the factor of m, in such a way that the image of R becomes an RST for A. Clearly, the size of the image is equal to $s(T_{mA})/m$, hence the result. □

Proposition 4.9. *Let A be a finite subset of $\mathbb{Z}^2$, $(b,c) \in A$, and $n = |A|$. Put $A^* = 10n(A - (b,c))$. Then $l_X(w_{A^*}) \in [20ns(A), 20ns(A) + 4n]$.*

Proof. Let u be a geodesic word for w_{A^*} relative to the basis X. Since $w_{A^*} \in F'$ the paths p_u and $p_{w_{A^*}}$ are closed paths in $\Gamma = \mathbb{Z}^2$ (viewed as the Cayley graph of M_2/M_2'). Hence u and w_{A^*} determine the same circulations $\pi_u = \pi_{A_{w^*}}$ on Γ. As described in Section 2.7, the flow π_u is associated with the subgraph Γ_u generated in Γ by $supp(\pi_u) \cup \{(0,0)\}$. It follows from the construction of the word w_{A^*} that the connected components of Γ_u are precisely the 1×1 squares in Γ, whose lower-left corners are located at the points from A^*. Notice that $(0,0) \in A - (b,c)$; hence $(0,0) \in A^*$. Now, if Q is a minimal forest for u (a subgraph of Γ of minimal size that makes the graph $\Gamma_u \cup Q$ connected in Γ), then by Theorem 2.11,

$$|u| = l_X(w_{A^*}) = \sum_{e \in supp(p_u)} \pi_u(e) + 2|E(Q)| = 4n + 2|E(Q)|. \tag{13}$$

Observe that an optimal RST T_{A^*} for A^* also makes the graph $\Gamma_u \cup Q$ connected in Γ; hence $|E(Q)| \leq s(A^*)$. Therefore, $l_X(w_{A^*}) \leq 20ns(A) + 4n$.

On the other hand assume that $|u| = l_X(w_{A^*}) < 20ns(A)$. Hence, from (13), there exists a minimal forest Q for A^* such that $2|E(Q)| < 20ns(A) - 4n$. Since every connected component has precisely 4 edges and there are n such components, it follows that there is an RST for A^* of size strictly less than $10ns(A)$, a contradiction with Corollary 4.8. This proves the proposition. □

Corollary 4.10. *Let A be a finite subset of $\mathbb{Z}^2$ and $k \in \mathbb{N}$. In the notation above,*

$$s(A) < k \Longleftrightarrow l_X(w_{A^*}) < 20nk + 4n.$$

In particular, this gives a polynomial reduction of RSTP to BGLP in M_2 relative to X.

Proof. Indeed, if $s(A) < k$, then by Proposition 4.9 $l_X(w_{A^*}) \leq 20ns(A) + 4n < 20nk + 4n$. On the other hand, suppose $s(A) \geq k$, say $s(A) = k + l$ for some positive $l \in \mathbb{N}$. Then, again by Proposition 4.9, $l_X(w_{A^*}) \geq 20ns(A) = 20n(k+l) > 20nk + 4n$, as required. □

Theorem 4.11. *GLDP in a free metabelian group M_2 is NP-complete.*

Proof. Corollary 4.10 gives a polynomial reduction of RSTP in $\mathbb{Z}^2$ to BGLP in M_2. Therefore BGLP in M_2 is NP-hard. Meanwhile, as was mentioned above, BGLP for M_2 is in NP, since WP in M_2 is polynomial. □

5. Open problems

Denote by $\mathcal{M}$ the class of all finitely generated metabelian groups.

Problem 5.1. Describe groups in $\mathcal{M}$ with GP in P. In particular, the following partial questions are of interest here:

- Are there any groups in $\mathcal{M}$ with GP not in P?
- Do polycyclic groups from $\mathcal{M}$ have GP in P?
- When do wreath products of two f.g. abelian groups have GP in P?

Notice, since WP is in P for groups from $\mathcal{M}$, it follows that GLP is at most in NP (Turing reducible in P time to BGLP which is in NP). This makes the following problem very interesting.

Problem 5.2. Classify groups in $\mathcal{M}$ with NP-complete BGLP. In particular, the following partial questions are of interest:

- Do polycyclic groups from $\mathcal{M}$ have GLP in P?
- Is it true that a wreath product $AwrB$ of finitely generated abelian groups has NP-complete BGLP if $A \neq 1$ and the torsion-free rank of B is at least 2?

Clearly, GLP is polynomial time reducible to GP. On the other hand, it is not clear if there are finitely presented (or finitely generated) groups where GP is not polynomial time reducible to GLP. It would be interesting to clarify the situation in the class $\mathcal{M}$. To this end we post the following.

Problem 5.3. Are there groups in $\mathcal{M}$ where GP is not polynomial time reducible to GLP?

Acknowledgement

We would like to thank M. Sapir and B. Steinberg who brought to our attention some relevant geometric ideas from semigroup theory.

References

[1] J. Almeida, *Semidirect products of pseudovarieties from the universal algebraist's point of view*, J. Pure Appl. Algebra 60 (1989), pp. 113–128. MR1020712 (91a:20068)
[2] J. Almeida and P. Weil, *Free profinite semigroups over semidirect products*, Izvestiya VUZ Matematika 39 (1995), pp. 3–31. MR1391317 (97e:20078)
[3] K. Auinger and B. Steinberg, *The geometry of profinite graphs with applications to free groups and finite monoids.*, Trans. Amer. Math. Soc. 356 (2004), pp. 805–851. MR2022720 (2005a:20040)
[4] ———, *A constructive version of the Ribes-Zalesskii product theorem*, Math. Z. 250 (2005), pp. 287–297. MR2178787 (2006h:20029)
[5] ———, *Constructing divisions into power groups*, Theoret. Comput. Sci. 341 (2005), pp. 1–21. MR2159641 (2007b:68127)
[6] J. Birman, *Braids, Links and Mapping Class Groups*, Annals of Math. Studies. Princeton University Press, 1974. MR0375281 (51:11477)
[7] B. Bollobas, *Graph Theory: An Introductory Course*, Graduate Texts in Mathematics. Springer, 1990. MR536131 (80j:05053)
[8] G. Chartrand and O. Oellermann, *Applied and Algorithmic Graph Theory*. McGraw-Hill, New York, 1993. MR1211413
[9] D. F. Cowan, *A class of varieties of inverse semigroups*, J. Algebra 141 (1991), pp. 115–142. MR1118319 (92f:20063)

[10] R. Crowell and R. Fox, *Introduction to Knot Theory*, Graduate Texts in Mathematics. Springer, 1984. MR0445489 (56:3829)
[11] CRyptography And Groups (CRAG) C++ Library, Available at http://www.acc.stevens.edu/downloads.php.
[12] R. Diestel, *Graph Theory*, Graduate Texts in Mathematics 173. Springer, 2005. MR2159259 (2006e:05001)
[13] C. Droms, J. Lewin, and H. Servatius, *The length of elements in free solvable groups*, Proc. Amer. Math. Soc. 119 (1993), pp. 27–33. MR1160298 (93k:20051)
[14] A. Dyubina, *Instability of the virtual solvability and the property of being virtually torsion-free for quasi-isometric groups*, Int. Math. Res. Notices 21 (2000), pp. 1097–1101. MR1800990 (2001j:20060)
[15] M. Elder, *A polynomial-time algorithm to compute geodesic length in BS(1,p)*, in preparation.
[16] D. B. A. Epstein, J. W. Cannon, D. F. Holt, S. V. F. Levy, M. S. Paterson, and W. P. Thurston, *Word processing in groups*. Jones and Bartlett Publishers, 1992. MR1161694 (93i:20036)
[17] A. Erschler, *On drift and entropy growth for random walks on groups*, Ann. of Prob. 31 (2003), pp. 1193–1204. MR1988468 (2004c:60018)
[18] A. Eskin, D. Fisher, and K. Whyte, *Quasi-isometries and Rigidity of Solvable Groups*, Pure Appl. Math. Quart. 3 (2007), pp. 927–947. MR2402598
[19] B. Farb and L. Mosher, *A rigidity theorem for the solvable Baumslag-Solitar groups. With an appendix by Daryl Cooper*, Invent. Math. 131 (1998), pp. 419–451. MR1608595 (99b:57003)
[20] ———, *Quasi-isometric rigidity for the solvable Baumslag-Solitar groups. II*, Invent. Math. 137 (1999), pp. 613–649. MR1709862 (2001g:20053)
[21] R. H. Fox, *Free differential calculus I*, Ann. of Math. (2) 57 (1953), pp. 547–560. MR0053938 (14:843d)
[22] ———, *Free differential calculus II*, Ann. of Math. (2) 59 (1954), pp. 196–210. MR0062125 (15:931e)
[23] ———, *Free differential calculus III*, Ann. of Math. (2) 64 (1956), pp. 407–419. MR0095876 (20:2374)
[24] ———, *Free differential calculus IV*, Ann. of Math. (2) 71 (1960), pp. 408–422. MR0111781 (22:2642)
[25] M. Garey and D. Johnson, *The Rectilinear Steiner Tree Problem is NP-complete*, SIAM J. Comput. 32 (1977), pp. 826–834. MR0443426 (56:1796)
[26] M. Gromov, *Groups of polynomial growth and expanding maps*, Publ. Math. IHES 53 (1981), pp. 53–73. MR623534 (83b:53041)
[27] ———, *Infinite groups as geometric objects*. Proceedings of the International Congress of Mathematicians, 1, pp. 385–395, 1983. MR804694 (87c:57033)
[28] K. W. Gruenberg, *Residual properties of infinite soluble groups*, Proc. London Math. Soc. 7 (1957), pp. 29–62. MR0087652 (19:386a)
[29] N. Gupta, *Free group rings*, Contemporary Mathematics 66. American Mathematical Society, 1987. MR895359 (88j:20030)
[30] M. Hanan, *On Steiner's problem with rectilinear distance*, SIAM J. Appl. Math. 14 (1966), pp. 255–265. MR0224500 (37:99)
[31] A. Hatcher, *Algebraic Topology*. Cambridge University Press, Cambridge, 2002. MR1867354 (2002k:55001)
[32] V. Kaimanovich and A. M. Vershik, *Random walks on discrete groups: Boundary and entropy.*, Ann. Probab. 11 (1983), pp. 457–490. MR704539 (85d:60024)
[33] M. I. Kargapolov and V. N. Remeslennikov, *The conjugacy problem for free solvable groups*, Algebra i Logika Sem. 5 (1966), pp. 15–25. Russian. MR0206080 (34:5905)
[34] R. M. Karp, *Reducibility Among Combinatorial Problems*. Complexity of Computer Computations, Computer Applications in the Earth Sciences, pp. 85–103. Springer, 1972. MR0378476 (51:14644)
[35] O. Kharlampovich, *A finitely presented solvable group with unsolvable word problem*, Izvest. Ak. Nauk, Ser. Mat. 45 (1981), pp. 852–873. MR631441 (82m:20036)
[36] O. Kharlampovich and M. Sapir, *Algorithmic problems in varieties*, Int. J. Algebr. Comput. 5 (1995), pp. 379–602. MR1361261 (96m:20045)
[37] W. Magnus, *On a theorem of Marshall Hall*, Ann. of Math. 40 (1939), pp. 764–768. MR0000262 (1:44b)

[38] S. W. Margolis and J. C. Meakin, *E-unitary inverse monoids and the Cayley graph of a group presentation*, J. Pure Appl. Algebra 58 (1989), pp. 45–76. MR996174 (90f:20096)
[39] S. W. Margolis, J. C. Meakin, and J. B. Stephen, *Free objects in certain varieties of inverse semigroups*, Canadian J. Math. 42 (1990), pp. 1084–1097. MR1099459 (92j:20057)
[40] J. Matthews, *The conjugacy problem in wreath products and free metabelian groups*, Trans. Amer. Math. Soc. 121 (1966), pp. 329–339. English transl., Soviet Math. Dokl. **8** (1967), 555–557. MR0193130 (33:1351)
[41] J. Milnor, *Growth of finitely generated solvable groups*, J. Differ. Geom. 2 (1968), pp. 447–449. MR0244899 (39:6212)
[42] L. Mosher, M. Sageev, and K. Whyte, *Quasi-actions on trees. I. Bounded valence*, Ann. of Math. (2) 158 (2003), pp. 115–164. MR1998479 (2004h:20055)
[43] W. D. Munn, *Free inverse semigroups*, Proc. London Math. Soc. 29 (1974), pp. 385–404. MR0360881 (50:13328)
[44] H. Newmann, *Varieties of groups*. Springer, 1967. MR0215899 (35:6734)
[45] W. Parry, *Growth Series of Some Wreath Products*, Trans. Amer. Math. Soc. 331 (1992), pp. 751–759. MR1062874 (92h:20061)
[46] V. N. Remeslennikov, *Certain properties of the Magnus embedding*, Algebra i Logika 8 (1969), pp. 72–76.
[47] V. N. Remeslennikov and N. S. Romanovskii, *Algorithmic problems for solvable groups*. Word Problems II: The Oxford book, pp. 337–346. North-Holland, 1980. MR579951 (81h:20044)
[48] V. N. Remeslennikov and V. G. Sokolov, *Certain properties of the Magnus embedding*, Algebra i Logika 9 (1970), pp. 566–578. MR0292920 (45:2001)
[49] J. Rhodes and B. Steinberg, *Profinite semigroups, varieties, expansions and the structure of relatively free profinite semigroups*, Internat. J. Algebra Comput. 11 (2001), pp. 627–672. MR1880372 (2002j:20114)
[50] A. L. Shmel′kin, *Wreath products and group varities*, Izvestiya AN SSSR, seriya matematika, 29 (1965), pp. 149–176.
[51] I. M. Singer and J. A. Thorpe, *Lectures Notes on Elementary Topology and Geometry*, Undergraduate Texts in Mathematics. Springer-Verlag, 1967. MR0213982 (35:4834)
[52] A. M. Vershik, *Geometry and dynamics on the free solvable groups*, preprint. Erwin Schroedinger Institute, Vienna, 1999, pp. 1–16.
[53] ———, *Dynamic theory of growth in groups: Entropy, boundaries, examples*, Uspekhi Mat. Nauk 55 (2000), pp. 59–128. MR1786730 (2001m:37019)
[54] A. M. Vershik and S. Dobrynin, *Geometrical approach to the free solvable groups*, Internat. J. Algebra Comput. 15 (2005), 1243–1260. MR2197831 (2006m:20049)
[55] B. A. F. Wehrfritz, *Two examples of soluble groups that are not conjugacy separable*, J. London Math. Soc. 2 (1973), pp. 312–316. MR0338176 (49:2942)
[56] ———, *Another example of a soluble group that is not conjugacy separable*, J. London Math. Soc. 14 (1976), pp. 380–382. MR0422429 (54:10418)

Department of Mathematics and Statistics, McGill University, Montreal, Canada H3A 2K6
E-mail address: amiasnikov@gmail.com

Department of Mathematics, Omsk State University, Omsk, Russia 644077
E-mail address: romankov@math.omsu.omskreg.ru

Department of Mathematical Sciences, Stevens Institute of Technology, Hoboken, New Jersey 07030
E-mail address: aushakov@stevens.edu

Petersburg Department of Steklov Institute of Mathematics, St. Petersburg, Russia 191023
E-mail address: vershik@pdmi.ras.ru