European Journal of Mathematics (2016) 2:614–643
DOI 10.1007/s40879-016-0101-9

RESEARCH ARTICLE

Surjectivity of certain word maps on PSL $(2, \mathbb{C})$ and SL $(2, \mathbb{C})$

Tatiana Bandman[1] · **Yuri G. Zarhin**[2]

Received: 6 July 2015 / Accepted: 1 February 2016 / Published online: 21 March 2016

Abstract Let $n \geqslant 2$ be an integer and F_n the free group on n generators, $F^{(1)}$, $F^{(2)}$ its first and second derived subgroups. Let K be an algebraically closed field of characteristic zero. We show that if $w \in F_n \setminus F^{(2)}$, then the corresponding word map $\mathrm{PSL}(2, K)^n \to \mathrm{PSL}(2, K)$ is surjective. We also describe certain word maps that are surjective on $\mathrm{SL}(2, \mathbb{C})$.

Keywords Special linear group · Word map · Trace map · Magnus embedding

Mathematics Subject Classification 20F70 · 20F14 · 20F45 · 20E32 · 20G20 · 14L10 · 14L24 · 14L35

T. Bandman was partially supported by the Ministry of Absorption (Israel), the Israeli Science Foundation (Israeli Academy of Sciences, Center of Excellence Program), and the Minerva Foundation (Emmy Noether Research Institute of Mathematics). Yu.G. Zarhin was partially supported by a grant from the Simons Foundation (#246625 to Yuri Zarkhin).
This work was started in September 2013 when the authors were visiting Max Planck Institute für Mathematik (Bonn, Germany) and was completed in May–June 2015 when they enjoyed the hospitality of the Department of Mathematics of the Weizmann Institute of Science (Rehovot, Israel). We are most grateful to both Institutions.

✉ Tatiana Bandman
bandman@macs.biu.ac.il

Yuri G. Zarhin
zarhin@math.psu.edu

1 Department of Mathematics, Bar-Ilan University, 5290002 Ramat Gan, Israel

2 Department of Mathematics, Pennsylvania State University, University Park, PA 16802, USA

1 Introduction

The surjectivity of word maps on groups became recently a vivid topic: the review on the latest activities may be found in [3,17,19,22].

Let $w \in F_n$ be an element of the free group F_n on $n > 1$ generators $g_1, \dots, g_n$:

$$w = \prod_{i=1}^{k} g_{n_i}^{m_i}, \qquad 1 \leqslant n_i \leqslant n.$$

For a group G by the same letter w we shall denote the corresponding word map $w\colon G^n \to G$ defined as a non-commutative product by the formula

$$w(x_1, \dots, x_n) = \prod_{i=1}^{k} x_{n_i}^{m_i}. \tag{1}$$

We call $w(x_1, \dots, x_n)$ both a *word in n letters* if considered as an element of a free group and a *word map in n letters* if considered as the corresponding map $G^n \to G$. We assume that it is reduced, i.e. $n_i \neq n_{i+1}$ for every $1 \leqslant i \leqslant k-1$ and $m_i \neq 0$ for $1 \leqslant i \leqslant k$.

Let K be a field and H a connected semisimple linear algebraic group that is defined over K. If w is not the identity then, by the Borel theorem [6], the regular map of (affine) K-algebraic varieties

$$w\colon H^n \to H, \qquad (h_1, \dots, h_n) \mapsto w(h_1, \dots, h_n)$$

is *dominant*, i.e., its image is a Zariski dense subset of H. Let us consider the group $G = H(K)$ and the image

$$w_G = w(G^n) = \big\{z \in G : z = w(x_1, \dots, x_n) \text{ for some } (x_1, \dots, x_n) \in G^n\big\}.$$

We say that a word (word map) w is *surjective* on G if $w_G = G$.

In [18, Problem 7], [19, Question 2.1 (i)], the following question is formulated: *Assume that w is not a power of another reduced word and $G = H(K)$. Is w surjective when $K = \mathbb{C}$ is a field of complex numbers and H is of adjoint type?*

According to [19], Question 2.1 (i) is still open, even in the simplest case $G =$ PSL(2, $\mathbb{C}$), even for words in two letters.

We consider word maps on groups $G = \mathrm{SL}(2, K)$ and $\widetilde{G} = \mathrm{PSL}(2, K)$. Put

$$F = F_n, \qquad F^{(1)} = [F, F], \qquad F^{(2)} = \big[F^{(1)}, F^{(1)}\big].$$

As usual, $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ stand for the ring of integers and fields of rational, real and complex numbers respectively. $\mathbb{A}(K)^m_{x_1,\dots,x_m}$ or, simply, $\mathbb{A}^m$, stands for the m-dimensional affine space over a field K with coordinates $x_1, \dots, x_m$. If $K = \mathbb{C}$, we use the notation $\mathbb{C}^m_{x_1,\dots,x_m}$.

Let $w \in F$. For the corresponding word map on $G = \mathrm{SL}(2, K)$ we check the following properties of the image w_G.

Properties 1.1

(a) w_G contains all semisimple elements x with $\mathrm{tr}(x) \neq 2$;
(b) w_G contains all unipotent elements x with $\mathrm{tr}(x) = 2$;
(c) w_G contains all minus unipotent elements x with $\mathrm{tr}(x) = -2$ and $x \neq -\mathrm{id}$;
(d) w_G contains $-\mathrm{id}$.

The word map w is surjective on $G = \mathrm{SL}(2, K)$ if all Properties 1.1 are satisfied. For the surjectivity on $\widetilde{G} = \mathrm{PSL}(2, K)$ it is sufficient that only Properties 1.1 (a), (b) are valid.

Definition 1.2 (cf. [2]) We say that the word map w is *almost surjective* on $G = \mathrm{SL}(2, K)$ if it has Properties 1.1 (a)–(c), i.e. $w_G \supset \mathrm{SL}(2, K) \setminus -\{\mathrm{id}\}$.

The goal of the paper is to describe certain words $w \in F$ such that the corresponding word maps are surjective or almost surjective on G and/or $\widetilde{G}$. Assume that the field K is algebraically closed. If $w(x_1, \ldots, x_d) = x_i^n$ then w is surjective on G if and only if n is odd (see [10,11]). Indeed, the element

$$x = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

is not a square in $\mathrm{SL}(2, K)$. Since only the elements with $\mathrm{tr}(x) = -2$ may be outside w_G [10,11], the induced by w word map $\widetilde{w}$ is surjective on $\widetilde{G}$.

Consider a word map (1). For an index $j \leqslant n$ let $S_j = \sum_{n_i=j} m_i$. If, say, $S_1 \neq 0$, then $w(x_1, \mathrm{id}, \ldots, \mathrm{id}) = x_1^{S_1}$, hence the word w is surjective on $\mathrm{PSL}(2, K)$. If $S_j = 0$ for all $1 \leqslant j \leqslant n$, then $w \in F^{(1)} = [F, F]$. In Sect. 5 we prove (see Corollary 5.4) the following:

> *The word map defined by a word $w \in F^{(1)} \setminus F^{(2)}$ is surjective on* $\mathrm{PSL}(2, K)$ *if K is an algebraically closed field with* $\mathrm{char}(K) = 0$.

The proof makes use of a variation on the Magnus Embedding Theorem, which is stated in Sect. 3 and proven in Sect. 4.

In Sects. 6–8, we consider words in two variables, i.e. $n = 2$. In this case we give explicit formulas for $w(x, y)$, where $x, y \in \mathrm{SL}(2, \mathbb{C})$ are upper triangular matrices. Using explicit formulas, in Sects. 7–8 we provide criteria for surjectivity and almost surjectivity of a word map on $G = \mathrm{SL}(2, \mathbb{C})$. In Sect. 7, these criteria are formulated in terms of exponents $a_i, b_i, i = 1 \ldots, k$, of the word

$$w(x, y) = \prod_{i=1}^{k} x^{a_i} y^{b_i},$$

where $a_i \neq 0$ and $b_i \neq 0$ for all $i = 1, \ldots, k$. A sample of such criteria is (Corollary 7.4)

If all b_i are positive, then the word map w is either surjective or the square of another word $v \neq \mathrm{id}$.

In Sect. 8, we connect the almost surjectivity of a word map with a property of the corresponding trace map. The last sections contain explicit examples.

2 Semisimple elements

Let K be an algebraically closed field with $\mathrm{char}(K) = 0$, and $G = \mathrm{SL}(2, K)$. Consider a word map $w: G^n \to G$ defined by (1). We consider G as an affine set

$$G = \{ad - bc = 1\} \subset \mathbb{A}^4_{a,b,c,d}.$$

The following lemma is, may be, known, but the authors do not have a proper reference.

Lemma 2.1 *A regular non-constant function on G^n omits no values in K.*

Proof Since all sets are affine, a function f regular on G^k is a restriction of a polynomial P_f onto G^k. We use induction on k.

Step 1. $k = 1$. G is an irreducible quadric. Assume that $f \in K[G]$ omits a value. Let $p: G \to \mathbb{A}^1_a$ be a projection defined by $p(a, b, c, d) = a$. If $a \neq 0$ then the fiber $F_a = p^{-1}(a) \cong \mathbb{A}^2_{b,c}$ is an affine space with coordinates b, c because $d = (1+bc)/a$. Since f omits a value, the restriction $f|_{F_a}$ is constant for every $a \neq 0$. Therefore it is constant on every fiber (note that the fiber $a = 0$ is connected). On the other hand, f has to be constant along the curve

$$C = \{(a, 1, -1, 0)\} \cong \mathbb{A}^1_a(K).$$

Since the curve $C \subset G$ intersects every fiber F_a of projection p, the function f is constant on G.

Step 2. Assume that the statement of the lemma is valid for all $k \leqslant n$. Let $f \in K[G^n]$ omit a value. We have $G^n = M \times N$, where $M = G^{n-1}$ and $N = G$. Let $p: G^n \to N$ be a natural projection. Then, by induction assumption, f is constant along every fiber of this projection. Take $x \in M$ and consider the set $C = x \times N \subset G^n$. Then $f|_C = \mathrm{const}$ and C intersects every fiber of p. Hence, f is constant. □

Proposition 2.2 *For every word $w(x_1, \ldots, x_k) \neq \mathrm{id}$ the image w_G contains every element $z \in G$ with $a = \mathrm{tr}(z) \neq \pm 2$.*

Proof We consider $G^n \subset \mathbb{A}^{4n}$ as the product of

$$G_i = \{a_i d_i - b_i c_i = 1\} \subset \mathbb{A}^4_{a_i,b_i,c_i,d_i},$$

$1 \leqslant i \leqslant n$. The function $f(a_1, b_1, c_1, d_1, \ldots, a_n, b_n, c_n, d_n) = \mathrm{tr}(w(x_1, \ldots, x_n))$ is a polynomial in $4n$ variables with integer coefficients, i.e., $f \in K[G^n]$. According to Lemma 2.1, it takes on all values in K. Thus for every value $A \in K$ there is an element

$u = w(y_1, \ldots, y_n) \in w_G$ such that $\mathrm{tr}(u) = A$. Let now $z \in G$, $A = \mathrm{tr}(z) \neq \pm 2$. Since $\mathrm{tr}(z) = \mathrm{tr}(u)$, z is conjugate to u, i.e., there is $v \in G$ such that $vuv^{-1} = z$. Hence $z = w(vy_1v^{-1}, \ldots, vy_nv^{-1})$. □

It follows that in order to check whether the word map w is surjective on G (or on $\widetilde{G}$) it is sufficient to check whether the elements z with $\mathrm{tr}(z) = \pm 2$ (or the elements z with $\mathrm{tr}(z) = 2$, respectively) are in the image. For that we need a version of the Embedding Magnus Theorem.

3 Variation on the Magnus Embedding Theorem: statements

Let $n \geqslant 2$ be an integer and $\Lambda_n = \mathbb{Z}[t_1, t_1^{-1}, \ldots, t_n, t_n^{-1}]$ be the ring of Laurent polynomials in n independent variables $t_1, \ldots, t_n$ over $\mathbb{Z}$. Let $F = F_n$ be a free group of rank n with generators $\{g_1, \ldots, g_n\}$. Recall: we write $F^{(1)}$ for the derived subgroup of F and $F^{(2)}$ for the derived subgroup of $F^{(1)}$. We have

$$F^{(2)} \subset F^{(1)} \subset F;$$

both $F^{(1)}$ and $F^{(2)}$ are normal subgroups in F. The quotient $A = F/F^{(1)} = \mathbb{Z}^n$ is a free abelian group of rank n with (standard) generators $\{e_1, \ldots, e_n\}$ where each e_i is the image of g_i, $1 \leqslant i \leqslant n$. The group ring $\mathbb{Z}[A]$ of A is canonically isomorphic to Λ_n: under this isomorphism each $e_i \in A \subset \mathbb{Z}[A]$ goes to

$$t_i \in \mathbb{Z}\left[t_1, t_1^{-1}, \ldots, t_n, t_n^{-1}\right] = \Lambda_n.$$

We write R_n for the ring of polynomials

$$\Lambda_n[s_1, \ldots, s_n] = \mathbb{Z}\left[t_1, t_1^{-1}, \ldots, t_n, t_n^{-1}; s_1, \ldots, s_n\right]$$

in n independent variables $s_1, \ldots, s_n$ over Λ_n. If R is a commutative ring with 1 then we write $\mathrm{T}(R)$ for the group of invertible 2×2 matrices of the form

$$\begin{bmatrix} a & 0 \\ b & 1 \end{bmatrix}$$

with $a \in R^*, b \in R$ and $\mathrm{ST}(R)$ for the group of unimodular 2×2 matrices of the form

$$\begin{bmatrix} a & 0 \\ b & a^{-1} \end{bmatrix}$$

with $a \in R^*, b \in R$. We have

$$\mathrm{T}(R) \subset \mathrm{GL}(2, R), \qquad \mathrm{ST}(R) \subset \mathrm{SL}(2, R).$$

Every homomorphism $R \to R'$ of commutative rings (with 1) induces the natural group homomorphisms

$$\mathrm{T}(R) \to \mathrm{T}(R'), \qquad \mathrm{ST}(R) \to \mathrm{ST}(R'),$$

which are injective if $R \to R'$ is injective.

The following assertion (that is based on the properties of the famous Magnus embedding [20]) was proven in [26, Lemma 2].

Theorem 3.1 *The assignment*

$$g_i \mapsto \begin{bmatrix} t_i & 0 \\ s_i & t_i^{-1} \end{bmatrix}, \qquad 1 \leqslant i \leqslant n,$$

extends to a group homomorphism $\mu_{\mathrm{W}} : F \to \mathrm{ST}(R_n)$ *with kernel* $F^{(2)}$ *and therefore defines an embedding*

$$F/F^{(2)} \hookrightarrow \mathrm{ST}(R_n) \subset \mathrm{SL}(2, R_n).$$

It follows from Theorem 3.1 that if K is a field of characteristic zero, whose transcendence degree over $\mathbb{Q}$ is, at least, $2n$ then there is an embedding

$$F/F^{(2)} \hookrightarrow \mathrm{ST}(K) \subset \mathrm{SL}(2, K).$$

(In particular, it works for $K = \mathbb{R}, \mathbb{C}$ or the field $\mathbb{Q}_p$ of p-adic numbers [26].) The aim of the following considerations is to replace in this statement the lower bound $2n$ by n.

Theorem 3.2 *The assignment*

$$g_i \mapsto \begin{bmatrix} t_i & 0 \\ 1 & t_i^{-1} \end{bmatrix}, \qquad 1 \leqslant i \leqslant n,$$

extends to a group homomorphism $\mu_1 : F \to \mathrm{ST}(\Lambda_n)$ *with kernel* $F^{(2)}$ *and therefore defines an embedding*

$$F/F^{(2)} \hookrightarrow \mathrm{ST}(\Lambda_n) \subset \mathrm{SL}(2, \Lambda_n).$$

Remark 3.3 Let

$$\mathrm{ev}_1 : R_n = \Lambda_n[s_1, \ldots, s_n] \to \Lambda_n$$

be the Λ_n-algebra homomorphism that sends all s_i to 1 and let

$$\mathrm{ev}_1{}^* : \mathrm{ST}(R_n) \to \mathrm{ST}(\Lambda_n)$$

be the group homomorphism induced by ev_1. Then μ_1 coincides with the composition

$$\mathrm{ev}_1{}^* \circ \mu_{\mathrm{W}} : F \to \mathrm{ST}(R_n) \to \mathrm{ST}(\Lambda_n).$$

Corollary 3.4 *Let K be a field of characteristic zero. Suppose that the transcendence degree of K over $\mathbb{Q}$ is, at least, n. Then there is a group embedding*

$$F/F^{(2)} \hookrightarrow \mathrm{ST}(K) \subset \mathrm{SL}(2, K).$$

The proof of Theorem 3.2 is based on the following observation.

Lemma 3.5 *Let K be a field of characteristic zero. Suppose that the transcendence degree of K over $\mathbb{Q}$ is, at least, n and let $\{u_1, \dots, u_n\} \subset K$ be an n-tuple of algebraically independent elements (over $\mathbb{Q}$). Let $\mathbb{Q}(u_1, \dots, u_n)$ be the subfield of K generated by $\{u_1, \dots, u_n\}$ and let us consider K as the $\mathbb{Q}(u_1, \dots, u_n)$-vector space. Let $\{y_1, \dots, y_n\} \subset K$ be an n-tuple that is linearly independent over $\mathbb{Q}(u_1, \dots, u_n)$. Let R be the subring of K generated by $3n$ elements $u_1, u_1^{-1}, \dots, u_n, u_n^{-1}; y_1, \dots, y_n$.*

Then the assignment

$$g_i \mapsto \begin{bmatrix} u_i & 0 \\ y_i & 1 \end{bmatrix} \in \mathrm{T}(R), \qquad 1 \leqslant i \leqslant n,$$

extends to a group homomorphism $\mu \colon F \to \mathrm{T}(R) \subset \mathrm{T}(K)$ with kernel $F^{(2)}$ and therefore defines an embedding

$$F/F^{(2)} \hookrightarrow \mathrm{T}(R) \subset \mathrm{T}(K).$$

Example 3.6 Let K be the field $\mathbb{Q}(t_1, \dots, t_n)$ of rational functions in n independent variables $t_1, \dots, t_n$ over $\mathbb{Q}$. One may view Λ_n as the subring of K generated by $2n$ elements $t_1, t_1^{-1}, \dots, t_n, t_n^{-1}$. By definition, the n-tuple $\{t_1, \dots, t_n\} \subset K$ is algebraically independent (over $\mathbb{Q}$). Clearly, the n-tuple

$$\left\{u_1 = t_1^2, \dots, u_i = t_i^2, \dots, u_n = t_n^2\right\} \subset K$$

is also algebraically independent. Then the n elements $y_1 = t_1, \dots, y_i = t_i, \dots, y_n = t_n$ are linearly independent over the (sub)field $\mathbb{Q}(t_1^2, \dots, t_n^2) = \mathbb{Q}(u_1, \dots, u_n)$. Indeed, if a rational function

$$f(t_1, \dots, t_n) = \sum_{i=1}^{n} t_i \cdot f_i$$

where all $f_i \in \mathbb{Q}(t_1^2, \dots, t_n^2)$ then

$$\begin{aligned} 2t_1 f_1 &= f(t_1, t_2, \dots, t_n) - f(-t_1, t_2, \dots, t_n), \dots, \\ 2t_i f_i &= f(t_1, \dots, t_i, \dots, t_n) - f(t_1, \dots, -t_i, \dots, t_n), \dots, \\ 2t_n f_n &= f(t_1, \dots, t_i, \dots, t_n) - f(t_1, \dots, t_i, \dots, -t_n). \end{aligned}$$

This proves that if $f = 0$ then all f_i are also zero, i.e., the set $\{t_1, \dots, t_n\}$ is linearly independent over $\mathbb{Q}(t_1^2, \dots t_n^2)$.

By definition, R coincides with the subring of K generated by $3n$ elements $t_1^2, t_1^{-2}, \dots, t_n^2, t_n^{-2}; t_1, \dots, t_n$. This implies easily that $R = \Lambda_n$. Applying Lemma 3.5, we conclude the example by the following statement.

The assignment

$$g_i \mapsto \begin{bmatrix} t_i^2 & 0 \\ t_i & 1 \end{bmatrix} \in \mathrm{T}(\Lambda_n), \qquad 1 \leqslant i \leqslant n,$$

extends to a group homomorphism $\mu\colon F \to \mathrm{T}(R) = \mathrm{T}(\Lambda_n)$ *with kernel* $F^{(2)}$ *and therefore defines an embedding*

$$F/F^{(2)} \hookrightarrow \mathrm{T}(\Lambda_n).$$

We prove Lemma 3.5, Theorem 3.2 and Corollary 3.4 in Sect. 4.

4 Variation on the Magnus Embedding Theorem: proofs

Proof of Lemma 3.5 Let $\Lambda \subset \mathbb{Q}(u_1, \dots, u_n) \subset K$ be the subring generated by $2n$ elements $u_1, u_1^{-1}, \dots, u_n, u_n^{-1}$. Since u_i are algebraically independent over $\mathbb{Q}$, the assignment

$$t_i \mapsto u_i, \qquad t_i^{-1} \mapsto u_i^{-1}$$

extends to a ring isomorphism $\Lambda_n \cong \Lambda$. The linear independence of y_i over $\mathbb{Q}(u_1, \dots, u_n)$ implies that $M = \Lambda \cdot y_1 + \dots + \Lambda \cdot y_n \subset R \subset K$ is a free Λ-module of rank n. On the other hand, let

$$U \subset \Lambda^* \subset \mathbb{Q}(u_1, \dots, u_n)^* \subset K^*$$

be the multiplicative (sub)group generated by all u_i. The assignment $g_i \mapsto u_i$ extends to the surjective group homomorphism $\delta\colon F \twoheadrightarrow U$ with kernel $F^{(1)}$ and gives rise to the group isomorphism $A \cong U$, which sends e_i to u_i and allows us to identify the group ring $\mathbb{Z}[U]$ of U with $\Lambda \cong \Lambda_n = \mathbb{Z}[A]$. Notice that M carries the natural structure of free $\mathbb{Z}[U]$-module of rank n defined by

$$\lambda(m) = \lambda \cdot m \in K, \qquad \lambda \in \mathbb{Z}[U] = \Lambda \subset K, \quad m \in M \subset K.$$

We have

$$\mu(F) \subset \begin{bmatrix} U & 0 \\ M & 1 \end{bmatrix} \subset \mathrm{T}(R) \subset \mathrm{GL}_2(R).$$

It follows from [27, Lemma 1 (c), p. 175] that $\ker(\mu)$ coincides with the derived subgroup of $\ker(\delta)$. Since $\ker(\delta) = F^{(1)}$, we conclude that $\ker(\mu) = F^{(2)}$ and we are done. □

Proof of Theorem 3.2 Let us return to the situation of Example 3.6. In particular, the group homomorphism (we know its kernel, thanks to already proven Lemma 3.5) $\mu : F \to \mathrm{T}(\Lambda_n) \subset \mathrm{GL}_2(\Lambda_n)$ is defined by

$$\mu(g_i) = \begin{bmatrix} t_i^2 & 0 \\ t_i & 1 \end{bmatrix} \in \mathrm{T}(\Lambda_n)$$

for all g_i. Let us consider the group homomorphism

$$\kappa : F \to \Lambda_n^*, \qquad g_i \mapsto t_i.$$

Since t_i are algebraically independent, they are multiplicatively independent and $\ker(\kappa) = F^{(1)}$. We claim that $\mu_1 : F \to \mathrm{ST}(\Lambda_n)$ coincides with the group homomorphism

$$g \mapsto \kappa(g)^{-1} \cdot \mu(g).$$

Indeed, we have for all g_i

$$\kappa(g_i)^{-1} \cdot \mu(g_i) = t_i^{-1} \cdot \begin{bmatrix} t_i^2 & 0 \\ t_i & 1 \end{bmatrix} = \begin{bmatrix} t_i & 0 \\ 1 & t_i^{-1} \end{bmatrix} = \mu_1(g_i) \subset \mathrm{ST}(\Lambda_n),$$

which proves our claim. Recall that we need to check that $\ker(\mu_1) = F^{(2)}$. In order to do that, first notice that $\mu_1(g)$ is of the form

$$\begin{bmatrix} \kappa(g) & 0 \\ * & \kappa(g)^{-1} \end{bmatrix}$$

for all $g \in F$ just because it is true for all $g = g_i$. This implies that $\ker(\mu_1) \subset \ker(\kappa) = F^{(1)}$. But $\mu = \mu_1$ on $F^{(1)}$. This implies that $\ker(\mu_1) = \ker(\mu) \cap F^{(1)}$. However, as we have seen in Example 3.6, $\ker(\mu) = F^{(2)} \subset F^{(1)}$. This implies that

$$\ker(\mu_1) = F^{(2)} \cap F^{(1)} = F^{(2)}$$

and we are done. □

Proof of Corollary 3.4 There exists an n-tuple $\{x_1, \ldots, x_n\} \subset K$ that is algebraically independent over $\mathbb{Q}$. The assignment

$$t_i \mapsto x_i, \qquad t_i^{-1} \mapsto x_i^{-1}$$

extends to an *injective* ring homomorphism

$$\Lambda_n = \mathbb{Z}\left[t_1, t_1^{-1}, \ldots, t_n, t_n^{-1}\right] \hookrightarrow K.$$

This implies that $\mathrm{ST}(\Lambda_n)$ is isomorphic to a subgroup of $\mathrm{ST}(K)$. Thanks to Theorem 3.2, $F/F^{(2)}$ is isomorphic to a subgroup of $\mathrm{ST}(\Lambda_n)$. This implies that $F/F^{(2)}$ is isomorphic to a subgroup of $\mathrm{ST}(K)$. □

Similar arguments prove the following generalization of Theorem 3.2.

Theorem 4.1 *Let $\{b_1, \ldots, b_n\}$ be an n-tuple of non-zero integers. Then the assignment*

$$g_i \mapsto \begin{bmatrix} t_i & 0 \\ b_i & t_i^{-1} \end{bmatrix}, \qquad 1 \leqslant i \leqslant n,$$

extends to a group homomorphism $F \to \mathrm{ST}(\Lambda_n)$ with kernel $F^{(2)}$.

5 Word maps and unipotent elements

Lemma 5.1 *Let w be an element of $F^{(1)}$ that does not belong to $F^{(2)}$. Then there exists a non-zero Laurent polynomial*

$$\mathcal{L}_w = \mathcal{L}_w(t_1, \ldots t_n) \in \mathbb{Z}\left[t_1, t_1^{-1}, \ldots, t_n, t_n^{-1}\right] = \Lambda_n$$

such that

$$\mu_1(w) = \begin{bmatrix} 1 & 0 \\ \mathcal{L}_w & 1 \end{bmatrix}.$$

Proof We have seen in the course of the proof of Theorem 3.2 that for all $g \in F$

$$\mu_1(g) = \begin{bmatrix} \kappa(g) & 0 \\ * & \kappa(g)^{-1} \end{bmatrix} \in \mathrm{ST}(\Lambda_n).$$

This means that there exists a Laurent polynomial $\mathcal{L}_g \in \Lambda_n$ such that

$$\mu_1(g) = \begin{bmatrix} \kappa(g) & 0 \\ \mathcal{L}_g & \kappa(g)^{-1} \end{bmatrix}.$$

We have also seen that if $g \in F^{(1)}$ then $\kappa(g) = 1$. Since $w \in F^{(1)}$,

$$\mu_1(w) = \begin{bmatrix} 1 & 0 \\ \mathcal{L}_w & 1 \end{bmatrix}$$

with $\mathcal{L}_w \in \Lambda_n$. On the other hand, by Theorem 3.2, $\ker(\mu_1) = F^{(2)}$. Since $w \notin F^{(2)}$, $\mathcal{L}_w \neq 0$ in Λ_n. □

Corollary 5.2 *Let w be an element of $F^{(1)}$ that does not belong to $F^{(2)}$. Suppose that $\mathbf{a} = \{a_1, \ldots, a_n\}$ is an n-tuple of non-zero rational numbers such that $c =$*

$\mathcal{L}_w(a_1,\ldots,a_n) \neq 0$. *(Since* $\mathcal{L}_w \neq 0$*, such an n-tuple always exists.) Let us consider the group homomorphism*

$$\gamma_{\mathbf{a}}\colon F \to \mathrm{ST}(\mathbb{Q}) \subset \mathrm{SL}(2,\mathbb{Q}), \qquad g_i \mapsto \begin{bmatrix} a_i & 0 \\ 1 & a_i^{-1} \end{bmatrix} = Z_i.$$

Then

$$\gamma_{\mathbf{a}}(w) = \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} = w(Z_1,\ldots,Z_n).$$

is a unipotent matrix that is not the identity matrix.

Proof One has only to notice that $\gamma_{\mathbf{a}}$ is the composition of μ_1 and the group homomorphism $\mathrm{ST}(\Lambda_n) \to \mathrm{ST}(\mathbb{Q})$ induced by the ring homomorphism $\Lambda_n \to \mathbb{Q}$, $t_i \mapsto a_i$, $t_i^{-1} \mapsto a_i^{-1}$. □

Corollary 5.3 *Let* w *be an element of* $F^{(1)}$ *that does not belong to* $F^{(2)}$*. Let* K *be a field of characteristic zero. Then for every unipotent matrix* $X \in \mathrm{SL}(2,K)$ *there exists a group homomorphism* $\psi_{w,X}\colon F \to \mathrm{SL}(2,K)$ *such that* $\psi_{w,X}(w) = X$*. In other words, there exist* $Z_1,\ldots,Z_n \in \mathrm{SL}(2,K)$ *such that* $w(Z_1,\ldots,Z_n) = X$*.*

Proof We have

$$\mathbb{Q} \subset K, \qquad \mathrm{SL}(2,\mathbb{Q}) \subset \mathrm{SL}(2,K) \lhd \mathrm{GL}(2,K).$$

We may assume that X is *not* the identity matrix. Let $\mathbf{a} = \{a_1,\ldots,a_n\}$ and $\gamma_{\mathbf{a}}$ be as in Corollary 5.2. Recall that $c = \mathcal{L}_w(a_1,\ldots,a_n) \neq 0$. Then there exists a matrix $S \in \mathrm{GL}(2,K)$ such that

$$X = S \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} S^{-1}.$$

Let us consider the group homomorphism $\psi_{w,X}\colon F \to \mathrm{SL}(2,K)$, $g \mapsto S\gamma_a(g)S^{-1}$. Then $\psi_{w,X}$ sends w to

$$S\gamma_{\mathbf{a}}(w)S^{-1} = S \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} S^{-1} = X.$$ □

Corollary 5.4 *Let* w *be an element of* $F^{(1)}$ *that does not belong to* $F^{(2)}$*. Let* K *be an algebraically closed field of characteristic zero. Then the word map* w *is surjective on* $\mathrm{PSL}(2,K)$.

Proof Consider w as a word map on $G = \mathrm{SL}(2,K)$. Due to Corollary 5.3, the image w_G contains all unipotents. According to Proposition 2.2, the image contains all semisimple elements as well. Thus, the word map w has Properties 1.1 (a) and (b). It follows that it is surjective on $\mathrm{PSL}(2,K)$. □

Remark 5.5 In [12], the words from $F^{(1)} \setminus F^{(2)}$ are proved to be surjective on $\mathrm{SU}(n)$ for an infinite set of integers n.

Theorem 5.6 *Let w be an element of $F^{(1)}$ that does not belong to $F^{(2)}$. Let G be a connected semisimple linear algebraic group of positive dimension over a field K of characteristic zero. If $u \in G(K)$ is a unipotent element then there exists a group homomorphism $F \to G(K)$ such that the image of w coincides with u. In other words, there exist $Z_1, \dots, Z_n \in G(K)$ such that $w(Z_1, \dots, Z_n) = u$.*

Proof Let $\mathbf{a} = \{a_1, \dots, a_n\}$, $\gamma_{\mathbf{a}}$ and $c = \mathcal{L}_w(a_1, \dots, a_n) \neq 0$ be as in Corollary 5.2. By Lemma 5.7 below, there exists a group homomorphism $\phi \colon \mathrm{ST}(K) \to G(K)$ such that $u = \phi(\mathbf{u}_1)$ for

$$\mathbf{u}_1 = \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \in \mathrm{ST}(K).$$

Now the result follows from Corollary 5.2: the desired homomorphism is the composition $\phi \circ \gamma_{\mathbf{a}} \colon F \to \mathrm{ST}(K) \to G(K)$. □

Lemma 5.7 *Let K be a field of characteristic zero, G a connected semisimple linear algebraic K-group of positive dimension, and u a unipotent element of $G(K)$. Then for every non-zero $c \in K$ there is a group homomorphism $\phi \colon \mathrm{ST}(K) \to G(K)$ such that u is the image of*

$$\mathbf{u}_1 = \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \in \mathrm{ST}(K).$$

Proof Let us identify the additive algebraic K-group $\mathbb{G}_{\mathrm{a}}$ with the closed subgroup H of all matrices of the form

$$v(t) = \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}$$

in $\mathrm{SL}(2)$. Its Lie subalgebra $\mathrm{Lie}(H)$ is the one-dimensional K-vector subspace $\mathrm{Lie}(H) = \{\lambda \mathbf{x}_0 : \lambda \in K\}$ of $\mathfrak{sl}_2(K)$ generated by the matrix

$$\mathbf{x}_0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \subset \mathfrak{sl}_2(K).$$

Here we view the K-Lie algebra $\mathfrak{sl}_2(K)$ of 2×2 traceless matrices as the Lie algebra of the algebraic K-group $\mathrm{SL}(2)$. Moreover, $\exp(\lambda \mathbf{x}_0) = v(\lambda)$ for all $\lambda \in K$.

We may view G as a closed algebraic K-subgroup of the linear algebraic group $\mathrm{GL}(N) = \mathrm{GL}(V)$, where V is an N-dimensional K-vector space for a suitable positive integer N. Then

$$u \in G(K) \subset \mathrm{Aut}_K(V) = \mathrm{GL}(N, K).$$

Thus the K-Lie algebra $\mathrm{Lie}(G)$ becomes a certain *semisimple* Lie subalgebra of $\mathrm{End}_K(V)$. Here we view $\mathrm{End}_K(V)$ as the Lie algebra $\mathrm{Lie}(\mathrm{GL}(V))$ of the K-algebraic group $\mathrm{GL}(V)$. As usual, we write

$$\mathrm{Ad}\colon G(K) \to \mathrm{Aut}_K(\mathrm{Lie}(G))$$

for the adjoint action of G. We have

$$\mathrm{Ad}(g)(u) = gug^{-1}$$

for all $g \in G(K) \subset \mathrm{Aut}_K(V)$ and $u \in \mathrm{Lie}(G) \subset \mathrm{End}_K(V)$. Since u is a unipotent element, the linear operator $u - 1\colon V \to V$ is nilpotent. Let us consider the nilpotent linear operator

$$x = \log(u) = \sum_{i=1}^{\infty} (-1)^{i+1} \frac{(u-1)^i}{i} \in \mathrm{End}_K(V)$$

([7, Section 7, p. 106], [24, Section 23, p. 336]) and the corresponding homomorphism of algebraic K-groups

$$\varphi_u\colon H \to \mathrm{GL}(V), \qquad v(t) \mapsto \exp(tx) = v(0) + tx + \cdots$$

In particular, since $\mathbf{u}_1 = v(1)$, $\varphi_u(\mathbf{u}_1) = u$. Clearly, the differential of φ_u

$$d\varphi_u\colon \mathrm{Lie}(H) \to \mathrm{Lie}(\mathrm{GL}(V)) = \mathrm{End}_K(V)$$

is defined as

$$d\varphi_u(\lambda \mathbf{x}_0) = \lambda x \qquad \text{for all} \quad \lambda \in K,$$

and sends $\mathbf{x}_0$ to $x \in \mathrm{Lie}(\mathrm{GL}(V))$. Since $\varphi_u(m) = u^m \in G(K)$ for all integers m and G is closed in $\mathrm{GL}(V)$ in Zariski topology, the image $\varphi_u(H)$ of H lies in G and therefore one may view φ_u as a homomorphism of algebraic K-groups $\varphi_u\colon H \to G$. This implies

$$d\varphi_u(\mathrm{Lie}(H)) \subset \mathrm{Lie}(G);$$

in particular, $x \in \mathrm{Lie}(G)$.

There exists a *cocharacter* $\Phi\colon \mathbb{G}_m \to G \subset \mathrm{GL}(V)$ of K-algebraic group G such that for each $\beta \in K^* = \mathbb{G}_m(K)$

$$\mathrm{Ad}(\Phi(\beta))(x) = \beta^2 x$$

(see [21, Section 6, pp. 402–403], here $\mathbb{G}_m$ is the multiplicative algebraic K-group). This means that for all $\lambda \in K$

$$\Phi(\beta)\lambda x\,\Phi(\beta)^{-1} = \mathrm{Ad}(\Phi(\beta))(\lambda x) = \lambda\beta^2 x = \beta^2\lambda x \in \mathrm{Lie}(G) \subset \mathrm{End}_K(V),$$

which implies that

$$\Phi(\beta)(\exp(\lambda x))\,\Phi(\beta)^{-1} = \exp\big(\Phi(\beta)\lambda x\Phi(\beta)^{-1}\big) = \exp(\beta^2\lambda x).$$

It follows that

$$\Phi(\beta)\left(\exp\left(\frac{\lambda}{c}\,x\right)\right)\Phi(\beta)^{-1} = \exp\left(\beta^2\frac{\lambda}{c}\,x\right).$$

Recall that $\mathrm{ST}(K)$ is a *semi-direct product* of its normal subgroup $H(K)$ and the torus

$$\mathrm{T}_1(K) = \left\{\begin{bmatrix}\beta^{-1} & 0\\ 0 & \beta\end{bmatrix} : \beta \in K^*\right\} \subset \mathrm{ST}(K).$$

In addition,

$$\begin{bmatrix}\beta^{-1} & 0\\ 0 & \beta\end{bmatrix}\begin{bmatrix}1 & 0\\ \lambda & 1\end{bmatrix}\begin{bmatrix}\beta^{-1} & 0\\ 0 & \beta\end{bmatrix}^{-1} = \begin{bmatrix}1 & 0\\ \beta^2\lambda & 1\end{bmatrix} \qquad \text{for all} \quad \lambda \in K, \quad \beta \in K^*.$$

It follows from [8, Chapter III, Proposition 27, p. 240] that there is a group homomorphism $\phi\colon \mathrm{ST}(K) \to G(K)$ that sends each $\left(\begin{smallmatrix}1 & 0\\ \lambda & 1\end{smallmatrix}\right)$ to $\exp(\lambda x/c)$ and each $\left(\begin{smallmatrix}\beta^{-1} & 0\\ 0 & \beta\end{smallmatrix}\right)$ to $\Phi(\beta)$. Clearly, ϕ sends $\mathbf{u}_1 = \left(\begin{smallmatrix}1 & 0\\ c & 1\end{smallmatrix}\right)$ to $\exp(cx/c) = \exp(x) = u$. □

6 Words in two letters on SL$(2,\mathbb{C})$

In this section we consider words in two letters. We provide the explicit formulas for $w(x, y)$, where x, y are upper triangular matrices. This enables us to extract some additional information on the image of words in two letters.

Consider a word map $w(x, y) = x^{a_1}y^{b_1}\dots x^{a_k}y^{b_k}$, where $a_i \neq 0$ and $b_i \neq 0$ for all $i = 1,\dots,k$. Let $A(w) = \sum_{i=1}^k a_i$ and $B(w) = \sum_{i=1}^k b_i$.

If $A(w) = B(w) = 0$, then $w \in F^{(1)} = [F, F]$. Since $F^{(1)}$ is a free group generated by elements $w_{n,m} = [x^n, y^m]$, $n \neq 0$, $m \neq 0$ [23, Chapter 1, Section 1.3], the word w with $A(w) = B(w) = 0$ may be written as a (non-commutative) product (with $s_i \neq 0$)

$$w = \prod_{i=1}^{r} w_{n_i,m_i}^{s_i}. \tag{2}$$

Moreover, the shortest (reduced) representation of this kind is unique. We denote by $S_w(n, m)$ the number of appearances of $w_{n,m}$ in representation (2) of w and by $R_w(n, m)$ the sum of exponents at all appearances. We denote by $\mathrm{Supp}(w)$ the set of all pairs (n, m) such that $w_{n,m}$ appears in the product. For example, if $w = w_{1,1} w_{2,2}^5 w_{1,1}^{-1}$, then

$$\begin{aligned} &\mathrm{Supp}(w) = \{(1,1), (2,2)\}, \\ &S_w(1,1) = 2, \qquad S_w(2,2) = 1, \\ &R_w(1,1) = 0, \qquad R_w(2,2) = 5. \end{aligned}$$

The subgroup

$$F^{(2)} = \left[F^{(1)}, F^{(1)}\right] = \left\{w \in F^{(1)} : R_w(n,m) = 0,\ (n,m) \in \mathrm{Supp}(w)\right\}.$$

Example 6.1 The Engel word

$$e_n = \underbrace{[\ldots[x, y], y], \ldots y]}_{n \text{ times}}$$

belongs to $F^{(1)} \setminus F^{(2)}$ (see also [12]). Indeed, the direct computation shows that

$$\begin{aligned} yw_{n,m} &= yx^n y^m x^{-n} y^{-m} \\ &= yx^n y^{-1} x^{-n} \cdot x^n y y^m x^{-n} y^{-m} y^{-1} \cdot y = w_{n,1}^{-1} w_{n,m+1} y, \\ yw_{n,m}^{-1} &= y \cdot y^m x^n y^{-m} x^{-n} \\ &= y^{(m+1)} x^n y^{-(m+1)} x^{-n} \cdot x^n y x^{-n} y^{-1} \cdot y = w_{n,m+1}^{-1} w_{n,1} y. \end{aligned}$$

It follows that

$$yw_{1,m}^s y^{-1} = \left(w_{1,1}^{-1} w_{1,m+1}\right)^s. \tag{3}$$

Let us prove by induction that $|R_{e_n}(1,n)| = 1$, $S_{e_n}(1,n) = 1$ and $S_{e_n}(r,m) = 0$ if $r \neq 1$ or $m > n$, i.e.

$$e_n = \prod_{i=1}^{s} w_{1,m_i}^{s_i} \cdot w_{1,n}^{\varepsilon} \cdot \prod_{j=1}^{t} w_{1,k_j}^{t_j} \tag{4}$$

for some integers $t \geqslant 0$, $s \geqslant 0$, $m_i < n$, $k_j < n$, and where $\varepsilon = \pm 1$.

Indeed $e_1 = w_{1,1}$. Assume that the claim is valid for all $k \leqslant n$. We have $e_{n+1} = e_n y e_n^{-1} y^{-1}$. Using (4), we get

$$e_{n+1} = e_n \left(\prod_{j=t}^{1} y w_{1,k_j}^{-t_j} y^{-1}\right) y w_{1,n}^{-\varepsilon} y^{-1} \left(\prod_{i=s}^{1} y w_{1,m_i}^{-s_i} y^{-1}\right).$$

Applying (3) to every factor of this product, we obtain that e_{n+1} has the needed form.

Thus the claim will remain valid for $n+1$. Since $|R_{e_n}(1,n)| = 1$, $e_n \notin F^{(2)}$. The surjectivity of the Engel words on simple algebraic groups was studied in [2,12,16]. There is a beautiful proof of surjectivity of e_n on $\mathrm{PSL}(2, \mathbb{C})$ in [18, Corollary 4].

Let us take

$$x = \begin{pmatrix} \lambda & c \\ 0 & 1/\lambda \end{pmatrix}, \tag{5}$$

$$y = \begin{pmatrix} \mu & d \\ 0 & 1/\mu \end{pmatrix}. \tag{6}$$

Then

$$x^n = \begin{pmatrix} \lambda^n & c \cdot h_{|n|}(\lambda)\,\mathrm{sgn}(n) \\ 0 & 1/\lambda^n \end{pmatrix}, \qquad y^m = \begin{pmatrix} \mu^m & d \cdot h_{|m|}(\mu)\,\mathrm{sgn}(m) \\ 0 & 1/\mu^m \end{pmatrix}.$$

Here sgn is the signum function, and (see [1, Lemma 5.2]) for $n \geqslant 1$

$$h_n(\zeta) = \frac{\zeta^{2n} - 1}{\zeta^{n-1}(\zeta^2 - 1)}.$$

Note that $h_n(1) = n$. Direct computations show that

$$x^n y^m = \begin{pmatrix} \lambda^n \mu^m & d \cdot \lambda^n\,\mathrm{sgn}(m)\,h_{|m|}(\mu) + c \cdot \mathrm{sgn}(n)\,h_{|n|}(\lambda)\,\mu^{-m} \\ 0 & \lambda^{-n}\mu^{-m} \end{pmatrix}, \tag{7}$$

$$x^{-n} y^{-m} = \begin{pmatrix} \lambda^{-n} \mu^{-m} & -d \cdot \lambda^{-n}\,\mathrm{sgn}(m)\,h_{|m|}(\mu) - c \cdot \mathrm{sgn}(n)\,h_{|n|}(\lambda)\,\mu^{m} \\ 0 & \lambda^{n}\mu^{m} \end{pmatrix},$$

$$w_{n,m}(x, y) = \begin{pmatrix} 1 & f(c, d, n, m) \\ 0 & 1 \end{pmatrix},$$

where

$$f(c, d, n, m) = ch_{|n|}(\lambda)\,\mathrm{sgn}(n)\,\lambda^n(1 - \mu^{2m}) + dh_{|m|}(\mu)\,\mathrm{sgn}(m)\,\mu^m(\lambda^{2n} - 1).$$

Hence,

$$w(x, y) = \prod_{i=1}^{r} w_{n_i, m_i}^{s_i}(x, y) = \begin{pmatrix} 1 & F_w(c, d, \lambda, \mu) \\ 0 & 1 \end{pmatrix},$$

where

$$F_w(c, d, \lambda, \mu) = \sum_{i=1}^{r} s_i\, f(c, d, n_i, m_i) = c\Phi_w(\lambda, \mu) + d\Psi_w(\lambda, \mu)$$

and

$$\Phi_w(\lambda, \mu) = \sum_{(\alpha,\beta)\in \mathrm{Supp}(w)} R_w(\alpha, \beta)\,\mathrm{sgn}(\alpha)(1 - \mu^{2\beta})\,\frac{(\lambda^{2|\alpha|} - 1)\,\lambda^{\alpha}}{\lambda^{|\alpha|-1}\,(\lambda^2 - 1)}, \tag{8}$$

$$\Psi_w(\lambda, \mu) = \sum_{(\alpha,\beta)\in \mathrm{Supp}(w)} R_w(\alpha, \beta)\,\mathrm{sgn}(\beta)(\lambda^{2\alpha} - 1)\,\frac{(\mu^{2|\beta|} - 1)\,\mu^{\beta}}{\mu^{|\beta|-1}\,(\mu^2 - 1)}. \tag{9}$$

(Since the order of factors in w is not relevant for (8) and (9), we use here α, β instead of n_i, m_i to simplify the formulas).

Proposition 6.2 *Rational functions $\Phi(\lambda,\mu)$ and $\Psi(\lambda,\mu)$ are non-zero linearly independent rational functions.*

Remark 6.3 It is evident from the Magnus Embedding Theorem that at least one of functions $\Phi(\lambda,\mu)$ and $\Psi(\lambda,\mu)$ is not identical zero. It follows from Proposition 6.2 that the same is valid for both of them.

Proof The proof is based on the following

Claim 6.4 *If $\Phi_w(\lambda,\mu)\equiv 0$ then $R_w(\alpha,\beta)=0$ for all $(\alpha,\beta)\in \mathrm{Supp}(w)$.*

Proof We use induction by the number $|\mathrm{Supp}(w)|$ of elements in $\mathrm{Supp}(w)$ for the word w. If $\mathrm{Supp}(w)$ contains only one pair (α,β), then there is nothing to prove, because

$$\Phi(\lambda,\mu)=R_w(\alpha,\beta)h_{|\alpha|}(\lambda)\,\mathrm{sgn}(\alpha)\lambda^{\alpha}(1-\mu^{2\beta}).$$

Assume that for words v with $|\mathrm{Supp}(v)|=l$ it is proved. Let w be such a word that $|\mathrm{Supp}(w)|=l+1$. Let $n=\max\{\alpha:(\alpha,\beta)\in\mathrm{Supp}(w)\}$.

Case 1. $n>0$. We have

$$\begin{aligned}\Phi_w(\lambda,\mu) &= \sum_{(\alpha,\beta)\in\mathrm{Supp}(w)} R_w(\alpha,\beta)\,\mathrm{sgn}(\alpha)(1-\mu^{2\beta})\frac{(\lambda^{2|\alpha|}-1)\lambda^{\alpha}}{\lambda^{|\alpha|-1}(\lambda^2-1)}\\ &= \sum_{(\alpha,\beta)\in\mathrm{Supp}(w)} R_w(\alpha,\beta)\,\mathrm{sgn}(\alpha)(1-\mu^{2\beta})\\ &\qquad\cdot\lambda^{a-|a|+1}\left(1+\lambda^2+\cdots+\lambda^{2(|\alpha|-1)}\right).\end{aligned}$$

It means that the coefficient of $\lambda^{2|n|-1}$ in the rational function $\Phi_w(\lambda,\mu)$ is

$$p(\mu)=\sum_{(n,\beta)\in\mathrm{Supp}(w)} R_w(n,\beta)(1-\mu^{2\beta}).$$

Hence, if $\Phi_w(\lambda,\mu)\equiv 0$, then $p(\mu)\equiv 0$, and all $R_w(n,\beta)=0$ for all β.

That means that $\Phi_w(\lambda,\mu)=\Phi_v(\lambda,\mu)$, where v is such a word that may be obtained from $w(x,y)=\prod_1^r w_{n_i,m_i}^{s_i}(x,y)$ by taking away every appearance of $w_{n,\beta}$:

$$v=\prod_{\substack{i=1\\ n_i\neq n}}^{r} w_{n_i,m_i}^{s_i}(x,y).$$

But $|\mathrm{Supp}(v)|\leqslant l$ and by the induction assumption $R_v(\alpha,\beta)=0$ for all $(\alpha,\beta)\in\mathrm{Supp}(v)$. Thus claim is valid for w in this case.

Case 2. $n < 0$. Let $-n' = \min\{\alpha : (\alpha, \beta) \in \mathrm{Supp}(w)\}$. We proceed as in Case 1 with $-n'$ instead of n: the coefficient of $\lambda^{-2n'+1}$ is

$$q(\mu) = \sum_{(-n',\beta)\in \mathrm{Supp}(w)} R_w(-n', \beta)(1 - \mu^{2\beta}).$$

If $\Phi_w(\lambda, \mu) \equiv 0$, then $q(\mu) \equiv 0$, and all $R_w(-n', \beta) = 0$ for all β. Once more, we may replace w by a word v with $|\mathrm{Supp}(v)| \leqslant l$. ■

Clearly, the similar statement is valid for $\Psi_w(\lambda, \mu)$. The functions Φ and Ψ are linearly independent, because Φ is odd with respect to λ and even with respect to μ, while Ψ has opposite properties. □

Proposition 6.5 *Assume that the word $w \in F^{(1)} \setminus F^{(2)}$ and that $\Phi_w(1, i) \neq 0$, where $i^2 = -1$. Then $-\mathrm{id} \in w_G$, where $G = \mathrm{SL}(2, \mathbb{C})$.*

Proof Assume that $\Phi(1, i) \neq 0$. From (8) we get

$$\Phi_w(1, i) = \sum_{\substack{(\alpha,\beta)\in \mathrm{Supp}(w) \\ \beta \text{ odd}}} 2R_w(\alpha, \beta)\alpha.$$

Take

$$x = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}, \qquad y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Then

$$[x, y] = \begin{pmatrix} a^2 & 0 \\ 0 & a^{-2} \end{pmatrix}.$$

Thus, if $w = \prod_1^r w_{n_j,m_j}^{s_j}$, then

$$w(x, y) = \prod_{m_j \text{ odd}} \begin{pmatrix} a^{2n_j s_j} & 0 \\ 0 & a^{-2n_j s_j} \end{pmatrix} = \begin{pmatrix} a^N & 0 \\ 0 & a^{-N} \end{pmatrix},$$

where $N = 2\sum_{m_j \text{ odd}} n_j s_j = \Phi_w(1, i) \neq 0$. Choose a such that $a^N = -1$. Then $w(x, y) = -\mathrm{id}$. □

Remark 6.6 The case $\Psi(i, 1) \neq 0$ may be treated in a similar way, one should only exchange roles of x and y.

Remark 6.7 Let $w = \prod_1^r w_{n_j,m_j}^{s_j}$, let $\gcd(m_j) = k = 2^d s$, s odd. Put $\mu_j = m_j/k$ and $u = \prod_1^r w_{n_j,\mu_j}^{s_j}$. Note that some of μ_j are odd. Let $z \in \mathrm{SL}(2, \mathbb{C})$ be such that

$$z^k = y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Then $w(x,z)=u(x,y)$, hence, if $\Phi_u(1,i)\neq 0$, then $-\mathrm{id}\in w_G$.

7 Surjectivity on $\mathbf{SL}(2,\mathbb{C})$

We keep the notation of Sect. 6.

Lemma 7.1 *Assume that $w=x^{a_1}y^{b_1}\dots x^{a_k}y^{b_k}$, $a_i\neq 0$, $b_i\neq 0$, $i=1,\dots,k$, $A=\sum a_i\neq 0$ or $B=\sum b_i\neq 0$ and x, y are defined by (5), (6) respectively. Then*

$$w(x,y)=\begin{pmatrix}\lambda^A\mu^B & \widetilde{F}_w(c,d,\lambda,\mu)\\ 0 & \lambda^{-A}\mu^{-B}\end{pmatrix},$$

where $\widetilde{F}_w(c,d,\lambda,\mu)=c\widetilde{\Phi}_w(\lambda,\mu)+d\widetilde{\Psi}_w(\lambda,\mu)$ and

$$\widetilde{\Phi}_w(\lambda,\mu)=\sum_{i=1}^k \operatorname{sgn}(a_i)h_{|a_i|}(\lambda)\frac{\lambda^{\sum_{j<i}a_j}\mu^{\sum_{j<i}b_j}}{\lambda^{\sum_{j>i}a_j}\mu^{\sum_{j\geqslant i}b_j}}, \tag{10}$$

$$\widetilde{\Psi}_w(\lambda,\mu)=\sum_{i=1}^k \operatorname{sgn}(b_i)h_{|b_i|}(\mu)\frac{\lambda^{\sum_{j\leqslant i}a_j}\mu^{\sum_{j<i}b_j}}{\lambda^{\sum_{j>i}a_j}\mu^{\sum_{j>i}b_j}}. \tag{11}$$

Proof We use induction on the complexity k of the word w. Using (7), we get

$$x^{a_1}y^{b_1}=\begin{pmatrix}\lambda^{a_1}\mu^{b_1} & d\cdot\lambda^{a_1}\operatorname{sgn}(b_1)h_{|b_1|}(\mu)+c\cdot\operatorname{sgn}(a_1)h_{|a_1|}(\lambda)\mu^{-b_1}\\ 0 & \lambda^{-a_1}\mu^{-b_1}\end{pmatrix}.$$

Thus for $k=1$ the lemma is valid. Assume that it is valid for $k'<k$. Let $u=x^{a_1}y^{b_1}\dots x^{a_{k-1}}y^{b_{k-1}}$ and $w=ux^{a_k}y^{b_k}$. By the induction assumption,

$$u(x,y)=\begin{pmatrix}\lambda^{A-a_k}\mu^{B-b_k} & \widetilde{F}_u(c,d,\lambda,\mu)\\ 0 & \lambda^{-A+a_k}\mu^{-B+b_k}\end{pmatrix}.$$

From (7) we get

$$x^{a_k}y^{b_k}=\begin{pmatrix}\lambda^{a_k}\mu^{b_k} & d\cdot\lambda^{a_k}\operatorname{sgn}(b_k)h_{|b_k|}(\mu)+c\cdot\operatorname{sgn}(a_k)h_{|a_k|}(\lambda)\mu^{-b_k}\\ 0 & \lambda^{-a_k}\mu^{-b_k}\end{pmatrix}.$$

Multiplying matrices u and $x^{a_k}y^{b_k}$, we get

$$\begin{aligned}\widetilde{F}_w(c,d,\lambda,\mu)=\lambda^{A-a_k}\mu^{B-b_k}\big(d\cdot\lambda^{a_k}\operatorname{sgn}(b_k)h_{|b_k|}(\mu)+c\cdot\operatorname{sgn}(a_k)h_{|a_k|}(\lambda)\mu^{-b_k}\big)\\ +\widetilde{F}_u(c,d,\lambda,\mu)\lambda^{-a_k}\mu^{-b_k}.\end{aligned}$$

Thus, the induction assumption implies that

$$\begin{aligned}\widetilde{\Phi}_w(\lambda,\mu) &= \operatorname{sgn}(a_k)h_{|a_k|}(\lambda)\mu^{-b_k}\lambda^{A-a_k}\mu^{B-b_k}\\ &\quad+\sum_{i=1}^{k-1}\operatorname{sgn}(a_i)h_{|a_i|}(\lambda)\,\frac{\lambda^{\sum_{j<i}a_j}\mu^{\sum_{j<i}b_j}}{\lambda^{\sum_{j=i+1}^{k}a_j}\mu^{\sum_{j=i}^{k}b_j}}\\ &=\sum_{i=1}^{k}\operatorname{sgn}(a_i)h_{|a_i|}(\lambda)\,\frac{\lambda^{\sum_{j<i}a_j}\mu^{\sum_{j<i}b_j}}{\lambda^{\sum_{j>i}a_j}\mu^{\sum_{j\geqslant i}b_j}}.\\ \widetilde{\Psi}_w(\lambda,\mu) &= \operatorname{sgn}(b_k)h_{|b_k|}(\mu)\lambda^{a_k}\lambda^{A-a_k}\mu^{B-b_k}\\ &\quad+\sum_{i=1}^{k-1}\operatorname{sgn}(b_i)h_{|b_i|}(\mu)\,\frac{\lambda^{\sum_{j\leqslant i}a_j}\mu^{\sum_{j<i}b_j}}{\lambda^{\sum_{j=i+1}^{k}a_j}\mu^{\sum_{j=i+1}^{k}b_j}}\\ &=\sum_{i=1}^{k}\operatorname{sgn}(a_i)h_{|a_i|}(\lambda)\,\frac{\lambda^{\sum_{j\leqslant i}a_j}\mu^{\sum_{j<i}b_j}}{\lambda^{\sum_{j>i}a_j}\mu^{\sum_{j>i}b_j}}.\end{aligned}$$

□

Denote

$$A_i=\sum_{j\leqslant i}a_i,\qquad B_i=\sum_{j<i}b_i,$$

and let C be a curve $C=\{\lambda^A\mu^B=-1\}\subset\mathbb{C}^2_{\lambda,\mu}$.

Multiplying (10) and (11) by $\lambda^A\mu^B$, we see that on C the following relations are valid:

$$\begin{aligned}\widetilde{\Phi}_w(\lambda,\mu)|_C &= -\sum_{i=1}^{k}\operatorname{sgn}(a_i)h_{|a_i|}(\lambda)\lambda^{2A_i-a_i}\mu^{2B_i},\\ \widetilde{\Psi}_w(\lambda,\mu)|_C &= -\sum_{i=1}^{k}\operatorname{sgn}(b_i)h_{|b_i|}(\mu)\lambda^{2A_i}\mu^{2B_i+b_i}.\end{aligned}$$

In particular, on C

$$\widetilde{\Phi}_w(1,\mu)|_C=-\sum_{i=1}^{k}a_i\mu^{2B_i}, \tag{12}$$

$$\widetilde{\Psi}_w(\lambda,1)|_C=-\sum_{i=1}^{k}b_i\lambda^{2A_i}. \tag{13}$$

Lemma 7.2 *Assume that $A\neq 0$ and the word map w is not surjective. Then*

$$\sum_{i=1}^{k}b_i\gamma^{2A_i}=0$$

for every root γ *of equation* $q(z) = z^A + 1 = 0$.

If $B \neq 0$ *and the word map* w *is not surjective, then*

$$\sum_{i=1}^{k} a_i \delta^{2B_i} = 0$$

for every root δ *of equation* $p(z) = z^B + 1 = 0$.

Proof The matrices z with $\mathrm{tr}(z) = 2$ are in the image, because $w(x, \mathrm{id}) = x^A$, $w(\mathrm{id}, y) = y^B$. It is evident that $-\mathrm{id}$ is in the image: one may take $c = d = 0$. Assume now that for complex numbers $\varkappa \neq 0$ the matrices

$$\begin{pmatrix} -1 & \varkappa \\ 0 & -1 \end{pmatrix}$$

are not in the image. This implies that $\widetilde{\Phi}_w(\lambda, \mu) \equiv 0$ and $\widetilde{\Psi}_w(\lambda, \mu) \equiv 0$ on the defined above curve $C = \{\lambda^A \mu^B = -1\} \subset \mathbb{C}^2_{\lambda,\mu}$.

If $A \neq 0$ or $B \neq 0$, then the pair $(\gamma, 1)$ or $(1, \delta)$ respectively belongs to the curve C. We have to use only (12), (13), respectively. □

Corollary 7.3 *Let* $2B_i = k_i B + T_i$*, where* k_i *are integers and* $0 \leqslant T_i < B \neq 0$*. If* w *is not surjective, then for every* $0 \leqslant T < B$

$$\sum_{i:T_i=T} a_i (-1)^{k_i} = 0.$$

Proof Indeed in this case

$$0 = \sum_{i=1}^{k} a_i \delta^{2B_i} = \sum_{T=0}^{B-1} \delta^T \sum_{i:T_i=T} a_i (-1)^{k_i}$$

for any root δ of equation $p(z) = z^B + 1 = 0$. Since $p(z)$ has no multiple roots, it implies that $p(z)$ divides the polynomial

$$p_1(z) = \sum_{T=0}^{B-1} z^T \sum_{i:T_i=T} a_i (-1)^{k_i}.$$

But since degree of polynomial $p(z)$ is bigger than degree of $p_1(z)$ that can be only if $p_1(z) \equiv 0$.

□

Corollary 7.4 *If all* b_i *are positive, then the word map* w *is either surjective or the square of another word* $v \neq \mathrm{id}$.

Proof In this case $0 \leqslant 2B_i < 2B$ and the sequence B_i is increasing. If w is not surjective, $p_1(z) \equiv 0$, by Corollary 7.3. Thus for every B_i there is B_j such that $2B_i = 2B_j + B$ and $a_i - a_j = 0$.

Thus, the sequence of $2B_i$ looks like

$$\begin{gathered}0 = 2B_1, \quad 2b_1 = 2B_2, \quad 2(b_1 + b_2) = 2B_3, \quad \dots,\\ 2(b_1 + \dots + b_s) = 2B_{s+1} = B,\\ 2(b_1 + \dots + b_{s+1}) = 2B_{s+2} = B + 2B_2 = B + 2b_1,\\ 2(b_1 + \dots + b_{s+2}) = 2B_{s+3} = B + 2B_3 = B + 2b_1 + 2b_2, \quad \dots,\\ 2(b_1 + \dots + b_{2s-1}) = 2B_{2s} = 2B_s + B,\\ 2(b_1 + \dots + b_{2s}) = 2B_{2s+1} = B + 2B_{s+1} = 2B.\end{gathered}$$

It follows that $k = 2s$ and

$$\begin{aligned}b_{s+1} &= B_{s+2} - B_{s+1} = B_2 - B_1 = b_1,\\ b_{s+2} &= B_{s+3} - B_{s+2} = B_3 - B_2 = b_2,\\ b_{2s-1} &= B_{2s} - B_{2s-1} = B_s - B_{s-1} = b_{s-1},\\ b_k = b_{2s} &= B_{2s+1} - B_{2s} = B_{s+1} - B_s = b_s.\end{aligned}$$

Thus,

$$b_i = b_{i+s}, \quad 2B_i = 2B_{i+s} + B, \quad a_i = a_{i+s}, \qquad i = 1, \dots, s.$$

Therefore the word is the square of $v = x^{a_1} y^{b_1} \dots x^{a_s} y^{b_s}$. □

Corollary 7.5 *If all b_i are negative, then the word map of the word w is either surjective or the square of another word $v \neq \mathrm{id}$.*

Proof We may change y to $z = y^{-1}$ and apply Corollary 7.4 to the word $w(x, z)$. □

Corollary 7.6 *If all a_i are positive, then the word map of the word w is either surjective or the square of another word $v \neq \mathrm{id}$.*

Proof Consider $v = x^{-1}$, $z = y^{-1}$, a word

$$w'(z, v) = w(x, y)^{-1} = y^{-b_k} x^{-a_k} \dots y^{-b_1} x^{-a_1} = z^{b_k} v^{a_k} \dots z^{b_1} v^{a_1},$$

and apply Corollary 7.4 to the word $w'(z, v)$. □

8 Trace criteria of almost surjectivity

For every word map $w(x, y)\colon G^2 \to G$ there are defined the trace polynomials $P_w(s, t, u) = \mathrm{tr}(w(x, y))$ and $Q_w = \mathrm{tr}(w(x, y)y)$ in three variables $s = \mathrm{tr}(x)$, $t = \mathrm{tr}(y)$, and $u = \mathrm{tr}(xy)$ [13–15,25].

In other words, the maps

$$\varphi_w\colon G^2 \to G^2, \qquad \varphi_w(x, y) = (w(x, y), y),$$
$$\psi_w\colon \mathbb{C}^3_{s,t,u} \to \mathbb{C}^3_{s,t,u}, \qquad \psi_w(s, t, u) = (P_w(s, t, u), t, Q_w(s, t, u))$$

may be included into the following commutative diagram:

$$\begin{tikzcd} G\times G \arrow[r, "\varphi_w"] \arrow[d, "\pi"'] & G\times G \arrow[d, "\pi"] \\ \mathbb{C}^3_{s,t,u} \arrow[r, "\psi_w"] & \mathbb{C}^3_{s,t,u}. \end{tikzcd}$$

Moreover, π is a surjective map [15]. For details, one can be referred to [3,5].

Since the coordinate t is invariant under ψ_w, for every fixed value $t = a \in \mathbb{C}$ we may consider the restriction $\psi_a(s, u) = (P_w(s, a, u), Q_w(s, a, u))$ of morphism ψ_w to the plane $\{t = a\} = \mathbb{C}^2_{s,u}$.

Definition 8.1 We say that $\psi_a(s, u)$ is BIG if the image $\psi_a(\mathbb{C}^2_{s,u}) = \mathbb{C}^2_{s,u} \setminus T_a$, where T_a is a finite set. We say that the trace map ψ_w of a word $w \in F$ is BIG if there is a value a such that $\psi_a(s, u)$ is BIG.

Proposition 8.2 *If the trace map ψ_w of a word $w \in F$ is* BIG *then the word map $w\colon G^2 \to G$ is almost surjective.*

Proof Let a be such a value of t that the map ψ_a is BIG. Let $S_a = T_a \cup \{(2, a)\} \cup \{(-2, -a)\}$. Consider lines $C_+ = \{s = 2\}$ and $C_- = \{s = -2\}$ in $\mathbb{C}^2_{s,u}$. Let $B_+ = C_+ \setminus (C_+ \cap S_a)$ and $B_- = C_- \setminus (C_- \cap S_a)$. Since S_a is finite, $B_+ \neq \varnothing$, $B_- \neq \varnothing$. Moreover, since these curves are outside S_a, we have: $D_+ = \psi^{-1}(B_+) \neq \varnothing$, $D_- = \psi^{-1}(B_-) \neq \varnothing$. Take $(s_0, u_0) \in D_+$ and $(s_1, u_1) \in D_-$. Then $\psi_w(s_0, a, u_0) = (2, a, b)$ with $a \neq b$ and $\psi_w(s_1, a, u_1) = (-2, a, d)$ with $a \neq -d$. The projection $\pi\colon G^2 \to \mathbb{C}^3_{s,t,u}$ is surjective, thus there is a pair $(x_0, y_0) \in G^2$ such that $\operatorname{tr}(x_0) = s_0$, $\operatorname{tr}(y_0) = a$, $\operatorname{tr}(x_0 y_0) = u_0$. Then $\pi(w(x_0, y_0)) = \psi_w(s_0, a, u_0) = (2, a, b)$. Hence, $\operatorname{tr}(w(x_0, y_0)) = 2$, but $w(x_0, y_0) \neq \mathrm{id}$, since $\operatorname{tr}(w(x_0, y_0) y_0) = b \neq a = \operatorname{tr}(y_0)$. Similarly, there is a pair $(x_1, y_1) \in G^2$ such that $\operatorname{tr}(x_1) = s_1$, $\operatorname{tr}(y_1) = a$, $\operatorname{tr}(x_1 y_1) = u_1$. Then $\pi(w(x_1, y_1)) = \psi_w(s_1, a, u_1) = (-2, a, d)$. Hence, $\operatorname{tr}(w(x_1, y_1)) = -2$, but $w(x_1, y_1) \neq -\mathrm{id}$, since $\operatorname{tr}(w(x_1, y_1) y_1) = d \neq -a = -\operatorname{tr}(y_1)$. It follows that all the elements $z \neq -\mathrm{id}$ with trace 2 and -2 are in the image of the word map w. □

Corollary 8.3 *Assume that the trace map ψ_w of a word w is* BIG. *Consider a sequence of words defined recurrently in the following way:*

$$v_1(x, y) = w(x, y), \qquad v_{n+1}(x, y) = w(v_n(x, y), y).$$

Then the word map $v_n\colon G^2 \to G$ is almost surjective for all $n \geqslant 1$.

Proof The trace map $\psi_n = \psi_{v_n}$ of the word map v_n is the n^{th} iteration $\psi_1^{(n)}$ of the trace map $\psi_1 = \psi_w$ (see [3,5]). Let us show by induction, that all maps ψ_n are BIG. Indeed ψ_1 is BIG by assumption, hence $(\psi_1)_a(\mathbb{C}^2_{s,u}) = \mathbb{C}^2_{s,u} - T_a$ for some value a and some finite set T_a. Assume now that ψ_{n-1} is BIG. Let for a value a of t the image $(\psi_{n-1})_a(\mathbb{C}^2_{s,u}) = \mathbb{C}^2_{s,u} \setminus N$ for some finite set N. Hence

$$\begin{aligned}(\psi_n)_a(\mathbb{C}^2_{s,u}) &= (\psi_1)_a((\psi_{n-1})_a(\mathbb{C}^2_{s,u})) = (\psi_1)_a(\mathbb{C}^2_{s,u} \setminus N)\\ &\supset (\psi_1)_a(\mathbb{C}^2_{s,u}) \setminus (\psi_1)_a(N) = \mathbb{C}^2_{s,u} \setminus (T_a \cup (\psi_1)_a(N)).\end{aligned}$$

Thus $(\psi_n)_a$ is BIG as well for the same value a.

According to Proposition 8.2, the word map v_n is almost surjective. □

Example 8.4 Consider the word $w(x, y) = v_1(x, y) = [yxy^{-1}, x^{-1}]$ and the corresponding sequence

$$v_n(x, y) = \left[yv_{n-1}y^{-1}, v_{n-1}^{-1}\right].$$

This is one of the sequences that were used for characterization of finite solvable groups (see [3,5,9]).

We have [5, Section 5.1]

$$\begin{aligned}\operatorname{tr}(w(x, y)) &= f_1(s, t, u) = \left(s^2 + t^2 + u^2 - ust - 4\right)\left(t^2 + u^2 - ust\right) + 2,\\ \operatorname{tr}(w(x, y)y) &= f_2(s, t, u) = f_1 t + (s(st - u) - t)\left(s^2 + t^2 + u^2 - ust - 4\right) - t.\end{aligned}$$

We want to show that for a general value $t = a$ the system of equations

$$f_1(s, a, u) = A, \qquad f_2(s, a, u) = B \tag{14}$$

has solutions for all pairs $(A, B) \in \mathbb{C}^2 \setminus T_a$, where T_a is a finite set.

Consider the system

$$\begin{aligned}h_1(s, u, a, C) &= \left(s^2 + a^2 + u^2 - usa - 4\right)\left(a^2 + u^2 - usa\right) = A - 2 = C,\\ h_2(s, u, a, D) &= (s(sa - u) - a)\left(s^2 + a^2 + u^2 - usa - 4\right)\\ &= B - a(C + 1) = D.\end{aligned} \tag{15}$$

Note that with respect to u the leading coefficients of h_1 and h_2 are 1 and $-s$ respectfully. The MAGMA computations show that the resultant (elimination of u) of $h_1 - C$ and $h_2 - D$ is of the form

$$R(s, a, C, D) = s^4 p_1(a, C, D) + s^2 p_2(a, C, D) + p_3(a, C, D).$$

It has a non-zero root $s \neq 0$ at any point (a, C, D), where at least two of three polynomials p_1, p_2, p_3 do not vanish. The MAGMA computations show that the ideals $J1 = \langle p_1, p_2\rangle \subset \mathbb{Q}[a, C, D]$, $J2 = \langle p_1, p_3\rangle \subset \mathbb{Q}[a, C, D]$, $J3 = \langle p_2, p_3\rangle \subset$

$\mathbb{Q}[a, C, D]$ generated, respectively, by $p_1(a, C, D)$ and $p_2(a, C, D)$, by $p_1(a, C, D)$ and $p_3(a, C, D)$, by $p_2(a, C, D)$ and $p_3(a, C, D)$, are one-dimensional. It follows that for a general value of a the set

$$\{p_1(a, C, D) = p_2(a, C, D) = 0\} \cup \{p_1(a, C, D) = p_3(a, C, D) = 0\} \\ \cup \{p_2(a, C, D) = p_3(a, C, D) = 0\}$$

is a finite subset $N_a \subset \mathbb{C}_{C,D}$. On the other hand, at any point (C, D) outside N_a the polynomial $R_a(s) = R(s, a, C, D)$ has a non-zero root, and, therefore system (15) has a solution. Thus, outside the finite set of points $T_a = \{(A = C + 2, B = D + a(C+1)) : (C, D) \in N_a\} \subset \mathbb{C}_{A,B}$, system (14) has a solution as well. Thus, $\psi_w = (f_1, t, f_2)$ is BIG and all word maps v_n are almost surjective on G. Let us cite the MAGMA computations for $t = a = 1$, where $p = h_1 - C$, $q = h_2 - D$ and R is the resultant of p, q with respect to u.

```
> r := u^2 + s^2 + 1 - u*s;
>
> p := (r - 4) * (r - s^2) - C;
>
> q := (r - 4) * (s * (s - u) - 1) - D;
>
> R := Resultant(p,q,u);
> R;
 - s^4*C^3 - 2*s^4*C^2*D + s^4*C^2 - 2*s^4*C*D^2 + s^4*C*D
 - s^4*D^3 + s^4*D^2 + 4*s^2*C^2*D - 4*s^2*C^2 + 8*s^2*C*D^2
 - 6*s^2*C*D + 6*s^2*D^3 - 8*s^2*D^2
 + C^2 - 2*C*D^2 + 8*C*D + D^4 - 8*D^3 + 16*D^2
>
> p1 := - C^3 - 2*C^2*D + C^2 - 2*C*D^2 + C*D - D^3 + D^2;
> p2 := 4*C^2*D - 4*C^2 + 8*C*D^2 - 6*C*D + 6*D^3 - 8*D^2;
> p3 := C^2 - 2*C*D^2 + 8*C*D + D^4 - 8*D^3 + 16*D^2;
> Factorization(p1);
[
    <C + D - 1, 1>,
    <C^2 + C*D + D^2, 1>
]
> Factorization(p2);
[
    <C^2*D - C^2 + 2*C*D^2 - 3/2*C*D + 3/2*D^3
                - 2*D^2, 1>
]
> Factorization(p3);
[
    <C - D^2 + 4*D, 2>
]
```

Clearly every pair among polynomials p_1, p_2, p_3 has only finite number of common zeros. For example, $p_1 = p_3 = 0$ implies $D^2 - 5D + 1 = 0$ or $(D^2-4D)^2 +$

$(D^2 - 4D)D + D^2 = 0$. Computations show also that the word $w(x, y)$ takes on value $-\mathrm{id}$. For example, one make take

$$x = \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}, \qquad y = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

where $t^2 = -1/2$. Therefore, the word v_1 is surjective. Here are computations:

```
> R<t> := PolynomialRing(Q);
> X := Matrix(R,2,2,[-1,1,-2,1]);
> Y := Matrix(R,2,2,[ 1,t,0,1]);
> X1 := Matrix(R,2,2,[1,-1,2,-1]);
> Y1 := Matrix(R,2,2,[1,-t,0,1]);
>
> Z := Y * X * Y1;
>
> p11 := Z[1,1];
> p12 := Z[1,2];
> p21 := Z[2,1];
> p22 := Z[2,2];
>
> Z1 := Matrix(R,2,2,[p22,-p12,-p21,p11]);
>
> W := Z * X1 * Z1 * X;
>
> q11 := W[1,1];
> q12 := W[1,2];
> q21 := W[2,1];
> q22 := W[2,2];
>
>
>  q11;
16*t^4 + 8*t^3 + 12*t^2 + 4*t + 1
>  q12;
-8*t^4 - 4*t^2
>  q21;
16*t^3 + 8*t
>   q22;
- 8*t^3 + 4*t^2 - 4*t + 1
```

Therefore, $t^2 = -1/2$ implies that $q_{11} = q_{22} = -1$, $q_{12} = q_{21} = 0$.

9 The word $v(x, y) = [[x, [x, y]], [y, [x, y]]]$

In this section we provide an example of a word v that is surjective though it belongs to $F^{(2)}$. The interesting feature of this word is the following: if we consider it as a polynomial in the Lie algebra $\mathfrak{sl}_2$ ($[x, y]$ being the Lie bracket) then it is not surjective [4, Example 4.9].

Theorem 9.1 *The word* $v(x, y) = [[x, [x, y]], [y[x, y]]]$ *is surjective on* $\mathrm{SL}(2, \mathbb{C})$ (*and, consequently, on* $\mathrm{PSL}(2, \mathbb{C})$).

Proof As it was shown in Proposition 2.2, for every $z \in \mathrm{SL}(2, \mathbb{C})$ with $\mathrm{tr}(z) \neq \pm 2$ there are $x, y \in \mathrm{SL}(2, \mathbb{C})^2$ such that $v(x, y) = z$.

Assume now that $a = \pm 2$. We have to show that $-\mathrm{id}$ is in the image and that there are matrices x, y in $\mathrm{SL}(2, \mathbb{C})$ such that

$$v(x, y) = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix}$$

has the following properties:

- $q_{12} + q_{22} = \pm 2$,
- $q_{12} \neq 0$.

We may look for these pairs among the matrices $x = \left(\begin{smallmatrix} 0 & b \\ c & d \end{smallmatrix}\right)$ and $y = \left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right)$.

In the following MAGMA calculations $C = [x, y]$, $D = [[x, y], x]$, $B = [[x, y], y]$, $A = [D, B]$. The ideal I in the polynomial ring $\mathbb{Q}[b, c, d, t]$ is defined by conditions $\det(x) = 1$, $\mathrm{tr}(A) = 2$. The ideal J in the polynomial ring $\mathbb{Q}[b, c, d, t]$ is defined by conditions $\det(x) = 1$, $\mathrm{tr}(A) = -2$. Let $T_+ \subset \mathrm{SL}(2)^2$ and $T_- \subset \mathrm{SL}(2)^2$ be, respectively, the corresponding affine subsets in the affine variety $\mathrm{SL}(2)^2$. The computations show that $q_{12}(b, c, d, t)$ does not vanish identically on T_+ or T_-.

```
> Q := Rationals();
> R<t,b,c,d> := PolynomialRing(Q,4);
> X := Matrix(R,2,2,[0,b,c,d]);
> Y := Matrix(R,2,2,[ 1,t,0,1]);
> X1 := Matrix(R,2,2,[d,-b,-c,0]);
> Y1 := Matrix(R,2,2,[1,-t,0,1]);
> C := X * Y * X1 * Y1;
> p11 := C[1,1];
> p12 := C[1,2];
> p21 := C[2,1];
> p22 := C[2,2];
> C1 := Matrix(R,2,2,[p22,-p12,-p21,p11]);
> D := C * X * C1 * X1;
>
> d11 := D[1,1];
> d12 := D[1,2];
> d21 := D[2,1];
> d22 := D[2,2];
> D1 := Matrix(R,2,2,[d22,-d12,-d21,d11]);
>
> B := C * Y * C1 * Y1;
```

```
>
> b11 := B[1,1];
> b12 := B[1,2];
> b21 := B[2,1];
> b22 := B[2,2];
> B1 := Matrix(R,2,2,[b22,-b12,-b21,b11]);
>
> A := D * B * D1 * B1;
>
> TA := Trace(A);
>
> q12 := A[1,2];
> I := ideal<R | b*c + 1, TA - 2>;
>
> IsInRadical(q12,I);
false
> J := ideal<R | b*c + 1, TA + 2>;
>
> IsInRadical(q12,J);
false
>
```

It follows that the function $q_{12}(b, c, d, t)$ does not vanish identically on the sets T_+ and T_-, hence, there are pairs with $\mathrm{tr}(v(x, y)) = 2$, $v(x, y) \neq \mathrm{id}$, and $\mathrm{tr}(v(x, y)) = -2$, $v(x, y) \neq -\mathrm{id}$.

In order to produce the explicit solutions for $v(x, y) = -\mathrm{id}$ and $v(x, y) = z$, $z \neq -\mathrm{id}$, $\mathrm{tr}(z) = -2$, consider the following matrices depending on one parameter d:

$$x = \begin{pmatrix} 1-d & 1 \\ -2/3 & d \end{pmatrix}, \qquad y = \begin{pmatrix} 2-3d & 0 \\ 0 & 3d-1 \end{pmatrix}.$$

Since the images of the commutator word on GL(2, $\mathbb{C}$) and SL(2, $\mathbb{C}$) are the same, we do not require that $\det(x) = 1$ or $\det(y) = 1$. We only assume that $\det(x) = d^2 - d - 2/3 \neq 0$ and $\det(y) = -9d^2 + 9d^2 - 2 \neq 0$. Let

$$A = v(x, y) = \begin{pmatrix} q_{11}(d) & q_{12}(d) \\ q_{21}(d) & q_{22}(d) \end{pmatrix}$$

and $TA = \mathrm{tr}(A)$. The MAGMA computations show that

$$q_{11}(d) + 1 = N_{11}\left(d^2 - d + \frac{1}{3}\right) H_{11}(d),$$
$$q_{22}(d) + 1 = N_{22}\left(d^2 - d + \frac{1}{3}\right) H_{22}(d),$$

$$q_{21}(d) = N_{21}\left(d-\frac{2}{3}\right)^2\left(d-\frac{1}{2}\right)^3\left(d-\frac{1}{3}\right)^2 \cdot \left(d^2-d-\frac{2}{3}\right)\left(d^2-d+\frac{1}{3}\right)H_{21}(d),$$

$$q_{12}(d) = N_{21}\left(d-\frac{2}{3}\right)^2\left(d-\frac{1}{2}\right)^3\left(d-\frac{1}{3}\right)^2 \cdot \left(d^2-d-\frac{2}{3}\right)\left(d^2-d+\frac{1}{3}\right)H_{12}(d),$$

$$TA+2 = N\left(d^2-d+\frac{1}{3}\right)H(d),$$

where N_{ij} and N are non-zero rational numbers; H_{ij} and H are polynomials with rational coefficients that are irreducible over $\mathbb{Q}$. Moreover $\deg H_{21} = \deg H_{12} = 25$ and $\deg H = 38$. It follows that if $d^2 - d + 1/3 = 0$ then $A = -\mathrm{id}$. If d is a root of H that is not a root of H_{21}, then A is a minus unipotent (which is *not* $-\mathrm{id}$). $\square$

Acknowledgments We thank Boris Kunyavskii for inspiring questions and useful comments, and Eugene Plotkin, Vladimir L. Popov, and Alexander Premet for help with references. We are grateful to both referees whose comments helped to correct the inaccuracies of the first version of the paper and for suggesting to use the Magnus Embedding Theorem.

References

1. Bandman, T., Garion, S.: Surjectivity and equidistribution of the word $x^a y^b$ on $\mathrm{PSL}(2,q)$ and $\mathrm{SL}(2,q)$. Internat. J. Algebra Comput. **22**(2), 1250017 (2012)
2. Bandman, T., Garion, S., Grunewald, F.: On the surjectivity of Engel words on $\mathrm{PSL}(2,q)$. Groups Geom. Dyn. **6**(3), 409–439 (2012)
3. Bandman, T., Garion, S., Kunyavskiĭ, B.: Equations in simple matrix groups: algebra, geometry, arithmetic, dynamics. Cent. Eur. J. Math. **12**(2), 175–211 (2014)
4. Bandman, T., Gordeev, N., Kunyavskiĭ, B., Plotkin, E.: Equations in simple Lie algebras. J. Algebra **355**, 67–79 (2012)
5. Bandman, T., Grunewald, F., Kunyavskiĭ, B., Jones, N.: Geometry and arithmetic of verbal dynamical systems on simple groups. Groups Geom. Dyn. **4**(4), 607–655 (2010)
6. Borel, A.: On free subgroups of semisimple groups. Enseign. Math. **29**(1–2), 151–164 (1983)
7. Borel, A.: Linear Algebraic Groups. Graduate Texts in Mathematics, vol. 126, 2nd edn. Springer, New York (1991)
8. Bourbaki, N.: General Topology, Chapters 1–4. Elements of Mathematics. Springer, Berlin (1989)
9. Bray, J.N., Wilson, J.S., Wilson, R.A.: A characterization of finite soluble groups by laws in two variables. Bull. London Math. Soc. **37**(2), 179–186 (2005)
10. Chatterjee, P.: On the surjectivity of the power maps of algebraic groups in characteristic zero. Math. Res. Lett. **9**(5–6), 741–756 (2002)
11. Chatterjee, P.: On the surjectivity of the power maps of semisimple algebraic groups. Math. Res. Lett. **10**(5–6), 625–633 (2003)
12. Elkasapy, A., Thom, A.: About Gotô's method showing surjectivity of word maps. Indiana Univ. Math. J. **63**(5), 1553–1565 (2014)
13. Fricke, R.: Über die Theorie der automorphen Modulgruppen. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, 91–101 (1896)
14. Fricke, R., Klein, F.: Vorlesungen über die Theorie der Automorphen Funktionen, vol. I–II. Teubner, Leipzig (1897, 1912)

15. Goldman, W.M.: Trace coordinates on Fricke spaces of some simple hyperbolic surfaces. In: Papadopoulos, A. (ed.) Handbook of Teichmüller Theory. Vol. II. IRMA Lectures in Mathematics and Theoretical Physics, vol. 13, pp. 611–684. European Mathematical Society, Zürich (2009). http://www.math.umd.edu/~wmg/publications.html
16. Gordeev, N.: On Engel words on simple algebraic groups. J. Algebra **425**, 215–244 (2015)
17. Kanel-Belov, A., Kunyavskii, B., Plotkin, E.: Word equations in simple groups and polynomial equations in simple algebras. Vestnik St. Petersburg Univ. Math. **46**(1), 3–13 (2013)
18. Klimenko, E., Kunyavskiĭ, B., Morita, J., Plotkin, E.: Word maps in Kac–Moody setting. Toyama Math. J. **37**, 25–53 (2015). arXiv:1506.01422
19. Kunyavskiĭ, B.: Complex and real geometry of word equations in simple matrix groups and algebras (2014). Private communication
20. Magnus, W.: Über den Beweis des Hauptidealsatzes. J. Reine Angew. Math. **170**, 235–240 (1934)
21. McNinch, G.J.: Optimal SL(2)-homomorphisms. Comment. Math. Helv. **80**(2), 391–426 (2005)
22. Segal, D.: Words. London Mathematical Society Lecture Note Series, vol. 361. Cambridge University Press, Cambridge (2009)
23. Serre, J.-P.: Trees. Springer Monographs in Mathematics. Springer, Berlin (2003)
24. Tauvel, P., Yu, R.W.T.: Lie Algebras and Algebraic Groups. Springer Monographs in Mathematics. Springer, Berlin (2005)
25. Vogt, H.: Sur les invariants fondamentaux des équations différentielles linéaires du second ordre. Annales Scientifiques de l'École Normale Supérieure **4**(Suppl.), 3–72 (1889)
26. Wehrfritz, B.A.F.: A residual property of free metabelian groups. Arch. Math. **20**, 248–250 (1969)
27. Wilson, J.S.: Free subgroups in groups with few relators. Enseign. Math. **56**(1–2), 173–185 (2010)