On the Parameterization of Solutions for Equations in Free Groups

Alexander A. Razborov
Steklov Mathematical Institute
Vavilova 42, 117966, GSP–1, Moscow, RUSSIA

to appear in *International Journal of Algebra and Computation*

Abstract

In this paper we study the question of representing the set of solutions of an equation in a free group as the set of values of finitely many parametric functions depending on free word variables and, possibly, on certain other parameters. The main result says that if we allow in our representations arbitrary superpositions of arbitrary basic parametric functions having less than g free word variables, then the set of solutions of the quadratic coefficient-free equation $x_1x_2\dots x_{2g} = x_{2g}\dots x_2x_1$ (which is equivalent to $[x_1, y_1]\dots[x_g, y_g] = 1$) can not be represented in this way. A similar statement holds for the same equation considered over a free semigroup.

1. Introduction

The most important class of algebraic systems for which the question of the algorithmic decidability of their elementary theories was left open after the remarkable breakthrough in the area achieved in the 60's, is the class of free groups. It is also open whether the elementary theories of non-abelian free groups with different number of generators coincide or not; both these questions are usually attributed to Tarski.

The simpler question of deciding whether a system of equations in a free group has at least one solution has met with much more success. The affirmative answer was known for a long time for the case of quadratic equations (see e.g. [5, 17]). In [6, 23, 24, 19] the algorithms deciding the solvability for other restricted classes of systems were constructed.

The complete solution for arbitrary systems in free groups was given by Makanin [16, 15][a]. Two immediate applications of this result should be noted here: the algorithmic decidability of positive and universal fragments of the elementary theories of free groups ([15], see also [14] where the latter result was slightly strengthened).

The next natural question is how to describe in external terms the structure of the solution set of a system provided this set is non-empty. Again, for the case of quadratic equations some description was known long ago. Recently it was substantially simplified by Comerford and Edmonds [3] and by Grigorchuk and Kurchanov [4].

Lyndon [6] suggested a general pattern to approach this problem. Namely, he introduced the notion of a *parametric word* as a function depending on integer parameters. Parametric words are built by taking finite superpositions of plain group operations along with the powering A^λ, where A is a word and λ is an integer parameter. Note that for any simple non-empty word A, $\left\{ A^\lambda \mid \lambda \in \mathbb{Z} \right\}$ equals the set of solutions of the equation

$$[x, A] = 1. \tag{1}$$

Lyndon proved that the set of solutions of arbitrary one-variable equations in free groups can be expressed as the union of the set of values of finitely many parametric words. Subsequently, his result was strengthened and simplified in [12, 2, 11].

Appel [1], however, showed that the set of all solutions of a simple quadratic equation in two variables already can not be represented in terms of Lyndon's parametric words. Hence the question of finding more complex parametric functions which would suffice for this and more complicated equations came to the surface.

Along these lines Khmelevskiĭ introduced a parametric function which parameterizes solutions of the equation

$$[x, y] = [a, b] \tag{2}$$

previously considered by Malcev in [17] (this function was afterwards called the *Nielsen-Khmelevskiĭ function*). Khmelevskiĭ [23, 24] proved that the set of solutions of those systems for which he established the algorithmic decidability, can be always represented in terms of finite superpositions of Lyndon's parametric words and Nielsen-Khmelevskiĭ functions.

Razborov [21] gave a partial answer to the problem of representing the solution set of an arbitrary system in a free group. That representation embraced for sure only solutions subjected to the condition that the period γ in every occurence of the form P^γ into their components is bounded by an arbitrary but fixed constant. One application of

[a] the original proof in [16] contained a gap shortly after fixed in [15]

this technique was the solution of the so-called rank problem: the rank of a coefficient-free system of equations in a free group is algorithmically computable (the *rank* of a system, by definition, is the maximal possible rank of subgroups generated by its solutions).

An unconditional description of the solution set for arbitrary systems was given by Razborov in [20] (for an English account see [10]). The final result in plain words says that this set can be represented as a union of finite superpositions of explicit parametric functions. These functions correspond to the set of solutions of "canonical" quadratic equations more or less in the same sense in which Lyndon's parametric words and Nielsen-Khmelevskiĭ functions correspond to solutions of equations (1), (2). In other words, arbitrary systems in free groups can in a sense be efficiently reduced to quadratic equations. Another application of some of the techniques used in [20] was given by Rips (personal communication).

In view of the above-mentioned description, it is important to study quadratic equations in a free group and corresponding parametric functions. Another good reason for this is their close connection to the Poincaré Conjecture (see e.g. [7, Section 4.4] or [10, §2.2]).

In this paper we are interested in the natural question if the hierarchy of parametric functions corresponding to quadratic equations with the number of unknowns growing to infinity is proper or not. We show that it is proper and in fact prove much more. Namely, we show that the set of solutions of the canonical quadratic equation $[x_1, y_1] \dots [x_g, y_g] = 1$ not only can not be represented as a finite union of superpositions of parametric functions corresponding to smaller g, but in fact can not be represented in this way even if we allow absolutely arbitrary parametric functions in less than g variables (Theorem 3.3). A similar statement holds also for the case of semigroups (Theorem 4.3). These results should greatly reduce the hope of ever getting a drastic simplification of the description from [20].

The paper is organized as follows. In Section 2 we fix some notation. The next section 3 contains formal definitions of basic notions and the statement of Theorem 3.3 which is the main result of this paper.

In Section 4 we introduce a hierarchy of words in a free group and show how it can be used for proving the main result (Lemma 4.2). The next section 5 contains some auxiliary facts and statements which are needed for the proof of Lemma 4.2 but are not directly related to the concepts involved in its statement. Based upon these preliminaries, we develop in Section 6 the machinery which allows us to prove lower bounds for the hierarchy defined in Section 4. After all this work is done, we comparatively easily complete the proof of Lemma 4.2 and Theorem 3.3 in Section 7.

A preliminary version of this paper appeared in [22].

2. Notation

The aim of this section is to agree upon some notation.

We will denote the free group with basis $x_1, \ldots, x_n$ by $\langle x_1, \ldots, x_n \rangle$. $\mathrm{rk}(F)$ is the rank of the free group F. $\mathrm{Gp}(A_1, \ldots, A_m)$ is the subgroup generated by elements $A_1, \ldots, A_m$. The symbol $=$ will stand for the group equality.

All words denoted by a single capital Latin letter are assumed to be reduced. Λ stands for the empty word. A word is *positive* if it does not contain inverse letters x_i^{-1}. Given words $W(u_1, \ldots, u_n)$ and $A_1, \ldots, A_n$, we denote by $W(A_1, \ldots, A_n)$ or sometimes by $W[A_1/u_1, \ldots, A_n/u_n]$ the result of substituting $A_1, \ldots, A_n$ for $u_1, \ldots, u_n$ into $W(u_1, \ldots, u_n)$. Note that we do *not* perform cancellations in the result of this substitution; this does *not* contradict the convention above as the expression $W(A_1, \ldots, A_n)$ is not a single letter. $W_1(A_1, \ldots, A_n) \overline{\circ} W_2(B_1, \ldots, B_m)$, by definition, means that both left- and right-hand sides are reduced and equal to each other. The same applies to the record $A_1 \ldots A_r \overline{\circ} B_1 \ldots B_s$. We will sometimes indicate the occurrences of the form $X \overline{\circ} REQ$ of a word E into a word X as $R * E * Q$ (this notation is borrowed from [8]). $A \subseteq B$ will mean the statement $\exists C, D(B \overline{\circ} CAD)$. Like with the notation $\overline{\circ}$, $W_1(A_1, \ldots, A_n) \subseteq W_2(B_1, \ldots, B_m)$ will also presume that the words on both sides are reduced. $\partial(A)$ is the length of the word A.

For a word A, we will denote by $\widehat{A}$ the unique cyclically reduced word such that $A \overline{\circ} P^{-1}\widehat{A}P$ for some P. $A \sim B$ will mean that A and B are conjugated which is equivalent to saying that $\widehat{A}$ and $\widehat{B}$ are cyclic shifts of each other. A cyclically reduced non-empty word P is *simple* if it can not be represented in the form $P \overline{\circ} Q^r$ with $r > 1$.

We will denote by E the functor from the category of groups into the category of vector spaces over $\mathbb{Q}$ defined by $E(G) \rightleftharpoons G/[G, G] \otimes \mathbb{Q}$. For an element A of a group F, $\widetilde{A}$ will stand for the image of A in $E(F)$. We say that a group homomorphism $\pi : F \longrightarrow G$ is *flat* if the induced homomorphism $E(\pi) : E(F) \longrightarrow E(G)$ is injective.

Vectors of letters like $\phi_1, \ldots, \phi_m$ or $X_1, \ldots, X_n$ etc. will be denoted by $\bar{\phi}, \bar{X}, \ldots$ When we need to consider simultaneously several vectors of the same type, we use superscripts: $\bar{X}^{(1)}, \ldots, \bar{X}^{(h)}, \ldots$

3. Systems of equations in free groups and their parametric solutions

We fix the following three group alphabets:

$$\begin{array}{lll} \Sigma_0 \rightleftharpoons \{x_1, \ldots, x_n, \ldots, x_1^{-1}, \ldots, x_n^{-1}, \ldots\} & - & \text{the alphabet of } \textit{unknowns}, \\ \Sigma_1 \rightleftharpoons \{a_1, \ldots, a_\omega, \ldots, a_1^{-1}, \ldots, a_\omega^{-1}, \ldots\} & - & \text{the alphabet of } \textit{coefficients}, \\ \Sigma_2 \rightleftharpoons \{u_1, \ldots, u_n, \ldots, u_1^{-1}, \ldots, u_n^{-1}, \ldots\} & - & \text{the alphabet of } \textit{free word variables}. \end{array}$$

In the sequel free word variables will stand for "generic" words over Σ_1.

A *system of equations in a free group* is a system of equalities of the form

$$\begin{cases} \phi_1(\bar{x}, \bar{a}) = 1, \\ \cdots \\ \phi_m(\bar{x}, \bar{a}) = 1, \end{cases} \tag{3}$$

where $\phi_1, \ldots, \phi_m$ are words over $\Sigma_0 \cup \Sigma_1$. A system is *coefficient-free* if its equations do not contain coefficients. A *solution* of the system (3) is a vector $\bar{X}$ of reduced words over Σ_1 such that $\phi_i(\bar{X}, \bar{a}) = 1$ for all $1 \leq i \leq m$. We denote by $\mathrm{Sol}(\bar{\phi})$ the set of all solutions of the system (3). A solution $\bar{X} = (X_1, \ldots, X_n)$ is *positive* if $X_1, \ldots, X_n$ are positive words.

Now we give the formal definition of parametric functions, terms and solutions. It corresponds to the usual practice of working with systems of equations in free groups. However, since our main result is negative, the definition presented here is quite general. In practice one deals only with very explicit parametric functions.

Let $P^{(1)}, P^{(2)}, \ldots, P^{(m)}, \ldots$ be some sets of absolutely arbitrary nature. We associate with each $P^{(i)}$ countably many *parameter variables* $p_1^{(i)}, \ldots, p_k^{(i)}, \ldots$ which afterwards will take values from $P^{(i)}$.

A *parametric function of rank* r is an arbitrary mapping $F(p_1^{(i_1)}, \ldots, p_l^{(i_l)})$ of the form $F : P^{(i_1)} \times \cdots \times P^{(i_l)} \longrightarrow \{u_1^{\pm 1}, \ldots, u_r^{\pm 1}\}^*$. As $u_1, \ldots, u_r$ stand for "generic" words over Σ_1^*, it is convenient to think alternatively of F as of the mapping

$$F : P^{(i_1)} \times \cdots \times P^{(i_l)} \times \underbrace{\Sigma_1^* \times \cdots \times \Sigma_1^*}_{r} \longrightarrow \Sigma_1^* \tag{4}$$

which takes $\boldsymbol{p}_1^{(i_1)} \in P^{(i_1)}, \ldots, \boldsymbol{p}_l^{(i_l)} \in P^{(i_l)}$ and $U_1, \ldots, U_r \in \Sigma_1^*$ to

$$F(\boldsymbol{p}_1^{(i_1)}, \ldots, \boldsymbol{p}_l^{(i_l)})[U_1/u_1, \ldots, U_r/u_r].$$

Thus we may consider F as a function symbol $F(p_1^{(i_1)}, \dots, p_l^{(i_l)}, u_1, \dots, u_r)$ in the first order many-sorted logic with the interpretation provided by (4).

Now we define in the usual way terms in the language consisting of constants from Σ_1, plain group operations and some parametric functions. Namely, given a set $\mathcal{L}$ of parametric functions, we define *parametric terms over* $\mathcal{L}$ recursively by

a) $a_1^{\pm 1}, \dots, a_\omega^{\pm 1}, \dots$ are parametric terms,

b) $u_1^{\pm 1}, \dots, u_n^{\pm 1}, \dots$ are parametric terms,

c) if S and T are parametric terms then S^{-1} and ST are parametric terms,

d) if $F(p_1^{(i_1)}, \dots, p_l^{(i_l)}, u_1, \dots, u_r) \in \mathcal{L}$, $q_1^{(i_1)}, \dots, q_l^{(i_l)}$ are parameter variables and $T_1, \dots, T_r$ are parametric terms then $F(q_1^{(i_1)}, \dots, q_l^{(i_l)}, T_1, \dots, T_r)$ is a parametric term.

The interpretation (4) is extended in the obvious way to arbitrary parametric terms. Also, similarly to the case of function symbols, another way to look at a parametric term $T(p_1^{(i_1)}, \dots, p_l^{(i_l)}, u_1, \dots, u_r)$ is to interpret it as a mapping

$$T(p_1^{(i_1)}, \dots, p_l^{(i_l)}) : P^{(i_1)} \times \dots \times P^{(i_l)} \longrightarrow \{a_1^{\pm 1}, \dots, a_\omega^{\pm 1}, \dots, u_1^{\pm 1}, \dots, u_r^{\pm 1}\}^*$$

with the property

$$T(\boldsymbol{p}_1^{(i_1)}, \dots, \boldsymbol{p}_l^{(i_l)}, U_1, \dots, U_r) = T(\boldsymbol{p}_1^{(i_1)}, \dots, \boldsymbol{p}_l^{(i_l)})[U_1/u_1, \dots, U_r/u_r].$$

We say that n parametric terms $T_1(\bar{p}, \bar{u}), \dots, T_n(\bar{p}, \bar{u})$ (over $\mathcal{L}$) form a *parametric solution* (over $\mathcal{L}$) of a system $\phi(x_1, \dots, x_n, \bar{a}) = 1$ if

$$\left\{ \bar{X} \;\middle|\; \exists \boldsymbol{p}_\nu^{(i_\nu)} \in P^{(i_\nu)} \exists \bar{U} \in \Sigma_1^* \; \bar{X} = \bar{T}(\bar{\boldsymbol{p}}, \bar{U}) \right\} \subseteq \mathrm{Sol}(\bar{\phi}). \tag{5}$$

Another convenient way to look at parametric solutions is provided by the following easy

Proposition 3.1. *Parametric terms $\bar{T}(\bar{p}, \bar{u})$ form a parametric solution of a system $\bar{\phi}(\bar{x}, \bar{a}) = 1$ if and only if for each $\bar{\boldsymbol{p}} \in \bar{P}$, $\bar{T}(\bar{\boldsymbol{p}})$ is an ordinary solution of the same system over the extended alphabet $\Sigma_1 \cup \Sigma_2$.*

This is an easy consequence of the following well-known fact:

Statement 3.2 (see e.g. [18]). *A vector $\bar{X}(\bar{u}, \bar{a})$ is a solution over the extended alphabet $\Sigma_1 \cup \Sigma_2$ of a system $\bar{\phi}(\bar{x}, \bar{a}) = 1$ if and only if for each $\bar{U} \in \Sigma_1^*$, $\bar{X}(\bar{U}, \bar{a})$ is an ordinary solution of the same system.*

We say that a finite collection $\bar{T}^{(1)}(\bar{p},\bar{u}),\dots,\bar{T}^{(h)}(\bar{p},\bar{u})$ of parametric solutions of a system $\bar{\phi}(\bar{x},\bar{a})=1$ *represents* (or *parameterizes*) the solution set of this system if

$$\bigcup_{i=1}^{h}\left\{\bar{X}\;\middle|\;\exists \boldsymbol{p}_\nu^{(i_\nu)}\in P^{(i_\nu)}\exists\bar{U}\in\Sigma_1^*\;\bar{X}=\bar{T}^{(i)}(\bar{\boldsymbol{p}},\bar{U})\right\}=\mathrm{Sol}(\phi). \tag{6}$$

Let $\mathcal{L}_r$ denote the language consisting of *all* parametric functions of rank at most r. The main result whose proof will be given in the rest of this paper is the following theorem.

Theorem 3.3. *Let $g\geq 3$. Then there is no finite collection of parametric solutions over $\mathcal{L}_{g-1}$ representing the solution set of the single quadratic coefficient-free equation*

$$[x_1,y_1]\dots[x_g,y_g]=1. \tag{7}$$

A similar result holds for semigroups, it will be stated in Section 4 (Theorem 4.3).

Corollary 3.4. *For any* **finite** *family $\mathcal{L}$ of parametric functions there exists $g>0$ such that the solution set of the equation* (7) *can not be represented by any finite collection of parametric solutions over $\mathcal{L}$.*

Remark 3.5. The bound on the rank in Theorem 3.3 is tight since there exist trivial parametric functions from $\mathcal{L}_g$ representing the solution set of the equation (7). These functions merely enumerate all solutions of (7) over $\{u_1^{\pm 1},\dots,u_g^{\pm 1}\}^*$.

Remark 3.6. In fact, we can prove even a slightly stronger statement. Namely, one can impose arbitrary connections between different parameters. That is, the full domain $P^{(i_1)}\times\dots\times P^{(i_l)}$ in (5), (6) from which the values of parameter variables are taken, can be replaced by an *arbitrary* subset of $P^{(i_1)}\times\dots\times P^{(i_l)}$. It will be clear from the proof of Theorem 3.3 that it is still valid in this more general setting.

A parametric term is *coefficient-free* if it does not contain occurrences of coefficients. The following simple fact says that, while describing the solution set of a coefficient-free system, we can restrict ourselves to coefficient-free parametric terms.

Proposition 3.7. *For every parametric solution $\bar{T}$ of a coefficient-free system $\bar{\phi}(\bar{x})=1$ there exists another parametric solution $\bar{T}'$ of the same system whose components are coefficient-free and such that*

$$\left.\begin{aligned}&\left\{\bar{X}\;\middle|\;\exists \boldsymbol{p}_\nu^{(i_\nu)}\in P^{(i_\nu)}\exists\bar{U}\in\Sigma_1^*\;\bar{X}=\bar{T}(\bar{\boldsymbol{p}},\bar{U})\right\}\\ &\quad\subseteq\left\{\bar{X}\;\middle|\;\exists \boldsymbol{p}_\nu^{(i_\nu)}\in P^{(i_\nu)}\exists\bar{U}\in\Sigma_1^*\;\bar{X}=\bar{T}(\bar{\boldsymbol{p}},\bar{U})\right\}.\end{aligned}\right\} \tag{8}$$

Proof. $\bar{T}'$ is obtained from $\bar{T}$ by replacing all occurrences $a_1, \ldots, a_\omega, \ldots$ of coefficients by new free word variables $v_1, \ldots, v_\omega, \ldots$ not appearing in $\bar{T}$. Then (8) is obvious, and the fact that $\bar{T}'$ is a parametric solution readily follows from Proposition 3.1.■

4. A hierarchy of words in free groups

Fix $r \geq 2$. Let $F_1 \rightleftharpoons \langle a_1, \ldots, a_\omega, \ldots \rangle$ and $F_2 \rightleftharpoons \langle u_1, \ldots, u_n, \ldots \rangle$. Our main goal in this section is to define a hierarchy of words $\mathcal{H}(0,r) \subseteq \mathcal{H}(1,r) \subseteq \ldots \subseteq F_2$ whose members majorize the sets of values of coefficient-free parametric terms over $\mathcal{L}_r$. The definition goes recursively as follows:

$$\mathcal{H}(0,r) \rightleftharpoons \Sigma_2,$$
$$\mathcal{H}(d+1,r) \rightleftharpoons \cup \{ \mathrm{Gp}(A_1, \ldots, A_r) \mid A_1, \ldots, A_r \in \mathcal{H}(d,r) \}.$$

Lemma 4.1. *For every $r \geq 2$ and every coefficient-free parametric term $T(\bar{p}, \bar{u})$ over $\mathcal{L}_r$ there exists $d \geq 0$ such that $\left\{ T(\bar{\boldsymbol{p}}) \,\middle|\, \bar{\boldsymbol{p}} \in \bar{P} \right\} \subseteq \mathcal{H}(d,r)$.*

Proof. Obvious induction on complexity of T (d in fact can be chosen as the depth of the term T).■

In the subsequent sections we will develop machinery which will allow us to prove "lower bounds" for the hierarchy $\mathcal{H}(d,r)$ that is to show that some explicit words do not belong to certain members of this hierarchy. Unfortunately, even this is not enough since we also should take care somehow of the substitutions $u_1 \to U_1, \ldots, u_r \to U_r$ when we do not have any a priori information about the words $U_1, \ldots, U_r$ (this is the main complication comparatively with the result of Appel [1] where such substitutions do not occur). In the next lemma we formulate the lower bound for our hierarchy which *is* sufficient for completing the proof of Theorem 3.3.

Lemma 4.2 (Main). *For any $g \geq 3$ and $d \geq 0$ there exists a positive solution $\bar{X} = (X_1, \ldots, X_{2g})$ of the equation*

$$x_1 x_2 \ldots x_{2g} x_1^{-1} x_2^{-1} \ldots x_{2g}^{-1} = 1 \tag{9}$$

with the following properties:

a) $\mathrm{Gp}(X_1, \ldots, X_{2g}) = \langle a_1, \ldots, a_g \rangle$,

b) *for every* **flat** *homomorphism* $\pi : \langle a_1, \ldots, a_g \rangle \longrightarrow F_2$, $\pi(X_2X_1) \notin \mathcal{H}(d, g-1)$.

The proof of this lemma will be given in the next three sections. Now we deduce from it our main result.

Proof of Theorem 3.3 from Lemma 4.2. Assume on the contrary that a finite collection of parametric solutions over $\mathcal{L}_{g-1}$ represents the solution set of the equation (7). Applying [7, Proposition 1.7.8 + Proposition 1.7.6] to the word $x_1x_2 \ldots x_{2g}x_1^{-1}x_2^{-1} \ldots x_{2g}^{-1}$, we find that there exists an automorphism of the group $\langle x_1, \ldots, x_{2g} \rangle$ which takes this word to $[x_1, x_2] \ldots [x_{2g-1}, x_{2g}]$. This implies that the solution set of the equation (9) is also representable over $\mathcal{L}_{g-1}$ by a finite collection of parametric solutions $\bar{T}^{(1)}(\bar{p}, \bar{u}), \ldots, \bar{T}^{(h)}(\bar{p}, \bar{u})$.

By Proposition 3.7, we may assume that all parametric terms $T_j^{(i)}$ involved in this representation are coefficient-free. By Lemma 4.1, there exists an uniform bound d such that for every i and j and every value of parameter variables $\bar{\boldsymbol{p}}$, $T_j^{(i)}(\bar{\boldsymbol{p}}) \in \mathcal{H}(d, g-1)$. As $g \geq 3$, this in particular implies $T_2^{(i)}(\bar{\boldsymbol{p}})T_1^{(i)}(\bar{\boldsymbol{p}}) \in \mathcal{H}(d+1, g-1)$ for every $i \leq h$.

Choose now the solution $\bar{X}$ according to Lemma 4.2 with $d := d+1$. We claim that it is not in $\bigcup_{i=1}^{h} \left\{ \bar{X} \;\middle|\; \exists \boldsymbol{p}_\nu^{(i_\nu)} \in P^{(i_\nu)} \exists \bar{U} \in \Sigma_1^* \; \bar{X} = \bar{T}^{(i)}(\bar{\boldsymbol{p}}, \bar{U}) \right\}$ which will give the desired contradiction.

Assume on the contrary that $\bar{X} = \bar{T}^{(i)}(\bar{\boldsymbol{p}}, U_1, \ldots, U_l)$ for some $i \leq h$, $\bar{\boldsymbol{p}} \in \bar{P}$ and $U_1, \ldots, U_l \in F_1$. Let $\phi : F_2 \longrightarrow F_1$ be the homomorphism defined by

$$\phi(u_\nu) = \begin{cases} U_\nu, \; 1 \leq \nu \leq l, \\ 1, \; \nu \geq l+1. \end{cases}$$

Then $\phi\left(\mathrm{Gp}\left(\left\{ \bar{T}_j^{(i)}(\bar{\boldsymbol{p}}) \mid 1 \leq j \leq 2g \right\}\right)\right) = \mathrm{Gp}(X_1, \ldots, X_{2g}) = \langle a_1, \ldots, a_g \rangle$.

On the other hand, by Proposition 3.1, $\bar{T}^{(i)}(\bar{\boldsymbol{p}})$ itself is a solution of (9). Hence [7, Proposition 1.7.13] gives us $\mathrm{rk}\left(\mathrm{Gp}\left(\bar{T}^{(i)}(\bar{\boldsymbol{p}})\right)\right) \leq g$. By [7, Proposition 1.2.7] this implies that in fact ϕ induces an isomorphism between $\mathrm{Gp}\left(\bar{T}^{(i)}(\bar{\boldsymbol{p}})\right)$ and $\langle a_1, \ldots, a_g \rangle$. Let

$$\pi : \langle a_1, \ldots, a_g \rangle \xrightarrow{\phi^{-1}} \mathrm{Gp}\left(\bar{T}^{(i)}(\bar{\boldsymbol{p}})\right) \xrightarrow{\mathrm{id}} F_2.$$

We have $\phi\pi = \mathrm{id}$, hence $E(\phi)E(\pi) = \mathrm{id}$ which implies that π is flat. Also, $\pi(X_2X_1) = T_2^{(i)}(\bar{\boldsymbol{p}})T_1^{(i)}(\bar{\boldsymbol{p}}) \in \mathcal{H}(d+1, g-1)$. This contradicts property b) of the solution $\bar{X}$ in Lemma 4.2 which completes the proof of Theorem 3.3 modulo this lemma.■

We conclude this section with an extension of our main theorem 3.3 to the case of semigroups. Definitions of parametric functions, terms, solutions etc. are extended to the case of semigroups in the obvious way. Let $\mathcal{L}_r^+$ be the analogue of $\mathcal{L}_r$ for the case of semigroups.

Theorem 4.3. *For each $g \geq 3$, there is no finite collection of parametric semigroup solutions over $\mathcal{L}^+_{g-1}$ representing the solution set of the coefficient-free equation*

$$x_1 x_2 \dots x_{2g} \stackrel{\circ}{=} x_{2g} \dots x_2 x_1 \tag{10}$$

in the free semigroup.

Proof. Each parametric semigroup solution of the equation (10) can be also considered as a parametric (group) solution of the equation (9). Applying the argument from the proof of Theorem 3.3 to the resulting collection of parametric group solutions of (9), we will find a positive solution $\bar{X}$ that is not represented by this collection. Since $\bar{X}$ is positive, it is also a solution of the semigroup equation (10). It is easy to see that $\bar{X}$ is not represented by the original collection of parametric semigroup solutions.■

5. Some useful facts

We start proving Lemma 4.2. This section contains in a convenient form several auxiliary facts and concepts not directly connected with this work.

Proposition 5.1 (see e.g. [8, 3.2.3]). *Let A, B be non-empty words, and $A^s A' \stackrel{\circ}{=} B^t B'$, where A' is an initial segment of A and B' is an initial segment of B. Assume that $\partial(A^s A') \geq \partial(A) + \partial(B)$. Then for some simple word P and for some integers k, l we have $A \stackrel{\circ}{=} P^k$ and $B \stackrel{\circ}{=} P^l$.*

Remark 5.2. Quite often we will use the conclusion of this lemma only in the weak form which says that $\widetilde{A}$ and $\widetilde{B}$ are linearly dependent over $\mathbb{Q}$.

The following observation is well known.

Proposition 5.3. *If $A \subseteq P^\infty$ and $A^{-1} \subseteq P^\infty$ then $\partial(A) \leq \frac{1}{2}\partial(P)$.*

Proof. Truncating A if necessary, we may assume that $\partial(A) \leq \partial(P)$. Then A is an initial segment of P', and A^{-1} is an initial segment of P'', where P' and P'' are cyclic shifts of P. Let $P' \stackrel{\circ}{=} P_1 P_2$, $P'' \stackrel{\circ}{=} P_2 P_1$. Assume w.l.o.g. that $\partial(P_1) \leq \partial(P_2)$.

$\partial(A) > \frac{1}{2}\partial(P)$ would imply $P_2 \stackrel{\circ}{=} P_2' P_2''$, $P_2' \neq \Lambda$, $A \stackrel{\circ}{=} P_1 P_2'$ and then $A^{-1} \stackrel{\circ}{=} (P_2')^{-1} P_1^{-1}$. Hence $(P_2')^{-1}$ would be also an initial segment of $P'' \stackrel{\circ}{=} P_2 P_1$. From this

we would have $P_2' \overline{\circ} (P_2')^{-1}$ which is impossible since $P_2' \neq \Lambda$. Hence the assumption $\partial(A) > \frac{1}{2}\partial(P)$ leads to a contradiction which completes the proof.■

For the definitions of *elementary Nielsen transformations*, *Nielsen transformations* and *Nielsen reduced* (or *N-reduced*) *vectors* see e.g. [7, Chapter 1.2]. The main property of a Nielsen reduced vector $(A_1, \ldots, A_n)$ is that in every product $A_{i_1}^{\epsilon_1} A_{i_2}^{\epsilon_2} A_{i_3}^{\epsilon_3}$ the middle syllable $A_{i_2}^{\epsilon_2}$ does not cancel completely unless $i_1 = i_2,\ \epsilon_1 = -\epsilon_2$ or $i_2 = i_3,\ \epsilon_2 = -\epsilon_3$. The following theorem refines [7, Proposition 1.2.3].

Theorem 5.4. *Let $(A_1, \ldots, A_n)$ be a Nielsen reduced vector. Then there exist words $W_1(\bar{u}), \ldots, W_n(\bar{u})$ and $\bar{U}$ such that:*

a) $\partial(W_i(\bar{u})) \leq 4n$ $(1 \leq i \leq n)$,

b) $A_i \overline{\circ} W_i(\bar{U})$ $(1 \leq i \leq n)$,

c) *for every words A and $H(a_1, \ldots, a_n)$ such that $A = H(A_1, \ldots, A_n)$ we have*

$$A \overline{\circ} H'(\bar{U}),$$

where $H'(\bar{u}) = H(W_1(\bar{u}), \ldots, W_n(\bar{u}))$.

Proof. Write down explicitly cancellations in all products of the form $A_i^{\epsilon} A_j^{\delta}$ $((i,\epsilon) \neq (j,-\delta))$. That is to say,

$$\begin{cases} A_i^{\epsilon} \overline{\circ} B_{i\epsilon j\delta} C_{i\epsilon j\delta}, \\ A_j^{\delta} \overline{\circ} C_{i\epsilon j\delta}^{-1} D_{i\epsilon j\delta}, \end{cases} \tag{11}$$

where the product $B_{i\epsilon j\delta} D_{i\epsilon j\delta}$ is reduced. For every fixed $i \leq n$, the occurrences $B_{i1j\delta} * C_{i1j\delta} *$ and $* C_{i(-1)j\delta}^{-1} * B_{i(-1)j\delta}^{-1}$ into A_i taken over all permissible (j,δ) split A_i into at most $4n$ pieces. Do the same for all A_i from our vector. We will get a table of the form

$$\begin{cases} A_1 \overline{\circ} W_1(\bar{U}), \\ \quad \cdots \\ A_n \overline{\circ} W_n(\bar{U}), \end{cases} \tag{12}$$

where $\bar{U}$ are pieces of the above defined partition of $A_1, \ldots, A_n$, and all $U_i^{\pm 1}$ are pairwise distinct.

We claim that the representation (12) has the required properties. The only thing to check is c). This readily follows from the main property of Nielsen reduced vectors and the observation

$$\partial(C_{i\epsilon k\nu}) < \partial(C_{i\epsilon j\delta}) \Rightarrow C_{i\epsilon k\nu} \overline{\circ} C_{k(-\nu)j\delta}$$

which says that the two occurrences of $C_{i\epsilon j\delta}$ in (11) are synchronously subdivided in (12).■

In this paper we will actually need only the following application of Theorem 5.4.

Theorem 5.5. *Let a vector $(A_1, \dots, A_n)$ be carried into a Nielsen reduced vector $(B_1, \dots, B_m)$ by a Nielsen transformation. Then for any $i \le m$ there exists a partition*

$$B_i \overline{\circ} B_{i1} \dots B_{il} \tag{13}$$

of the word B_i with $l \le 4m \le 4n$ such that for any $j \le l$ there exists $\nu \le n$ with the property $B_{ij} \subseteq A_\nu^{\pm 1}$.

Proof. We apply Theorem 5.4 to the vector $(B_1, \dots, B_m)$ and take as (13) the resulting representation b).

In order to prove that (13) has the desired property, we reverse elementary Nielsen transformations in the sequence leading from $(A_1, \dots, A_n)$ to $(B_1, \dots, B_m)$ and apply this reversed sequence to the vector $(b_1, \dots, b_m)$. We will find words $H_1(b_1, \dots, b_m), \dots, H_n(b_1, \dots, b_m)$ such that $A_\nu = H_\nu(B_1, \dots, B_m)$. Hence, by the property c) in Theorem 5.4, we have

$$A_\nu \overline{\circ} H'_\nu(\bar{U}), \tag{14}$$

where $H'_\nu(\bar{u}) = H_\nu(W_1(\bar{u}), \dots, W_m(\bar{u}))$. The vector $(W_1(\bar{u}), \dots, W_m(\bar{u}))$ is carried by the reversed sequence of elementary Nielsen transforms into $(H'_1(\bar{u}), \dots, H'_n(\bar{u}))$. Since the set of letters appearing in a vector is invariant under Nielsen transformations, we conclude that $(H'_1(\bar{u}), \dots, H'_n(\bar{u}))$ contains all letters which occur in $(W_1(\bar{u}), \dots, W_m(\bar{u}))$. Now the theorem immediately follows from (14).■

The last topic discussed in this section is a simplified form of one argument due to Bulitko [9].

Fix a simple word P. An occurrence $B * P^s * C,\ s \in \mathbb{Z}$ is *stable* if B ends with P and C starts with P. Proposition 5.1 ensures that any two intersecting stable occurrences into the same word can be joined to get a bigger stable occurrence. Hence, if we display all *maximal* stable occurrences into a word A, we will get the unique representation of A in the form

$$A \overline{\circ} B_1 P^{s_1} B_2 P^{s_2} \dots P^{s_k} B_{k+1}.$$

Given a positive integer s, we define

$$\partial_{P,s}(A) \rightleftharpoons \sum_{i=1}^{k} \operatorname{sign}(s_i) \cdot (|s_i| \dot{-} s),$$

where

$$a \dot{-} b \rightleftharpoons \begin{cases} a - b \text{ if } a \geq b, \\ 0 \text{ if } a < b. \end{cases}$$

Lemma 5.6. **a)** $\partial_{P,s}(A^{-1}) = -\partial_{P,s}(A)$,

b) *if* $A \overset{\circ}{=} BC$ *then* $|\partial_{P,s}(A) - \partial_{P,s}(B) - \partial_{P,s}(C)| \leq s + 4$,

c) *if* $A = A_1 \dots A_n$ *then*

$$\left| \partial_{P,s}(A) - \sum_{i=1}^{n} \partial_{P,s}(A_i) \right| \leq 3(s+4)(n-1).$$

Proof. a) follows from the symmetry of definitions.

b) Let $B \overset{\circ}{=} B_1 B_2$ and $C \overset{\circ}{=} C_1 C_2$, where $\partial(B_2) = \min(2\partial(P), \partial(B))$ and $\partial(C_1) = \min(2\partial(P), \partial(C))$. Then every single-letter occurrence into B_1 or C_2 belongs to a stable occurrence into A if and only if it belongs to a stable occurrence into B [C respectively]. Now the required inequality readily follows from the observation $(x + y + 4) \dot{-} s \leq (x \dot{-} s) + (y \dot{-} s) + s + 4$.

c) Consider first the case $n = 2$. We have

$$A_1 \overset{\circ}{=} BC, \ A_2 \overset{\circ}{=} C^{-1}D, \ A \overset{\circ}{=} BD \tag{15}$$

for some B, C, D. The required inequality $|\partial_{P,s}(A) - \partial_{P,s}(A_1) - \partial_{P,s}(A_2)| \leq 3(s+4)$ follows from the already proven part b) applied thrice to the representation (15).

The general case is proved from the case $n = 2$ by obvious induction on n.■

6. Lower bounds for the hierarchy $\mathcal{H}(d, r)$

In this section we develop machinery which for some explicit words A will allow us to prove facts like "$A \notin \mathcal{H}(d, r)$". In fact we will be able to establish at once that A is "hard" for $\mathcal{H}(d, r)$ in the sense "for *all* X such that $A \subseteq X$, $X \notin \mathcal{H}(d, r)$". To that end we will show that if A contains words $P_1^\beta, \dots, P_{r+1}^\beta$, where $\beta = \beta(r)$ is sufficiently large, $P_1, \dots, P_{r+1}$ are already known to be hard for $\mathcal{H}(d-1, r)$ and moreover are such that $\widetilde{P}_1, \dots, \widetilde{P}_{r+1}$ are linearly independent, then A is hard for $\mathcal{H}(d, r)$. This will allow us in the next section to find out words which are hard for $\mathcal{H}(d, r)$ by induction on d.

For technical reasons we start with a simpler semigroup version of the corresponding statement (Lemma 6.1) from which it will be easy to deduce its group analogue (Lemma 6.2).

Let

$$W(u_1, \dots, u_r) \mathrel{\overline{\circ}=} u_{i_1} u_{i_2} \dots u_{i_n} \tag{16}$$

be a positive word and $A_1, \dots, A_r \in F_2$. We denote by $E(W, \bar{A})$ the linear subspace in $E(F_2)$ spanned by the sums

$$\left\{ \sum_{j=k+1}^{l} \widetilde{A}_{i_j} \mid 1 \le k < l \le n,\ i_k = i_l \right\}.$$

Let

$$\alpha(r) \rightleftharpoons \frac{1}{5}(21 \cdot 16^r - 16) \cdot (r!)^2.$$

Lemma 6.1. *Let $P^{\alpha(r)} \subseteq W(A_1, \dots, A_r)$, where W is a positive word and $P \not\subseteq A_i$ $(1 \le i \le r)$. Then $\widetilde{P} \in E(W, \bar{A})$.*

Proof. We proceed by induction on r. As the case $r = 0$ is trivial, we assume that $r > 0$ and that the lemma is already established for all smaller values of r.

Representing W in the form (16), we have

$$A_{i_1} A_{i_2} \dots A_{i_n} \mathrel{\overline{\circ}=} B P^{\alpha(r)} C.$$

For $1 \le \nu \le \alpha(r) - 2$ split this word into BP^ν and $P^{\alpha(r)-\nu}C$ as follows:

$$A_{i_{f(\nu)}} \mathrel{\overline{\circ}=} A'_\nu A''_\nu, \tag{17}$$

$$A_{i_1} \dots A_{i_{f(\nu)-1}} A'_\nu \mathrel{\overline{\circ}=} B P^\nu, \tag{18}$$

$$A''_\nu A_{i_{f(\nu)+1}} \dots A_{i_n} \mathrel{\overline{\circ}=} P^{\alpha(r)-\nu} C, \tag{19}$$

where $f(\nu)$ is a non-decreasing function. In fact, the assumption $P \not\subseteq A_i$ $(1 \le i \le r)$ implies that $f(\nu)$ is *strictly* increasing.

As $i_{f(\nu)}$ takes on at most r values, we can find $\Gamma \subseteq \{1, 2, \dots, \alpha(r) - 2\}$ with $|\Gamma| \ge \frac{\alpha(r)-2}{r}$ such that $i_{f(\nu)} = \text{const},\ \nu \in \Gamma$. W.l.o.g. assume that $i_{f(\nu)} = 1$ for all $\nu \in \Gamma$.

Note for the record that for every $\nu_1, \nu_2 \in \Gamma$ with $\nu_1 < \nu_2$ (19) implies

$$(\nu_2-\nu_1)\widetilde{P}+\left(\widetilde{A}''_{\nu_2}-\widetilde{A}''_{\nu_1}\right)=\sum_{j=f(\nu_1)+1}^{f(\nu_2)}\widetilde{A}_{i_j}\in E(W,\bar{A}). \tag{20}$$

If, in particular, $A''_{\nu_1}\overline{\circ}\,A''_{\nu_2}$ for two different $\nu_1,\nu_2\in\Gamma$ then we are done. Hence we may assume that all $A''_\nu,\ \nu\in\Gamma$ are pairwise distinct.

Either $\partial\left(A''_\nu\right)\le\frac{1}{2}\partial\left(A_1\right)$ or $\partial\left(A''_\nu\right)\ge\frac{1}{2}\partial\left(A_1\right)$ for at least half of all $\nu\in\Gamma$. Hence we may find $\Gamma_0\subseteq\Gamma$ with $|\Gamma_0|\ge\frac{\alpha(r)-2}{2r}$ and with the additional property

$$\left|\partial\left(A''_{\nu_1}\right)-\partial\left(A''_{\nu_2}\right)\right|\le\frac{1}{2}\partial\left(A_1\right);\ \nu_1,\nu_2\in\Gamma_0. \tag{21}$$

Let $\nu_0\in\Gamma_0$ be such that $\partial\left(A''_{\nu_0}\right)$ is minimal. We will abbreviate A'_{ν_0} to A' and A''_{ν_0} to A''.

For any other $\nu\in\Gamma_0$ we now have

$$A''_\nu\overline{\circ}\,A''D_\nu,\ A'\overline{\circ}\,E_\nu A'_\nu,\ D_\nu\neq\Lambda$$

since in view of (19), (18), A'', A''_ν are initial segments of P, and A', A'_ν are terminal segments of P. Substituting these equalities to (17), we find

$$A_1\overline{\circ}\,E_\nu(A'_\nu A'')\overline{\circ}\,(A'_\nu A'')D_\nu. \tag{22}$$

Hence there exists a simple word Q_ν and a positive integer $g(\nu)$ such that

$$D_\nu\overline{\circ}\,Q_\nu^{g(\nu)},\ Q_\nu\overline{\circ}\,Q'_\nu Q''_\nu,\ (A'_\nu A'')\overline{\circ}\,Q''_\nu Q_\nu^{s(\nu)},\ E_\nu\overline{\circ}\,(Q''_\nu Q'_\nu)^{g(\nu)}$$

for some $Q'_\nu, Q''_\nu, s(\nu)$.

Now, (21) implies $\partial\left(Q_\nu\right)\le\partial\left(D_\nu\right)\le\frac{1}{2}\partial\left(A_1\right)$. Hence for every $\nu_1,\nu_2\in\Gamma_0$ we may apply to (22) Proposition 5.1 to conclude that in fact $Q_{\nu_1}\overline{\circ}\,Q_{\nu_2}$. We will abbreviate the common value of Q_ν $(\nu\in\Gamma_0)$ to Q. In view of (20), we are only left to show that $\widetilde{Q}\in E(W,\bar{A})$. Note for the record that $g(\nu)\cdot\partial\left(Q\right)=\partial\left(D_\nu\right)\le\partial\left(A''_\nu\right)\le\partial\left(P\right)$, hence

$$g(\nu)\le\frac{\partial\left(P\right)}{\partial\left(Q\right)},\ \nu\in\Gamma_0. \tag{23}$$

Also, $g(\nu_0)=0$, and all $g(\nu),\ \nu\in\Gamma_0$ are pairwise different since all A''_ν are so.

Let us consider now the two infinite periodic words $A''Q^\infty$ and P^∞.

Case 1. $A''Q^\infty\overline{\circ}\,P^\infty$.

Then, as Q is simple, $P \overset{\circ}{=} Q_1^t$ with $Q_1 \sim Q$. Choose $\nu \in \Gamma_0$ so that $|\nu - \nu_0| \geq 2$. Applying (20) to the pair ν_0, ν, we conclude $(\nu - \nu_0)\widetilde{P} + g(\nu)\widetilde{Q} \in E(W, \bar{A})$ or $\widetilde{Q}(t(\nu - \nu_0) + g(\nu)) \in E(W, \bar{A})$. As $g(\nu) \leq t$ by (23), $t(\nu - \nu_0) + g(\nu) \neq 0$ which implies $\widetilde{Q} \in E(W, \bar{A})$ (and hence also $\widetilde{P} \in E(W, \bar{A})$).

Case 2. *$A''Q^\infty$ and P^∞ are not graphically equal.*

Let $A''Q^tQ'$ be their maximal common initial segment, where $Q \overset{\circ}{=} Q'Q''$, $Q'' \neq \Lambda$. In other words, P^∞ begins with $A''Q^tQ'u$, where u is *not* the initial letter of Q''. $\partial(P) \geq \partial(Q)$ by (23) hence, by Proposition 5.1, $\partial(A''Q^tQ') < 2\partial(P)$. This implies that $A''Q^tQ'u$ is an initial segment of P^2 and hence of the word (19) for each $\nu \in \Gamma_0$. Moreover, $\partial(A''_\nu) \leq \partial(A''Q^tQ')$ since A''_ν is an initial segment of both $A''Q^\infty$ and P. This implies $g(\nu) \leq t$ and that $Q^{t-g(\nu)}Q'u$ is an initial segment of the word $A_{i_{f(\nu)+1}} \dots A_{i_n}$ for all $\nu \in \Gamma_0$. That is to say

$$A_{i_{h(\nu)}} \overset{\circ}{=} \bar{A}'_\nu \bar{A}''_\nu, \ \bar{A}''_\nu \neq \Lambda, \tag{24}$$

$$Q^{t-g(\nu)}Q' \overset{\circ}{=} A_{i_{f(\nu)+1}} \dots A_{i_{h(\nu)-1}} \bar{A}'_\nu, \tag{25}$$

$$\bar{A}''_\nu \text{ begins with } u \tag{26}$$

for some function $h(\nu)$.

Choose $\Gamma_1 \subseteq \Gamma_0$ such that $|\Gamma_1| \geq \frac{1}{r}|\Gamma_0| \geq \frac{\alpha(r)-2}{2r^2}$ and $i_{h_\nu} = \text{const}$ for $\nu \in \Gamma_1$. Note that if $i_{f_1} = i_{f_2}$ and $i_{h_1} = i_{h_2}$ then

$$\sum_{j=f_1+1}^{h_1-1} \widetilde{A}_{i_j} - \sum_{j=f_2+1}^{h_2-1} \widetilde{A}_{i_j} = \left(\sum_{j=f_1+1}^{n} \widetilde{A}_{i_j} - \sum_{j=f_2+1}^{n} \widetilde{A}_{i_j} \right) - \left(\sum_{j=h_1}^{n} \widetilde{A}_{i_j} - \sum_{j=h_2}^{n} \widetilde{A}_{i_j} \right) \in E(W, \bar{A}).$$

From this and (25) we conclude

$$(g(\nu_2) - g(\nu_1))\widetilde{Q} + \left(\widetilde{\bar{A}'_{\nu_2}} - \widetilde{\bar{A}'_{\nu_1}} \right) \in E(W, \bar{A}); \ \nu_1, \nu_2 \in \Gamma_1. \tag{27}$$

Now, suppose for a moment that for two different $\nu_1, \nu_2 \in \Gamma_1$ we have $\partial\left(\bar{A}'_{\nu_1}\right) \geq 2\partial(Q)$, $\partial\left(\bar{A}'_{\nu_2}\right) \geq 2\partial(Q)$. Then (25), (24) along with Proposition 5.1 would imply $\bar{A}'_{\nu_1} \overset{\circ}{=} Q_1 Q^{r_1} Q'$ and $\bar{A}'_{\nu_2} \overset{\circ}{=} Q_1 Q^{r_2} Q'$, where Q_1 is an initial segment of Q. This implies that, unless $r_1 = r_2$, the longer of the two words $\bar{A}''_{\nu_1}$, $\bar{A}''_{\nu_2}$ begins with Q'' which contradicts (26). In the remaining case $r_1 = r_2$ we have $\bar{A}'_{\nu_1} \overset{\circ}{=} \bar{A}'_{\nu_2}$ which, by (27), proves the required inclusion $\widetilde{Q} \in E(W, \bar{A})$.

Hence we may assume that $\partial\left(\bar{A}'_\nu\right) \geq 2\partial(Q)$ for at most one $\nu \in \Gamma_1$. Removing if necessary from Γ_1 that particular ν, we come up with $\Gamma_2 \subseteq \Gamma_1$ such that $|\Gamma_2| \geq \frac{\alpha(r)-2r^2-2}{2r^2}$ and $\partial\left(\bar{A}'_\nu\right) \leq 2\partial(Q),\ \nu \in \Gamma_2$.

Let us call A_i *large* if $\partial(A_i) \geq 2\partial(Q)$ and *small* otherwise. Let r' be the number of small A_is. A_1 is large hence $r' < r$. Consider two cases.

Case 2.1. *For some $\nu \in \Gamma_2$ and some large A_i, i appears at least twice in the multiset $\left\{i_{f(\nu)+1}, \ldots, i_{h(\nu)-1}\right\}$.*

If, say, $i_a = i_b = i$ $(f(\nu)+1 \leq a < b \leq h(\nu)-1)$ then $A_{i_a} \ldots A_{i_b}$ is a Q-periodic word which implies $\widetilde{Q} = \gamma \cdot \sum_{j=a+1}^{b} \widetilde{A}_{i_j} \in E(W, \bar{A})$ for some $\gamma > 0$.

Case 2.2. *The assumption of Case* 2.1 *does not take place.*

Then for each $\nu \in \Gamma_2$ we set

$$I(\nu) \rightleftharpoons \left\{ i \,\middle|\, A_i \text{ is large and appears in } A_{i_{f(\nu)+1}}, \ldots, A_{i_{h(\nu)-1}} \right\}.$$

$I(\nu)$ takes on at most $2^{r-r'}$ values hence for some $I \subseteq \{i \mid A_i \text{ is large}\}$ we may find $\Gamma_3 \subseteq \Gamma_2$ with $|\Gamma_3| \geq \frac{\alpha(r)-2r^2-2}{2r^2 \cdot 2^{r-r'}} \geq \frac{\alpha(r)}{2r^2 \cdot 2^{r-r'}} - 1$ such that $I(\nu) = I$ for $\nu \in \Gamma_3$. Let $\nu_1, \nu_2 \in \Gamma_3$ be determined by $g(\nu_1) = \min_{\nu \in \Gamma_3} g(\nu),\ g(\nu_2) = \max_{\nu \in \Gamma_3} g(\nu)$. Then

$$g(\nu_2) - g(\nu_1) \geq |\Gamma_3| - 1 \geq \frac{\alpha(r)}{2r^2 \cdot 2^{r-r'}} - 2. \tag{28}$$

Comparing lengths in (25), we will find the following two inequalities:

$$\sum_{i \in I} \partial(A_i) \leq \partial(Q)\,(t - g(\nu_2)) + \partial(Q') \tag{29}$$

and

$$\sum_{i \in I} \partial(A_i) + \sum_{\substack{j \in [f(\nu_1)+1, h(\nu_1)-1] \\ A_{i_j} \text{ small}}} \partial\left(A_{i_j}\right) + 2\partial(Q) \geq \partial(Q)\,(t - g(\nu_1)) + \partial(Q'). \tag{30}$$

From (28), (29) and (30) we get

$$\sum_{\substack{j \in [f(\nu_1)+1, h(\nu_1)-1] \\ A_{i_j} \text{ small}}} \partial\left(A_{i_j}\right) \geq \partial(Q)\left(\frac{\alpha(r)}{2r^2 \cdot 2^{r-r'}} - 4\right). \tag{31}$$

We have only $|I| \leq r - r'$ large occurrences into $A_{f(\nu_1)+1} \ldots A_{h(\nu_1)-1}$ which split this word into at most $(r - r' + 1)$ pieces consisting entirely of small A_is. Hence we have an

interval $[a,b] \subseteq [f(\nu_1)+1, h(\nu_1)-1]$ such that all $A_{i_j},\ j \in [a,b]$ are small and

$$\sum_{j=a}^{b} \partial\left(A_{i_j}\right) \geq \partial(Q)\left(\frac{\alpha(r)}{2r^2(r-r'+1)\cdot 2^{r-r'}} - 2\right).$$

This implies that $Q^{\alpha'} \subseteq A_{i_a} \ldots A_{i_b}$, where $\alpha' \geq \frac{\alpha(r)}{2r^2(r-r'+1)\cdot 2^{r-r'}} - 4$.

Now we are in position to apply the inductive assumption. Namely, we let $P' \rightleftharpoons Q^2$ and $W' \rightleftharpoons u_{i_a} \ldots u_{i_b}$. Then $(P')^{\lfloor \alpha'/2 \rfloor} \subseteq W'(\bar{A})$ and it is easy to check that $\lfloor \alpha'/2 \rfloor \geq \frac{\alpha(r)}{4r^2(r-r'+1)\cdot 2^{r-r'}} - 3 \geq \alpha(r')$. $P' \not\subseteq A_i$ for a small A_i merely because $\partial(A_i) < 2\partial(Q) = \partial(P')$.

All premises of our lemma are fulfilled for $r := r'$. Hence we conclude $\widetilde{P}' \in E(W', \bar{A}) \subseteq E(W, \bar{A})$ which also shows $\widetilde{Q} \in E(W, \bar{A})$ and completes the proof of Lemma 6.1.■

Now it is comparatively easy to establish the analogue of this lemma for the case of groups. Let

$$\beta(r) \rightleftharpoons 17r \cdot 16^{8r^3} \cdot \left\{(8r^3)!\right\}^2.$$

Lemma 6.2. *Let* $A \in \mathrm{Gp}(A_1, \ldots, A_r)$, $P^{\beta(r)} \subseteq A$ *and* $P \not\subseteq A_i^{\pm 1}$ $(1 \leq i \leq r)$. *Then* $\widetilde{P} \in \mathrm{Span}(\widetilde{A}_1, \ldots, \widetilde{A}_r)$.

Proof. Carry the vector $(A_1, \ldots, A_r)$ into a Nielsen reduced vector $(B_1, \ldots, B_s)$ $(s \leq r)$ by a Nielsen transformation. Then $A = W(B_1, \ldots, B_s)$, where $W \overline{\circ} u_{i_1}^{\epsilon_1} \ldots u_{i_n}^{\epsilon_n}$.

For every triple (p_1, p_2, p_3), where $p_2 \in \left\{u_1^{\pm 1}, \ldots, u_s^{\pm 1}\right\}$; $p_1, p_3 \in \left\{u_1^{\pm 1}, \ldots, u_s^{\pm 1}, 1\right\}$ and $p_1 \neq p_2^{-1}$, $p_3 \neq p_2^{-1}$, we introduce a new variable $v_{p_1p_2p_3}$. Note that altogether we have $(2s)^3 \leq 8r^3$ v-variables.

Now we transform $W(\bar{u})$ to another word $W'(\bar{v})$ by replacing in W all occurrences $u_{i_j}^{\epsilon_j}$ with $v_{i_j} \rightleftharpoons v_{p_1p_2p_3}$, where $(p_1, p_2, p_3) \rightleftharpoons \left(u_{i_{j-1}}^{\epsilon_{j-1}}, u_{i_j}^{\epsilon_j}, u_{i_{j+1}}^{\epsilon_{j+1}}\right)$. The main property of Nielsen reduced vectors then says that

$$A \overline{\circ} W'(\bar{C}),$$

where for $(p_1, p_2, p_3) = \left(u_{\nu_1}^{\epsilon_1}, u_{\nu_2}^{\epsilon_2}, u_{\nu_3}^{\epsilon_3}\right)$, $C_{p_1p_2p_3}$ is the part remaining from the middle syllable $B_{\nu_2}^{\epsilon_2}$ in the product $B_{\nu_1}^{\epsilon_1} B_{\nu_2}^{\epsilon_2} B_{\nu_3}^{\epsilon_3}$ after performing cancellations.

Now we are going to apply Lemma 6.1. Namely, we set in its statement $P := P^{4r}$, $r := 8r^3$ and $W(\bar{A}) := W'(\bar{C})$. Let us check its premises.

Note that $\beta(r) \geq 4r \cdot \alpha(8r^3)$.

To check the condition $P^{4r} \not\subseteq C_{p_1p_2p_3}$, it suffices to see that $P^{4r} \not\subseteq B_j^{\pm 1}$ $(1 \leq j \leq s)$. Indeed, if $P^{4r} \subseteq B_j^{\pm 1}$ then we would apply Theorem 5.5 and find that at least one $A_i^{\pm 1}$ would contain P. This would contradict the assumption.

So, all premises of Lemma 6.1 are fulfilled and we conclude that $\widetilde{P} \in E(W', \bar{C})$. In the case $\left(u_{i_{k-1}}^{\epsilon_{k-1}}, u_{i_k}^{\epsilon_k}, u_{i_{k+1}}^{\epsilon_{k+1}}\right) = \left(u_{i_{l-1}}^{\epsilon_{l-1}}, u_{i_l}^{\epsilon_l}, u_{i_{l+1}}^{\epsilon_{l+1}}\right)$ we have, however, $\sum_{j=k+1}^{l} \widetilde{C}_{i_j} = \sum_{j=k+1}^{l} \widetilde{B}_{i_j}^{\epsilon_j}$. Hence $E(W', \bar{C}) \subseteq \mathrm{Span}(\widetilde{B}_1, \ldots, \widetilde{B}_s) = \mathrm{Span}(\widetilde{A}_1, \ldots, \widetilde{A}_r)$ which completes the proof of Lemma 6.2.■

To make use of Lemma 6.2 in the next section, we will have to be able to establish facts like $A \subseteq B$ or $\partial(A) \leq \partial(B)$ from only a very indirect information about words A and B. In the rest of this section we prove two lemmas to that end.

Lemma 6.3. *Let $A, B, C_1, \ldots, C_k, \ldots, C_l, \ldots, C_n$ be words and s be an integer such that $\partial\left(\widehat{B}\right) \geq \frac{1}{s}\partial\left(\widehat{C}_i\right)$ $(1 \leq i \leq n)$. Let $t \geq (s+4)(9n+22)$ and*

$$D = C_1 \ldots C_k A^t C_{k+1} \ldots C_l B^t C_{l+1} \ldots C_n.$$

a) *If $\widetilde{A}$ and $\widetilde{B}$ are linearly independent, then $\partial\left(\widehat{D}\right) \geq s \cdot \partial\left(\widehat{A}\right)$ and $\partial\left(\widehat{D}\right) \geq s \cdot \partial\left(\widehat{B}\right)$;*

b) *if $A = 1$, then $\widehat{B}^s \subseteq \widehat{D}$.*

Proof. We prove both parts of the lemma simultaneously. Obviously, in both cases we may assume w.l.o.g. that

$$\partial\left(\widehat{A}\right) \leq \partial\left(\widehat{B}\right).$$

Let $A \overline{\circ} X^{-1}\widehat{A}X$, $B \overline{\circ} Y^{-1}\widehat{B}Y$, $C_i \overline{\circ} Z_i^{-1}\widehat{C}_i Z_i$ and $D \overline{\circ} T^{-1}\widehat{D}T$. Then

$$\widehat{D} = T\left(\prod_{i=1}^{k} Z_i^{-1}\widehat{C}_i Z_i\right) \cdot X^{-1}\widehat{A}^t X \cdot \left(\prod_{i=k+1}^{l} Z_i^{-1}\widehat{C}_i Z_i\right) \cdot Y^{-1}\widehat{B}^t Y \cdot \left(\prod_{i=l+1}^{n} Z_i^{-1}\widehat{C}_i Z_i\right) T^{-1}. \quad (32)$$

Let $\widehat{B} \overline{\circ} P^h$ where P is simple.

We apply to (32) Lemma 5.6 a), c) with $s := hs$, $n := 3n + 8$. We then have

$$\left|\partial_{P,hs}\left(\widehat{A}^t\right) + \partial_{P,hs}\left(\widehat{B}^t\right) + \sum_{i=1}^{n} \partial_{P,hs}\left(\widehat{C}_i\right) - \partial_{P,hs}\left(\widehat{D}\right)\right| \leq 3(hs+4)(3n+7). \quad (33)$$

Obviously,

$$\partial_{P,hs}\left(\widehat{B}^t\right) = \partial_{P,hs}\left(\widehat{P}^{ht}\right) = h(t-s) - 2 > 3(hs+4)(3n+7). \quad (34)$$

$\partial_{P,hs}\left(\widehat{C}_i\right) = 0$ merely because $\partial\left(\widehat{C}_i\right) \leq hs \cdot \partial(P)$.

We claim that also $\partial_{P,hs}\left(\widehat{A}^t\right) = 0$. This is obvious if $A = 1$ (part b)). For part a) note that in view of Proposition 5.1 and observation $\partial(P) + \partial\left(\widehat{A}\right) \leq \partial(P) + \partial\left(\widehat{B}\right) = (h+1)\partial(P)$, $P^{\pm(hs+1)} \subseteq \widehat{A}^t$ would imply that $\widetilde{P}$ and $\widetilde{A}$ are linearly dependent contrary to the assumption of part a).

Now (33) and (34) imply $\partial_{P,hs}\left(\widehat{D}\right) > 0$ from which we have $\widehat{B}^s \subseteq \widehat{D}$. This inclusion proves both parts of the lemma.■

Lemma 6.4. *Let $\widetilde{A}, \widetilde{B}$ be linearly independent and α be a positive integer. Then:*

a) $\partial\left(\widehat{A^2B^\alpha}\right) \geq (\alpha - 5) \cdot \partial\left(\widehat{B}\right)$,

b) $\partial\left(\widehat{A^2B^\alpha}\right) \geq \left(1 - \frac{5}{\alpha}\right) \cdot \partial\left(\widehat{A}\right)$.

Proof. Replacing if necessary the pair (A, B) with a conjugate, we may assume that

$$\partial(A) + \partial(B) \leq \partial\left(P^{-1}AP\right) + \partial\left(P^{-1}BP\right) \tag{35}$$

for each P. Let $A \overset{\circ}{=} U^{-1}\widehat{A}U$, $B \overset{\circ}{=} V^{-1}\widehat{B}V$. We apply case analysis.

Case 1. $U \neq 1$ *or* $V \neq 1$.

In this case the product $U^{-1}\widehat{A}^2UV^{-1}\widehat{B}^\alpha V$ is cyclically reduced (since otherwise $\partial(A) + \partial(B)$ could be decreased contrary to (35)) and the statement becomes obvious.

Case 2. $A \overset{\circ}{=} \widehat{A}$, $B \overset{\circ}{=} \widehat{B}$.

As $\widetilde{A}$ and $\widetilde{B}$ are linearly independent, the infinite words $B^{-\infty}$ and A^∞ can not coincide. Let P be their maximal common initial segment. Conjugating A and B by P, we may additionally assume that BA is reduced. Let $C \rightleftharpoons \widehat{A^2B^\alpha}$.

Case 2.1. B^α *completely cancels in the product* A^2B^α.

This means $A^2 \overset{\circ}{=} W^{-1}CWB^{-\alpha}$ for some W.

Case 2.1.1. $C \overset{\circ}{=} C_1C_2$, $A \overset{\circ}{=} W^{-1}C_1 \overset{\circ}{=} C_2WB^{-\alpha}$.

$\partial(C) \geq \partial(C_1) \geq \alpha \cdot \partial(B)$ which proves part a). Also, as in the proof of Proposition 5.3 we have $\partial(W) \leq \partial(C_2)$, hence $\partial(C) = \partial(C_1) + \partial(C_2) \geq \partial(C_1) + \partial(W) = \partial(A)$ which proves b).

Case 2.1.2. $W \overset{\circ}{=} W_1W_2$, $W_2 \neq \Lambda$, $A \overset{\circ}{=} W^{-1}CW_1 \overset{\circ}{=} W_2B^{-\alpha}$.

This case is actually impossible since W_2 and W_2^{-1} can not both be initial segments of A.

Case 2.1.3. $B^{-\alpha} \doteq B_1 A,\ A \doteq W^{-1}CWB_1$.

We have $B^{-\alpha} \doteq B_1 W^{-1}CWB_1$. Proposition 5.3 implies $\partial(W) \leq \frac{1}{2}\partial(B)$. Also $\partial(B_1) \leq 2\partial(B)$ since otherwise the standard argument would show $\widetilde{A} = \widetilde{C} + \widetilde{B}_1 = \gamma\widetilde{B}$ for some $\gamma \in \mathbb{Q}$ contrary to the assumption that $\widetilde{A}$ and $\widetilde{B}$ are linearly independent. These two facts imply $\partial(C) \geq (\alpha - 5)\partial(B)$. On the other hand, $\alpha \cdot \partial(B) \geq \partial(A)$. Hence also $\partial(C) \geq \left(1 - \frac{5}{\alpha}\right)\partial(A)$.

Case 2.2. *A^2 completely cancels in the product A^2B^α.*

In other words, $B^\alpha \doteq A^{-2}W^{-1}CW$. By Proposition 5.1, $\partial(A^2) \leq \partial(A) + \partial(B)$ that is $\partial(A) \leq \partial(B)$. By Proposition 5.3, $\partial(W) \leq \frac{1}{2}\partial(B)$. Hence $\partial(C) \geq (\alpha - 3)\partial(B)$ and so $\partial(C) \geq (\alpha - 3)\partial(A)$.

Case 2.3. *Neither A^2 nor B^α completely cancel in the product A^2B^α.*

Along with the above mentioned assumption that BA is reduced, this gives us

$$A^2 \doteq UV^{-1},\ B^\alpha \doteq VW,\ C \doteq UW.$$

Once again, by Proposition 5.1 we have $\partial(V) \leq \partial(A) + \partial(B) = \frac{1}{2}\partial(U) + \left(\frac{1}{2} + \frac{1}{\alpha}\right)\partial(V) + \frac{1}{\alpha}\partial(W)$. Therefore $\partial(V) \leq \frac{\alpha}{\alpha-2}\partial(U) + \frac{2}{\alpha-2}\partial(W)$. Now $\partial(C) = \partial(U) + \partial(W) \geq \frac{\alpha-2}{\alpha}(\partial(V) + \partial(W)) = (\alpha - 2)\partial(B)$ and $\partial(C) \geq \frac{\alpha-2}{2\alpha-2}(\partial(U) + \partial(V)) = \left(1 - \frac{1}{\alpha-1}\right) \cdot \partial(A) \geq \left(1 - \frac{5}{\alpha}\right)\partial(A)$.■

7. Proof of the main lemma

After all this tedious preliminary work it is comparatively easy to finish the proof of the main lemma which we do in this section. We will be less careful about explicit bounds on the constants involved in the construction than in the previous sections. Still the proof is absolutely constructive and these bounds can be extracted from it if desired.

For typographical reasons we will be denoting by $\boldsymbol{A}$ the value $\pi(A)$ for a flat homomorphism $\pi : \langle a_1, \ldots, a_g\rangle \longrightarrow F_2$, the latter will be always clear from the context.

The construction is naturally subdivided into two stages. At the first stage we achieve the following.

Lemma 7.1. *There exist a positive solution $\bar{X}$ of the equation* (9) *and a basis $\mathcal{B}$ of the free group $\langle a_1, \ldots, a_g\rangle$ with the following properties:*

a) $\mathrm{Gp}(\bar{X}) = \langle a_1, \ldots, a_g\rangle$,

b) *for any flat homomorphism* $\pi : \langle a_1, \ldots, a_g \rangle \longrightarrow F_2$ *we have* $\partial\left(\widehat{\boldsymbol{X_2 X_1}}\right) \geq 2$ *and, moreover, for any* $B \in \mathcal{B}$,

$$\partial\left(\widehat{\boldsymbol{X_2 X_1}}\right) \geq \frac{1}{2}\partial\left(\widehat{\boldsymbol{B}}\right).$$

Proof. We define the following sequence $\bar{X}^{(0)}, \bar{X}^{(1)}, \ldots, \bar{X}^{(2g-3)}$ of positive solutions of the equation (9):

$$\begin{aligned}
X_i^{(0)} &\rightleftharpoons \begin{cases} a_i, \ 1 \leq i \leq g, \\ a_{2g+1-i}, \ g+1 \leq i \leq 2g; \end{cases} \\
X_1^{(q+1)} &\rightleftharpoons \left(A^{(q)}\right)^{155} X_2^{(q)}; \\
X_i^{(q+1)} &\rightleftharpoons X_{i+1}^{(q)}; \ (2 \leq i \leq 2g-1) \\
X_{2g}^{(q+1)} &\rightleftharpoons \left(A^{(q)}\right)^{154} X_2^{(q)},
\end{aligned}$$

where

$$A^{(q)} \rightleftharpoons X_2^{(q)} X_1^{(q)},$$

borrowed from [13]. In order to see that $\bar{X}^{(q)}$ are indeed solutions of the equation (9), we observe that for any integer α the automorphism

$$x_1 \to (x_2 x_1)^{\alpha} x_2, \ x_i \to x_{i+1} \ (2 \leq i \leq 2g-1), \ x_{2g} \to (x_2 x_1)^{\alpha-1} x_2$$

of the free group $\langle x_1, \ldots, x_{2g} \rangle$ preserves the cyclic word $x_1 x_2 \ldots x_{2g} x_1^{-1} x_2^{-1} \ldots x_{2g}^{-1}$ and apply an obvious induction on q. For the record note also that

$$X_i^{(q)} \underline{\circ} \begin{cases} a_{i+q} \text{ if } i \geq 2 \text{ and } i+q \leq g, \\ a_{2g+1-i-q} \text{ if } i \geq 2 \text{ and } g+1 \leq i+q \leq 2g. \end{cases}$$

The desired solution $\bar{X}$ will be $\bar{X}^{(2g-3)}$, and the basis $\mathcal{B}$ will be

$$\left\{a_g, A^{(0)}, A^{(1)}, \ldots, A^{(g-2)}\right\}.$$

Let us first check that $\mathcal{B}$ is indeed a basis. The elements $a_1, \ldots, a_{g-1}$ are recursively expressed in terms of $\mathcal{B}$ as

$$a_i = \left(A^{(i-2)}\right)^{-155} a_{i+1}^{-1} A^{(i-1)}, \ (1 \leq i \leq g-1) \tag{36}$$

where we let $A^{(-1)} \rightleftharpoons 1$. Hence $\mathrm{Gp}(\mathcal{B}) = \langle a_1, \ldots, a_g \rangle$, and $\mathcal{B}$ is a basis e.g. by [7, Proposition 2.7].

Part a) of the lemma follows from the more general fact $\mathrm{Gp}(\bar{X}^{(q)}) = \langle a_1, \ldots, a_g \rangle$ which is readily proved by induction on q.

For part b), fix a flat homomorphism $\pi : \langle a_1, \ldots, a_g \rangle \longrightarrow F_2$. The first link in the chain of facts to be pulled out is the observation

$$A^{(g-1)} = a_g \left(A^{(g-2)} \right)^{155} a_g. \tag{37}$$

a_g and $A^{(g-2)}$ are members of $\mathcal{B}$, hence $\widetilde{a}_g$ and $\widetilde{A}^{(g-2)}$ are linearly independent. Since π is flat, the same holds for $\widetilde{\boldsymbol{a_g}}$ and $\widetilde{\boldsymbol{A^{(g-2)}}}$. Applying to the words $\boldsymbol{a_g}, \boldsymbol{A^{(g-2)}}$ Lemma 6.4, we find

$$\partial \left(\widehat{\boldsymbol{A^{(g-1)}}} \right) \geq \frac{30}{31} \cdot \partial \left(\widehat{\boldsymbol{a_g}} \right) \tag{38}$$

and

$$\partial \left(\widehat{\boldsymbol{A^{(g-1)}}} \right) \geq 150 \cdot \partial \left(\widehat{\boldsymbol{A^{(g-2)}}} \right). \tag{39}$$

Now we are going to show by induction on $q = g-1, \ldots, 2g-3$ that

$$\partial \left(\widehat{\boldsymbol{A^{(q)}}} \right) \geq \partial \left(\widehat{\boldsymbol{A^{(q-1)}}} \right) \geq \ldots \geq \partial \left(\widehat{\boldsymbol{A^{(g-1)}}} \right) \tag{40}$$

and

$$\partial \left(\widehat{\boldsymbol{A^{(q)}}} \right) \geq \partial \left(\widehat{\boldsymbol{A^{(2g-q-3)}}} \right). \tag{41}$$

Note that part b) of our lemma is implied by (40), (39), (38) and (41).

Base $q = g-1$ follows from (39).

Inductive step. Let $g-1 \leq q \leq 2g-4$ and suppose that we already have (40) and (41). We need to deduce $\partial \left(\widehat{\boldsymbol{A^{(q+1)}}} \right) \geq \partial \left(\widehat{\boldsymbol{A^{(q)}}} \right)$ and $\partial \left(\widehat{\boldsymbol{A^{(q+1)}}} \right) \geq \partial \left(\widehat{\boldsymbol{A^{(2g-q-4)}}} \right)$.

Using (36) with $i := 2g-q-2$, we have

$$\left.\begin{aligned} A^{(q+1)} &= X_2^{(q+1)} X_1^{(q+1)} = a_{2g-q-2} \left(A^{(q)} \right)^{155} a_{2g-q-1} \\ &= \left(A^{(2g-q-4)} \right)^{-155} a_{2g-q-1}^{-1} A^{(2g-q-3)} \left(A^{(q)} \right)^{155} a_{2g-q-1} \\ &\sim a_{2g-q-1} \left(A^{(2g-q-4)} \right)^{-155} a_{2g-q-1}^{-1} A^{(2g-q-3)} \left(A^{(q)} \right)^{155}. \end{aligned}\right\} \tag{42}$$

Now we are going to apply Lemma 6.3 a) with:

$$
\begin{aligned}
&A := \boldsymbol{a_{2g-q-1}} \left(\boldsymbol{A^{(2g-q-4)}}\right)^{-1} \boldsymbol{a_{2g-q-1}^{-1}},\\
&B := \boldsymbol{A^{(q)}},\\
&C_1 := \boldsymbol{A^{(2g-q-3)}},\\
&n := 1,\ k := 0,\ l := 1,\\
&s := 1,\ t := 155.
\end{aligned}
$$

To check the premises of Lemma 6.3 a), note that $\widetilde{A}^{(2g-q-4)}$ and $\widetilde{A}^{(q)}$ are linearly independent. Indeed, expanding them in terms of the linear basis $\widetilde{\mathcal{B}}$, we see that $\widetilde{A}^{(2g-q-4)} \in \widetilde{\mathcal{B}}$, whereas $\widetilde{A}^{(q)}$'s coefficient in front of $\widetilde{a}_g$ is non-zero. The latter fact is easily proved by induction on q using (37) and (42). Hence, as π is flat, $\widetilde{\boldsymbol{A^{(2g-q-4)}}}$ and $\widetilde{\boldsymbol{A^{(q)}}}$ are also linearly independent. Lastly, the premise $\partial\left(\widehat{\boldsymbol{A^{(2g-q-3)}}}\right) \leq \partial\left(\widehat{\boldsymbol{A^{(q)}}}\right)$ is exactly the inductive assumption (41).

Applying Lemma 6.3 a) gives us (40) and (41) with $q+1$ instead of q. This completes the inductive step.

So, we have (40) and (41), and, as we mentioned above, they imply the statement of Lemma 7.1.■

Now we finish the proof of Lemma 4.2. Let $\bar{X}^{(0)}$ and $\mathcal{B}$ be a positive solution of the equation (9) and a basis of $\langle a_1, \ldots, a_g \rangle$ respectively with the properties described in the statement of Lemma 7.1. We define the sequence $\bar{X}^{(q)}$ based upon $\bar{X}^{(0)}$ analogously to the sequence used in the proof of Lemma 7.1:

$$
\begin{aligned}
X_1^{(q+1)} &\rightleftharpoons \left(A^{(q)}\right)^{\alpha_q} X_2^{(q)};\\
X_i^{(q+1)} &\rightleftharpoons X_{i+1}^{(q)};\ (2 \leq i \leq 2g-1)\\
X_{2g}^{(q+1)} &\rightleftharpoons \left(A^{(q)}\right)^{\alpha_q - 1} X_2^{(q)},
\end{aligned}
$$

where again

$$A^{(q)} \rightleftharpoons X_2^{(q)} X_1^{(q)}.$$

The only difference is that this time $\{\alpha_q\}$ is a rapidly growing sequence of integers defined recursively (along with an auxiliary sequence $\{n_q\}$) by

$$
\begin{aligned}
&n_q \text{ is the sum of lengths of } X_2^{(q)} \text{ and } X_3^{(q)} \text{ with respect to the basis } \mathcal{B},\\
&\alpha_q \rightleftharpoons (\beta(g-1)+4)(9n_q+22).
\end{aligned}
$$

Here $\beta(r)$ is the function from Lemma 6.2.

Let $d \geq 0$. We will show that the solution $X^{(2gd)}$ has the properties required in Lemma 4.2.

Once again, $\mathrm{Gp}(X^{(q)}) = \langle a_1, \ldots, a_g \rangle$ is easily proved by induction on q. In particular,

$$\mathrm{Gp}(X^{(2gd)}) = \langle a_1, \ldots, a_g \rangle. \tag{43}$$

Fix a flat homomorphism $\pi : \langle a_1, \ldots, a_g \rangle \longrightarrow F_2$. We are left to show that $\boldsymbol{A^{(2gd)}} \notin \mathcal{H}(d, g-1)$. We will actually prove slightly more than this, namely

$$\forall X \left(\widehat{\boldsymbol{A^{(2gd)}}} \subseteq X \Rightarrow X \notin \mathcal{H}(d, g-1) \right). \tag{44}$$

As a first step toward this goal we establish by induction on q the following fact:

$$\widehat{\boldsymbol{A^{(q)}}}^{\beta(g-1)} \subseteq \widehat{\boldsymbol{A^{(q+1)}}}. \tag{45}$$

Indeed, assume that we already have (45) for all smaller values of q. Then

$$A^{(q+1)} = X_2^{(q+1)} \left(A^{(q)}\right)^{\alpha_q} X_2^{(q)} = X_3^{(q)} \left(A^{(q)}\right)^{\alpha_q} X_2^{(q)} = B_1^{\pm 1} \ldots B_l^{\pm 1} \left(A^{(q)}\right)^{\alpha_q} B_{l+1}^{\pm 1} \ldots B_{n_q}^{\pm 1},$$

where $B_1, \ldots, B_{n_q} \in \mathcal{B}$. The inductive assumption clearly implies

$$\partial\left(\widehat{\boldsymbol{A^{(q)}}}\right) \geq \partial\left(\widehat{\boldsymbol{A^{(q-1)}}}\right) \geq \ldots \geq \partial\left(\widehat{\boldsymbol{A^{(0)}}}\right),$$

and

$$\partial\left(\widehat{\boldsymbol{A^{(0)}}}\right) \geq \frac{1}{2}\partial\left(\widehat{\boldsymbol{B_i}}\right)$$

for all i by Lemma 7.1. Hence we may apply Lemma 6.3 b) with $B := \boldsymbol{A^{(q)}}$, $n := n_q$, $C_i := \boldsymbol{B_i}$, $s := \beta(g-1)$ to get (45).

Note that (45) actually implies the stronger fact

$$\widehat{\boldsymbol{A^{(q)}}}^{\beta(g-1)} \subseteq \widehat{\boldsymbol{A^{(q')}}} \tag{46}$$

for all $q' > q$.

Now we prove (44) by induction on d.

Base $d = 0$ follows from the property $\partial\left(\widehat{\boldsymbol{A^{(0)}}}\right) > 1$ (see Lemma 7.1).

Inductive step. Assume that (44) is already proved for some d. Let $\widehat{\boldsymbol{A}^{(2gd)}} \subseteq X$. We need to prove $X \notin \mathcal{H}(d+1, g-1)$. Assume the contrary, that is $X \in \mathrm{Gp}(A_1, \ldots, A_{g-1})$, where $\{A_1, \ldots, A_{g-1}\} \subseteq \mathcal{H}(d, g-1)$.

We apply Lemma 6.2 for every P from the set $\left\{ \widehat{\boldsymbol{A}^{(2gd+q)}} \mid 0 \le q \le 2g-1 \right\}$ with $r := g-1$. Note that for any such P, $\widehat{\boldsymbol{A}^{(2gd)}} \subseteq P$ by (46) and $\widehat{\boldsymbol{A}^{(2gd)}} \not\subseteq A_i^{\pm 1}$ due to $A_i \in \mathcal{H}(d, g-1)$ and the inductive assumption. Hence $P \not\subseteq A_i^{\pm 1}$ for all $1 \le i \le g-1$.

We conclude from Lemma 6.2 that $\widetilde{\boldsymbol{A}^{(2gd)}}, \ldots, \widetilde{\boldsymbol{A}^{(2gd+2g-1)}} \in \mathrm{Span}(\widetilde{A}_1, \ldots, \widetilde{A}_{g-1})$. Hence, to get the desired contradiction from the assumption $X \in \mathcal{H}(d+1, g-1)$, we only have to check that $E = \mathrm{Span}(\widetilde{a}_1, \ldots, \widetilde{a}_g)$, where $E \rightleftharpoons \mathrm{Span}(\widetilde{A}^{(2gd)}, \ldots, \widetilde{A}^{(2gd+2g-1)})$. Due to (43), it suffices to show that $\widetilde{X}_i^{(2gd)} \in E$ for all $1 \le i \le 2g$.

This is the matter of computation. We consequently find in E vectors

$$\widetilde{X}_1^{(2gd)} + \widetilde{X}_2^{(2gd)} = \widetilde{A}_2^{(2gd)},$$

$$\widetilde{X}_i^{(2gd)} + \widetilde{X}_{i+1}^{(2gd)} = \widetilde{X}_2^{(2gd+i-2)} + \widetilde{X}_2^{(2gd+i-1)} = \widetilde{A}^{(2gd+i-1)} - \alpha_{2gd+i-2}\widetilde{A}^{(2gd+i-2)} \ (2 \le i \le 2g-1)$$

and

$$\begin{aligned} \widetilde{X}_2^{(2gd)} + \widetilde{X}_{2g}^{(2gd)} &= \widetilde{X}_{2g}^{(2gd+1)} - \alpha_{2gd}\widetilde{A}^{(2gd)} + \widetilde{X}_{2g}^{(2gd)} = \widetilde{X}_2^{(2gd+g-1)} + \widetilde{X}_2^{(2gd+g-2)} - \alpha_{2gd}\widetilde{A}^{(2gd)} \\ &= A^{(2gd+2g-1)} - \alpha_{2gd+2g-2}\widetilde{A}^{(2gd+2g-2)} - \alpha_{2gd}\widetilde{A}^{(2gd)}. \end{aligned}$$

And these vectors span $\mathrm{Span}(\widetilde{X}_1^{(2gd)}, \ldots, \widetilde{X}_{2g}^{(2gd)})$ and hence span $\mathrm{Span}(\widetilde{a}_1, \ldots, \widetilde{a}_g)$.

We have proved that $E = \mathrm{Span}(\widetilde{a}_1, \ldots, \widetilde{a}_g)$, hence $\dim(\widetilde{\boldsymbol{E}}) = g$ (since π is flat) which contradicts $E \subseteq \mathrm{Span}(\widetilde{A}_1, \ldots, \widetilde{A}_{g-1})$.

This contradiction completes the inductive step in proving (44), the proof of Lemma 4.2 and also of Theorems 3.3 and 4.3.

8. Acknowledgment

I am grateful to Igor G. Lysionok for reading an earlier version of this paper and making several useful suggestions. I also thank Sergey I. Adian for a few historical comments.

References

[1] K. I. Appel. On two variable equations in free groups. *Proc. Amer. Math. Soc.*, 21(1):179–184, 1969.

[2] K. I. Appel. One variable equations in free groups. *Proc. Amer. Math. Soc.*, 19:912–918, 1968.

[3] L. P. Comerford and C. C. Edmunds. Solutions of equations in free groups. In *Group Theory, Proceedings of the Singapore Group Theory Conference held at the National University of Singapore, June* 8-19,1987, pages 347–356, Walter de Gruyter, 1989.

[4] R. I. Grigorchuk and P. F. Kurchanov. On quadratic equations in free groups. *Contemporary Mathematics*, 131:159-171, 1992 (Part 1).

[5] R. C. Lyndon. The equation $a^2b^2 = c^2$ in free groups. *Michigan Math. J.*, 6:89–95, 1959.

[6] R. C. Lyndon. Equations in free groups. *Trans. Amer. Math. Soc.*, 96:445–457, 1960.

[7] R. C. Lyndon and P. E. Shupp. *Combinatorial Group Theory.* Springer-Verlag, New York/Berlin, 1977. рус. пер.: Р. Линдон, П. Шупп, Комбинаторная теория групп, М.: Мир, 1980.

[8] С. И. Адян. Проблема Бернсайда и тождества в группах. Наука, Москва, 1975. S. I. Adian, The Burnside Problem and Identities in Groups, Springer-Verlag, 1979.

[9] В. К. Булитко. Об уравнениях и неравенствах в свободной группе и свободной полугруппе. *Ученые записки мат. кафедр. Тульский гос. пед. ин-т*, 2:242–253, 1970. V. K. Bulitko, On equations and inequalities in a free group and in a free semigroup, *Proceedings of the math. dep. Tula State Institute for Elementary School*, 2(1970), 242-253.

[10] Р. И. Григорчук и П. Ф. Курчанов. Некоторые вопросы теории групп, связанные с геометрией. In Итоги науки и техники. Современные проблемы математики, фундаментальные направления, 58, ВИНИТИ, 1990. R. I. Grigorchuk and P. F. Kurchanov, Some Questions of Group Theory Related to Geometry, to appear in *EMS*, 58.

[11] А. А. Лоренц. О представлении множеств решений систем уравнений с одним неизвестным в свободных группах. *Докл. АН СССР*, 178(2):290–292, 1968. A. A. Lorenc, On the representation of solution sets of systems of equations with one unknown in a free group, *Soviet Math. Dokl.*, 9(1968).

[12] А. А. Лоренц. Решение систем уравнений с одним неизвестным в свободных группах. *Докл. АН СССР*, 148(6):1253–1256, 1963. A. A. Lorenc, The solution of systems of equations in one unknown in free groups, *Soviet Math. Dokl.*, 4(1963), 262-266.

[13] Г. С. Маканин. Бескоэффициентные уравнения с четырьмя неизвестными в свободной полугруппе. In Тезисы 3 Всесоюзной конференции по математической логике, pages 128–130, Новосибирск, 1974. Coefficient-free equations with four unknowns in a free semigroup. In *Proceedings of the 3rd All-Union Conference on Mathematical Logic*, Novosibirsk, 1974, p. 128-130.

[14] Г. С. Маканин. Об одном разрешимом фрагменте элементарной теории свободной группы. In С. И. Адян, editor, Вопросы кибернетики. Сложность вычислений и прикладная математическая логика, pages 103–114, ВИНИТИ, Москва, 1988. G. S. Makanin, On one decidable fragment of the elementary theory of a free group. In *Problems of Cybernetics. Complexity Theory and Applied Mathematical Logic*, ed.: S. I. Adian. Moscow, 1988, pages 103-114.

[15] Г. С. Маканин. Разрешимость универсальной и позитивной теорий свободной группы. *Изв. АН СССР, сер. матем.*, 48(4):735–749, 1984. G. S. Makanin, Decidability of universal and positive theories of a free group, *Math. USSR Izv.*, 1984.

[16] Г.С. Маканин. Уравнения в свободной группе. *Изв. АН СССР, сер. матем.*, 46(6):1199–1273, 1982. G. S. Makanin, Equations in a free group, *Math. USSR Izv.*, 21(1983), 483-546.

[17] А. И. Мальцев. Об уравнении $zxyx^{-1}y^{-1}z^{-1} = aba^{-1}b^{-1}$ в свободной группе. *Алгебра и логика*, 1(5):45–50, 1962. A. I. Malcev, On the equation $zxyx^{-1}y^{-1}z^{-1} = aba^{-1}b^{-1}$ in a free group, *Algebra i Logika*, 1, 45-50 (1962).

[18] Ю. И. Мерзляков. Позитивные формулы на свободных группах. *Алгебра и логика*, 5(4):25–42, 1966. Yu. I. Merzlyakov, Positive formulae over free groups, *Algebra i Logika*, 5(4):25-42 (1966).

[19] Ю. И. Ожигов. Уравнения с двумя неизвестными в свободной группе. *Докл. АН СССР*, 268(4):808–814, 1983. Yu. I. Ozhigov, Equations with two unknowns in a free group, *Soviet Math. Dokl.*, 27(1983).

[20] А. А. Разборов. О системах уравнений в свободной группе. PhD thesis, МГУ, 1987. A. A. Razborov, On systems of equations in a free group, Moscow State University, 1987.

[21] А. А. Разборов. О системах уравнений в свободной группе. *Изв. АН СССР, сер. матем.*, 48(4):779–832, 1984. A. A. Razborov, On systems of equations in a free group, *Math. USSR Izvestiya*, 25(1):115-162, 1985.

[22] А. А. Разборов. Об уравнениях в свободной группе, общие решения которых не представимы в виде суперпозиции конечного числа параметрических функций. In *Тезисы 9 всесоюзного симпозиума по теории групп*, page 54, Москва, 1984. A. A. Razborov, An equation in a free group whose set of solutions does not allow a representation as a superposition of a finite number of parametric functions. In *Proceedings of the 9th All-Union Symposium on the Group Theory*, Moscow, 1984, p. 54.

[23] Ю. И. Хмелевский. Системы уравнений в свободной группе, I. *Изв. АН СССР, сер. матем.*, 35(6):1237–1268, 1971. Yu. I. Khmelevskiĭ, Systems of equations in a free group, *Math. USSR Izv.*, 5(1971).

[24] Ю. И. Хмелевский. Системы уравнений в свободной группе, II. *Изв. АН СССР, сер. матем.*, 36(1):110–179, 1972. Yu. I. Khmelevskiĭ, Systems of equations in a free group, *Math. USSR Izv.*, 6(1972).