arXiv:1201.5260v2 [math.GR] 15 Feb 2013

CRITERIA FOR EQUIDISTRIBUTION OF SOLUTIONS OF WORD EQUATIONS ON SL(2)

TATIANA BANDMAN AND BORIS KUNYAVSKIĬ

ABSTRACT. We study equidistribution of solutions of word equations of the form $w(x, y) = g$ in the family of finite groups $\mathrm{SL}(2, q)$. We provide criteria for equidistribution in terms of the trace polynomial of w. This allows us to get an explicit description of certain classes of words possessing the equidistribution property and show that this property is generic within these classes.

1. INTRODUCTION

Equidistribution of solutions of various (systems of) diophantine equations has been remaining one of central topics in number theory, arithmetic geometry, ergodic theory. It is not our goal to review vast literature in the area. The reader interested in evolution of ideas in this fascinating domain of mathematics may find instructive to overview materials of ICM's, starting from the foundational address by Linnik (Stockholm, 1962) until impressive contributions of the past two decades: Margulis, Sarnak (Kyoto, 1990); Dani, Ratner (Zürich, 1994); Eskin (Berlin, 1998); Ullmo (Beijing, 2002); Einsiedler–Lindenstrauss, Michel–Venkatesh, Tschinkel (Madrid, 2006); Oh, Shah (Hyderabad, 2010). Each of the approaches mentioned above assumes its own understanding of the notion of equidistribution. What most of them share in common is focusing on certain group actions arising in a natural way and allowing one to combine methods of number theory and dynamical systems with group-theoretic considerations.

Let us describe the circle of problems we are interested in. First, we want to study *polynomial matrix equations*. In the most general form, one can consider equations of the form $P(A_1, \dots, A_m, X_1, \dots, X_d) = 0$ where $n \times n$-matrices $A_1, \dots, A_m$ with entries from a ring R are *given*, $X_1, \dots, X_d$ are *unknowns*, and P is an *associative noncommutative* polynomial. We, however, restrict our attention to a particular class of equations of the form $P(X_1, \dots, X_d) = A$ where A is a given matrix, $X_1, \dots, X_d$ are unknowns, and a solution must belong to a fixed subset $\mathcal{M} \subset \mathrm{M}(n, R)^d$. There are several cases where such an equation has a solution for a "generic" A (here $R = K$ is an algebraically closed field):

- $\mathcal{M} = G(K)^d$ where $G(K)$ is the group of rational points of a connected semisimple algebraic group and $P = w \neq 1$ is a nontrivial word (=monomial in $X_1, X_1^{-1}, \dots, X_d, X_d^{-1}$) (Borel [Bo], Larsen [La]);
- $\mathcal{M} = \mathfrak{g}^d$ where the Lie algebra $\mathfrak{g}$ of a semisimple algebraic K-group and a Lie polynomial P satisfy some additional assumptions (Bandman, Gordeev, Kunyavskiĭ and Plotkin [BGKP]);
- $\mathcal{M} = \mathrm{M}(n, R)^d$ and P satisfies some additional assumptions (Kanel-Belov, Malev and Rowen [KBMR]).

If $R = \mathbb{Z}$, in all these cases we may interpret the situation as follows: the generic fibre of the morphism $\mathbb{P}\colon \mathbb{M}^d \to \mathbb{M}$ of $\mathbb{Z}$-schemes, induced by the polynomial P, is a *dominant* morphism of $\mathbb{Q}$-schemes.

One can ask whether the situation is similar in *special* fibres of the morphism P. As the notion of dominance does not make much sense for finite sets, we would like to formalize the following phenomena:

- the maps $P_q\colon (M_q)^d \to M_q$ have "asymptotically large" images;
- the number $\#\{(A_1, \dots, A_d) \in (M_q)^d : P_q(A_1, \dots, A_d) = A\}$ (where $q = p^n$; $p = 2, 3, 5, \dots$; A runs over a "large" subset of M_q) is, in some reasonable sense, almost independent of A.

(Here M_q denotes the set of $\mathbb{F}_q$-points of the fibre of the scheme $\mathbb{M}$ at q, and P_q is the fibre of the morphism $\mathbb{P}$ at q.)

The conditions formulated above mean that the equations $P(X_1, \dots, X_d) = A$, with the right-hand side running, for each q, over "almost whole" set M_q, have many and almost equally many solutions in $(M_q)^d$, respectively. We shall call such morphisms *p-almost equidistributed*, or *almost equidistributed* (depending on whether p in the second condition is or is not fixed); the word "almost" will often be dropped. See Section 2 for precise definitions.

According to Larsen [La], Larsen and Shalev [LS1], for any word $w \neq 1$ and any family of Chevalley groups G_q of fixed type, the images of the maps $P_{w,q}\colon (G_q)^d \to G_q$ are "asymptotically large". Note, however, that for any individual G_q the image of $P_{w,q}$ may be very small: say, w may be identically 1 on $(G_q)^d$; moreover, even if this is not the case, then, according to an observation of Kassabov and Nikolov [KN] (see also a subsequent paper of Levy [Le1]), the image of w may consist only of a single conjugacy class together with the identity element. Recently Lubotzky [Lu] proved that such a phenomenon can happen in any finite simple group, for any conjugacy class; Levy [Le2] extended this result to some almost simple and quasisimple groups.

Our main result (Theorem 2.13) provides a necessary and sufficient condition on the word w in two variables under which the morphism $\mathbb{P}_w\colon \mathrm{SL}_2 \times \mathrm{SL}_2 \to \mathrm{SL}_2$ is almost equidistributed. This result can be viewed, on the one hand, as a refinement (in the SL_2-case) of equidistribution theorems of Larsen and Pink [LP], Larsen and Shalev [LS2], Larsen, Shalev and Tiep [LST] on general words w and general Chevalley groups G, and, on the other hand, as a generalization of equidistribution theorems for some particular words: Garion and Shalev [GS] (commutator words on any G), Bandman, Garion and Grunewald [BGG] (Engel words on SL_2), Bandman and Garion [BG] (positive words on SL_2). As a consequence, we obtain a somewhat surprising conclusion: if the word morphism as above has a large image (in the sense that for almost all q the image of $P_{w,q}$ contains all noncentral semisimple elements of $\mathrm{SL}(2,q)$), then it is almost equidistributed (in the terminology of the preceding paragraph, "many" implies "almost equally many").

Acting in the spirit of [GS], we deduce a criterion for $w\colon \mathrm{SL}_2 \times \mathrm{SL}_2 \to \mathrm{SL}_2$ to be *almost measure-preserving*.

Note that certain word maps are measure-preserving in a much stronger sense. Namely, if w is *primitive*, i.e., is a part of a basis of the free d-generated group F_d, then the corresponding word map $G^d \to G$ is measure-preserving for *every finite group* G, i.e., all fibres of this map have the same cardinality. Only primitive words possess this property, this was proven for $d = 2$ by Puder [Pu] and extended to arbitrary d by Puder and Parzanchevski [PP]. (Note that the word map P_w induced by a primitive word w is obviously surjective.) It is well known (see, e.g., Myasnikov and Shpilrain [MS]) that primitive words are asymptotically rare (negligible, in the terminology of Kapovich and Schupp [KS]). We are looking for criteria for equidistribution for more general words.

The criteria we are talking about are formulated in terms of the *trace polynomial* of the word w. It turns out (see our main results in Section 2; they are proved in Section 3) that "good" (equidistributed, measure-preserving) words are essentially those whose trace polynomial cannot be represented as a composition of two other polynomials. Since a "bad" trace polynomial tends to be the trace polynomial of some power word (see Section 4), we conclude (see Section 5) that within certain natural classes of words a "random" word is "good" ("good" words, i.e., those whose trace map is p-equidistributed for all but finitely many primes p, form an exponentially generic set, in the sense of [KS]).

2. Main results

We start with precise definitions of notions described in the introduction. We will follow the approach to equidistribution adopted in [GS]:

Definition 2.1. (cf. [GS, §3]) Let $f\colon X \to Y$ be a map between finite non-empty sets, and let $\varepsilon > 0$. We say that f is *ε-equidistributed* if there exists $Y' \subseteq Y$ such that

(i) $\#Y' > \#Y(1-\varepsilon)$;
(ii) $|f^{-1}(y) - \frac{\#X}{\#Y}| < \varepsilon \frac{\#X}{\#Y}$ for all $y \in Y'$.

Our setting is as follows. Let a family of maps of finite sets $P_q\colon X_q \to Y_q$ be given for every $q = p^n$. Assume that for all sufficiently large q the set Y_q is non-empty. For each such q take $y \in Y_q$ and denote

$$P_y = \{x \in X_q : P_q(x) = y\}.$$

Definition 2.2. Fix a prime p. With the notation as above, we say that the family $P_q\colon X_q \to Y_q$, $q = p^n$, is *p-equidistributed* if there exist a positive integer n_0 and a function $\varepsilon_p\colon \mathbb{N} \to \mathbb{N}$ tending to 0 as $n \to \infty$ such that for all $q = p^n$ with $n > n_0$ the set Y_q contains a subset S_q with the following properties:

(i) $\#S_q < \varepsilon_p(q)\,(\#Y_q)$;
(ii) $|\#P_y - \frac{\#X_q}{\#Y_q}| < \varepsilon_p(q)\frac{\#X_q}{\#Y_q}$ for all $y \in Y_q \setminus S_q$.

Remark 2.3. Definition 2.2 means that for $q = p^n$ large enough, the map $X_q \to Y_q$ is $\varepsilon_p(q)$-equidistributed, in the sense of Definition 2.1.

Definition 2.4. We say that the family $P_q\colon X_q \to Y_q$ is *equidistributed* if it is p-equidistributed for all p and there exists a function $\varepsilon\colon \mathbb{N} \to \mathbb{N}$ tending to 0 as $n \to \infty$ such that for every p and every $q = p^n$ large enough, we have $\varepsilon_p(q) \le \varepsilon(q)$.

Let us now consider the case where $Y_q = G_q$ is a Chevalley group over $\mathbb{F}_q$, $X_q = (G_q)^d$ is a direct product of its d copies ($d \ge 2$ is fixed), and $P_q = P_{w,q}\colon (G_q)^d \to G_q$ is the map induced by some fixed word $w \in F_d$: to each d-tuple $(g_1, \dots, g_d)$ we associate the value $w(g_1, \dots, g_d)$.

In the present paper we focus our attention on a particular case $d = 2$, $G_q = \mathrm{SL}(2,q)$. It is convenient to view the maps $P_{w,q}\colon \mathrm{SL}(2,q) \times \mathrm{SL}(2,q) \to \mathrm{SL}(2,q)$

as fibres of the morphism $\mathbb{P}_w\colon \mathrm{SL}_{2,\mathbb{Z}}\times\mathrm{SL}_{2,\mathbb{Z}} \to \mathrm{SL}_{2,\mathbb{Z}}$ of group schemes over $\mathbb{Z}$. We say that the morphism $\mathbb{P}_w$ (or, for brevity, the word w) is equidistributed (or p-equidistributed) if so is the family $P_{w,q}$.

In such a situation, there is a natural way to associate to any word $w = w(x, y) \in F_2$ its *trace polynomial*. This construction goes back to the 19th century (Vogt, Fricke, Klein), see, e.g., [Ho] for a modern exposition. For $G = \mathrm{SL}(2, k)$ (k is any commutative ring with 1) denote by $\mathrm{tr}(w)\colon G^2 \to G$ the trace character, $(g_1, g_2) \mapsto \mathrm{tr}(w(g_1, g_2))$. Then $\mathrm{tr}(w) = f_w(s, u, t)$ where $f_w \in \mathbb{Z}[s, u, t]$ is an integer polynomial in three variables $s = \mathrm{tr}(x)$, $u = \mathrm{tr}(xy)$, $t = \mathrm{tr}(y)$. We denote by the same letters the induced morphisms of affine $\mathbb{Z}$-schemes

$$f_w\colon \mathbb{A}^3_{s,u,t} = \mathrm{Spec}\,\mathbb{Z}[s, u, t] \to \mathbb{A}^1_z = \mathrm{Spec}\,\mathbb{Z}[z],$$

of affine $\overline{\mathbb{F}}_p$-schemes:

$$f_{w,p}\colon\ \mathrm{Spec}\,\overline{\mathbb{F}}_p[s, u, t] \to \mathrm{Spec}\,\overline{\mathbb{F}}_p[z],$$

and also maps of sets of $\overline{\mathbb{F}}_p$-points:

$$f_{w,p}\colon \mathbb{A}^3_{s,u,t}(\overline{\mathbb{F}}_p) \to \mathbb{A}^1_z(\overline{\mathbb{F}}_p)$$

(here $\mathbb{A}^N_{x_1,\dots,x_N}$ stands for affine space with coordinates $x_1, \dots, x_N$).

Our criteria for equidistribution of w will be formulated in terms of the polynomial f_w. Some recollections and definitions on polynomials are on order.

Definition 2.5. Let $\mathbb{F}$ be a finite field. We say that $h \in \mathbb{F}[x]$ is a permutation polynomial if the set of its values $\{h(z)\}_{z\in\mathbb{F}}$ coincides with $\mathbb{F}$.

Theorem 2.6. [LN, Theorem 7.14] *Let $q = p^n$. A polynomial $h \in \mathbb{F}_q[x]$ is a permutation polynomial of all finite extensions of $\mathbb{F}_q$ if and only if $h = ax^{p^k} + b$, where $a \neq 0$ and k is a non-negative integer.*

The following notions are essential for our criteria.

Definition 2.7. Let $\mathbb{F}$ be a field. We say that a polynomial $P \in \mathbb{F}[x_1, \dots, x_n]$ is $\mathbb{F}$-*composite* if there exist $Q \in \mathbb{F}[x_1, \dots, x_n]$, $\deg Q \geq 1$, and $h \in \mathbb{F}[z]$, $\deg h \geq 2$, such that $P = h \circ Q$. Otherwise, we say that P is $\mathbb{F}$-*noncomposite*.

Note that if $\mathbb{E}/\mathbb{F}$ is a separable field extension, it is known [AP, Theorem 1 and Proposition 1] that P is $\mathbb{F}$-composite if and only if P is $\mathbb{E}$-composite. In particular, working over perfect ground fields, we may always assume, if needed, that $\mathbb{F}$ is algebraically closed.

Definition 2.8. Let $P \in \mathbb{Z}[x_1, \dots, x_n]$. Fix a prime p.

- We say that P is *p-composite* if the reduced polynomial $P_p \in \mathbb{F}_p[x_1, \dots, x_n]$ is $\mathbb{F}_p$-composite. Otherwise, we say that P is *p-noncomposite*.
- We say that a p-composite polynomial P is *p-special* if, in the notation of Definition 2.7, $P_p = h \circ Q$ where $h \in \mathbb{F}_p[x]$ is a permutation polynomial of all finite extensions of $\mathbb{F}_p$.

Definition 2.9. We say that a polynomial $P \in \mathbb{Z}[x_1, \dots, x_n]$ is *almost noncomposite* if for every prime p it is either p-noncomposite or p-special. Otherwise we say that P is *very composite*.

Remark 2.10. If a polynomial $P \in \mathbb{Z}[x_1, \dots, x_n]$ is $\mathbb{Q}$-noncomposite, it is p-noncomposite for all but finitely many primes p [BDN, 2.2.1]. If $P \in \mathbb{Z}[x_1, \dots, x_n]$ is $\mathbb{Q}$-composite, it is very composite.

Example 2.11. Consider the family of Dickson polynomials $\mathcal{D}_n(x, a)$. Denote $D_n(x) = \mathcal{D}_n(x, 1)$. We have $D_n(x) = 2T_n(x/2)$ where $T_n(x)$ is the n^{th} Chebyshev polynomial. If n is not prime then D_n is very composite (see, e.g., Section 4 below). If $n = p$ is prime, then D_n is almost noncomposite and p-special since $D_p(x) = x^p$ in $\mathbb{F}_p[x]$.

We can now formulate our main results.

Theorem 2.12. *Let $w \in F_2$. The morphism $\mathbb{P}_w \colon \mathrm{SL}_{2,\mathbb{Z}} \times \mathrm{SL}_{2,\mathbb{Z}} \to \mathrm{SL}_{2,\mathbb{Z}}$ is p-equidistributed if and only if the trace polynomial f_w is either p-noncomposite or p-special.*

Theorem 2.13. *Let $w \in F_2$. The morphism $\mathbb{P}_w \colon \mathrm{SL}_{2,\mathbb{Z}} \times \mathrm{SL}_{2,\mathbb{Z}} \to \mathrm{SL}_{2,\mathbb{Z}}$ is equidistributed if and only if the trace polynomial f_w is almost noncomposite.*

Corollary 2.14. *Let $w(x, y) = x^{a_1} y^{b_1} \dots x^{a_r} y^{b_r}$ be a reduced word such that we have $f_w(s, u, t) = D_r(q(s, u, t))$ over $\mathbb{Q}$. Then $w(x, y) = (x^{a_1} y^{b_1})^r$.*

For a given word $w \in F_2$, let us now consider the family of groups $\hat{G}_q = \mathrm{PSL}(2, q)$ and the corresponding word maps $\hat{P}_{w,q} \colon \hat{G}_q \times \hat{G}_q \to \hat{G}_q$.

Proposition 2.15. *If the morphism $\mathbb{P}_w \colon \mathrm{SL}_{2,\mathbb{Z}} \times \mathrm{SL}_{2,\mathbb{Z}} \to \mathrm{SL}_{2,\mathbb{Z}}$ is equidistributed (or p-equidistributed), then so is the family $\hat{P}_{w,q} \colon \hat{G}_q \times \hat{G}_q \to \hat{G}_q$.*

3. Proofs

Fix a word w in F_2. We slightly change the general notation, and for a group Γ and $g \in \Gamma$ we denote

$$W_{g,\Gamma} = \{(x, y) \in \Gamma \times \Gamma : w(x, y) = g\}.$$

We will omit the subscript Γ when no confusion may arise. For $\Gamma = G_q = \mathrm{SL}(2, q)$ we denote this set by $W_{g,q}$ (or just W_g).

Since $\#G_q = q(q^2 - 1)$, we will replace, if needed, $\#G_q$ by q^3 in all asymptotic estimates.

Proof of Theorem 2.12. Slightly rephrasing Definition 2.2, we are going to prove that there exist positive numbers n_0, A, B, α, β, all independent of $g \in G_q$, such that for every $q > q_0 = p^{n_0}$ there exists $S_q \subset G_q$ with the following properties:

$$\text{(i) } \#S_q/q^3 < Aq^{-\alpha};$$

(1)

$$\text{(ii) for every } g \in T_q := G_q \setminus S_q \text{ we have } \left| \frac{\#W_{g,q}}{q^3} - 1 \right| < Bq^{-\beta}.$$

Indeed, this is enough for proving that w is p-equidistributed: in Definition 2.2 one can then take $\varepsilon_p(q) := \max\{Aq^{-\alpha}, Bq^{-\beta}\}$.

Towards this end, we will use the following commutative diagram:

(2)

$$\begin{array}{ccc} G_q \times G_q & \xrightarrow{P_{w,q}} & G_q \\ \pi \downarrow & & \downarrow \mathrm{tr} \\ \mathbb{A}^3_{s,u,t}(\mathbb{F}_q) & \xrightarrow{f_{w,q}} & \mathbb{A}^1_z(\mathbb{F}_q) \end{array}$$

where

$$\pi(x,y) = (\mathrm{tr}(x), \mathrm{tr}(xy), \mathrm{tr}(y)). \tag{3}$$

"Typical" fibres of the maps in this diagram should consist of $O(q^3)$ elements (for $P_{w,q}$ and π), and of $O(q^2)$ elements (for tr and $f_{w,q}$). Below we will show how to attain this with error term of order $O(q^{-\beta})$ by throwing away $O(q^{-\alpha})$ elements.

We will use an explicit Lang–Weil estimate of the following form: if $H \subset \mathbb{A}^3_{\mathbb{F}_q}$ is an *absolutely irreducible* hypersurface of degree d, then

$$|\#H(\mathbb{F}_q) - q^2| \le (d-1)(d-2)q^{3/2} + 12(d+3)^4 q$$

(see, e.g., [GL, Remark 11.3]), or, equivalently, $\#H(\mathbb{F}_q) = q^2(1+r_1)$ with

$$|r_1| \le q^{-1/2}[(d-1)(d-2) + 12(d+3)^4 q^{-1/2}]. \tag{4}$$

(The remainder term $r_1 = r_1(H)$, as well as all remainder terms in the sequel, depend on the hypersurface under consideration. To ease the notation, we do not include this dependence in formulas.)

For $d > 4$ and $q > 16$, equation (4) gives

$$|r_1| < q^{-1/2}(d^2 + 12\cdot 2^4 d^4/4) < d^4 q^{-1/2}(1/d^2 + 48) < 50d^4 q^{-1/2}. \tag{5}$$

Moreover, if $d > 4$ and $q > 4(50d^4)^2$, then $|r_1| < 1/2$. This remains true also for $d \le 3$. Without loss of generality, we may and will assume that the latter inequality is valid.

Step 1. Suppose that the polynomial f_w is p-noncomposite.

Denote the degree of f_w by d, the degree of the reduced polynomial $f_{w,p}$ is then at most d. Consider the corresponding reduced map $f_{w,p}\colon \mathbb{A}^3_{s,u,t}(\overline{\mathbb{F}}_p) \to \mathbb{A}^1_z(\overline{\mathbb{F}}_p)$.

Denote by $\sigma(f_{w,p})$ the spectrum of $f_{w,p}$, i.e., the set of all points $z \in \mathbb{A}^1_z(\overline{\mathbb{F}}_p)$ such that the hypersurface $H_z \subset \mathbb{A}^3_{s,u,t}(\overline{\mathbb{F}}_p)$, defined by the equation $f_w(s,u,t) = z$, is reducible. By a generalized Stein–Lorenzini inequality [Na], this set contains at most $d-1$ points. The same is true for each $\sigma_q(f_w) := \sigma(f_{w,p}) \cap \mathbb{F}_q$. Without loss of generality, we may and will assume that ± 2 are inside $\sigma_q(f_w)$ (by enlarging $\#\sigma_q(f_w)$ to $d+1$).

Let $z \in \mathbb{A}^1_z(\overline{\mathbb{F}}_p) \setminus \sigma(f_{w,p})$. Then H_z is an irreducible hypersurface and hence (4), (5) are valid for H_z.

Lemma 3.1. *Let $H \subset \mathbb{A}^3_{s,u,t}(\overline{\mathbb{F}}_p)$ be a hypersurface of degree d. Let $D(s,u,t) = (t^2-4)(s^2-4)(s^2+t^2+u^2-ust-4)$, and let $\Delta \subset \mathbb{A}^3_{s,u,t}$ be defined by the equation $D = 0$. Assume that $H \not\subset \Delta$. Then (see (3)) we have $\#\pi^{-1}(H)(\mathbb{F}_q) = \#H(\mathbb{F}_q)q^3(1+r_2)$, where $|r_2| < 157d/q$.*

Proof. We use the following fact (see [BG, Proposition 7.2]):

$$\#\pi^{-1}(s,u,t)(\mathbb{F}_q) = q^3(1+\delta_1(s,u,t)),\ |\delta_1| \le 3/q,$$

if $(s,u,t) \notin \Delta(\mathbb{F}_q)$, and

$$\#\pi^{-1}(s,u,t)(\mathbb{F}_q) \le 2q^3(1+1/q)$$

if $(s,u,t) \in \Delta(\mathbb{F}_q)$.

Denote $H \cap \Delta$ by H_Δ. By Bezout's theorem, this is a curve of degree at most $7d$, hence $\#H_\Delta(\mathbb{F}_q) \le 7d(q+1)$. We have

$$\begin{aligned}\#\pi^{-1}(H)(\mathbb{F}_q) &= \#\pi^{-1}(H \setminus H_\Delta)(\mathbb{F}_q) + \#\pi^{-1}(H_\Delta)(\mathbb{F}_q)\\ &\le \#(H \setminus H_\Delta)(\mathbb{F}_q) q^3(1+\alpha_1) + \#H_\Delta(\mathbb{F}_q) q^3 \alpha_2,\end{aligned}$$

where $|\alpha_1| \le 3/q$ and $|\alpha_2| \le 2(1+1/q) \le 3$. Thus

$$\frac{\#\pi^{-1}(H)(\mathbb{F}_q)}{\#H(\mathbb{F}_q)} = q^3 \left[\left(1 - \frac{\#H_\Delta(\mathbb{F}_q)}{\#H(\mathbb{F}_q)}\right)(1+\alpha_1) + \frac{\#H_\Delta(\mathbb{F}_q)}{\#H(\mathbb{F}_q)}\alpha_2\right] = q^3(1+r_2)$$

with

$$\begin{aligned}|r_2| &\le \frac{\#H_\Delta(\mathbb{F}_q)}{\#H(\mathbb{F}_q)}(1+|\alpha_1|+|\alpha_2|) + |\alpha_1| \le \frac{7d(q+1)}{q^2(1+r_1)}(1+|\alpha_1|+|\alpha_2|)+|\alpha_1|\\ &\le \frac{7d \cdot 2q \cdot (11/2)}{q^2/2} + \frac{3}{q} \le \frac{157d}{q}.\end{aligned}$$

□

Let S'_q be the set of all $z \in \mathbb{F}_q$ such that $H_z \subset \Delta$ (see Lemma 3.1). This set is finite, and $\#S'_q \le 7$ since Δ is of degree 7 and thus cannot contain more than 7 irreducible components.

Let $\tau\colon G_q \to \mathbb{A}^1$ be the trace map, $\tau(g) = \mathrm{tr}(g)$. We have $\#\tau^{-1}(z) \le q(q+1)$.

We define $\tilde{S}_q := \sigma_q(f_w) \cup S'_q$ and $S_q := \tau^{-1}(\tilde{S}_q)$. By construction,

$$\#S_q \le (d+8)q(q+1) \le q^3 \frac{2(d+8)}{q}.$$

According to Lemma 3.1, for any $z \in T_q$ we have

$$\#\pi^{-1}(H_z)(\mathbb{F}_q) = \#H_z(\mathbb{F}_q) q^3(1+r_2) = q^5(1+r_1)(1+r_2).$$

On the other hand, all $g \in G_q$ with $\mathrm{tr}(g) = z \in T_q$ are conjugate, and there are $\#\tau^{-1}(z) = q(q \pm 1)$ such elements. Hence for every such g (see diagram (2)), we have

$$\#W_g = \frac{\#\pi^{-1}(H_z)(\mathbb{F}_q)}{q(q\pm 1)} = \frac{q^5(1+r_1)(1+r_2)}{q(q \pm 1)} = q^3(1+r_3)$$

with

$$|r_3| \le 2(|r_1| + |r_2| + |r_1 r_2|) \le 2|r_1| + 3|r_2|.$$

Recall that $q \ge 4(50d^4)^2$, hence

$$|r_3| \le 2 \cdot 50 d^4 q^{-1/2} + 3 \cdot 157d/q \le q^{-1/2}(100d^4+1).$$

So for $q > q_0 = 4(50d^4)^2$, in equation (1) we can take

$$A = 2(d+8), \alpha = 1, B = 100d^4+1, \beta = 1/2. \tag{6}$$

Thus f_w is p-equidistributed.

Remark 3.2. Note that q_0 and all numbers in (6) depend only on w (through d, the degree of the trace polynomial f_w) and not on p.

Step 2. Suppose that the polynomial f_w is p-composite.

This means that $f_w(s,u,t) = h(Q(s,u,t))$ where $h \in \mathbb{F}_p[x]$ is a polynomial in one variable of degree $d_1 \ge 2$ and $Q \in \mathbb{F}_p[s,u,t]$ is a noncomposite polynomial in three variables.

Consider three separate cases.

Case 1. f_w is p-special, i.e., h is a permutation polynomial of all fields $\mathbb{F}_q$, $q = p^n$. For any $z \in \mathbb{F}_q$ there is a unique $x \in \mathbb{F}_q$ such that the hypersurface $H_z \subset \mathbb{A}^3$, defined by the equation $f_w(s,u,t) = z$, coincides with the hypersurface $\tilde{H}_x$, defined by the equation $Q(s,u,t) = x$. Since Q is noncomposite, **Step 1** implies that w is p-equidistributed in this case.

Remark 3.3. In this case, the parameters q_0, A, B, α, β also do not depend on p. They depend on the word w, this time through the degree of Q which is less than the degree of the trace polynomial of w.

Case 2. h is not a permutation polynomial for $\mathbb{F}_q$, $q = p^n$. Then it is not a permutation polynomial for any extension $\mathbb{F}_{q^m}$ of $\mathbb{F}_q$.

According to [Wa], [WSC], there exists a subset $U_m \subset \mathbb{A}^1_z(\mathbb{F}_{q^m})$ such that

- $\#U_m \geq (q^m - 1)/d_1$;
- $h^{-1}(s)(\mathbb{F}_{q^m}) = \emptyset$ for every $s \in U_m$.

It follows that $f_w^{-1}(s)(\mathbb{F}_{q^m}) = \emptyset$ for every m and every $s \in U_m$. So the polynomial $\pi \circ f_{w,q^m}$ also omits at least $(q^m - 1)/d_1$ values, and hence so does $P_{w,q^m} \circ \mathrm{tr}$ (see diagram (2)), i.e., $P_{w,q^m}(G_{q^m} \times G_{q^m})$ contains no elements $g \in G_{q^m}$ with $\mathrm{tr}(g) \in U_m$. For every $s \in \mathbb{F}_{q^m}$, $s \neq \pm 2$, the group G_{q^m} contains at least $(q^m)^2 - q^m$ elements with trace s. Thus P_{w,q^m} omits at least

$$q^m(q^m - 1)[(q^m - 1)/d_1 - 2] \approx (q^m)^3/d_1$$

values. Hence w is not p-equidistributed.

Case 3. h is a permutation polynomial for $\mathbb{F}_q$ but not for an extension $\mathbb{F}_{q^m}$. Then we can start with $\mathbb{F}_{q^m}$ and proceed as in **Case 2.**

Theorem 2.12 is proved. □

Proof of Theorem 2.13. If f_w is almost noncomposite, then, according to Remarks 3.2 and 3.3, the word w is equidistributed.

If f_w is very composite, then for some p it is p-composite but not p-special and, by Theorem 2.12, the word w is not p-equidistributed. Hence it is not equidistributed. □

Corollary 3.4. *Suppose that for each p and all n big enough the image of the map $P_{w,p^n}\colon \mathrm{SL}(2,p^n) \times \mathrm{SL}(2,p^n) \to \mathrm{SL}(2,p^n)$ contains all noncentral semisimple elements of* $\mathrm{SL}(2,p^n)$*. Then w is equidistributed.*

Proof. Assume the contrary. Then by Theorem 2.13, the polynomial f_w is very composite, i.e., for some p it is p-composite but not p-special. As in Case 2 considered above, we see that for big n the polynomial f_{w,p^n} omits at least $(p^n - 1)/d_1$ values. This contradicts the assumption of the corollary according to which f_{w,p^n} omits at most two values, 2 and -2. □

Remark 3.5. The converse statement is not true. Indeed, let $f_w = \mathrm{tr}(w(x,y))$ be the trace polynomial of a word $w(x,y)$. Let $a \in \mathbb{Q}$ be a rational point in the spectrum $\sigma(f_w)$, which means that the surface H_a, given by the equation $f_w(s,u,t) = a$, is not absolutely irreducible. Then for all p big enough the reduction a_p lies in $\sigma(f_{w,p})$ (see [BDN, 2.2.1]). It follows that the set of numbers q such that $H_a(\mathbb{F}_q) = \emptyset$ may be infinite.

Examples of such words were provided by Jambor, Liebeck and O'Brien [JLO]. For instance, let $w(x,y) = x^2(x^2yx^{-2}y^{-1})^2$. Let us show that the trace polynomial

f_w is $\mathbb{Q}$-noncomposite. Assume to the contrary that f_w is $\mathbb{Q}$-composite. Then, according to Proposition 4.10 below, $f_w(s,u,t) = D_2(p(s,u,t))$ for some polynomial p, where $D_2(z) = z^2 - 2$ is the second Dickson polynomial. We conclude that $f_w + 2 = p^2$. However, the factorization of f_w (say, on MAGMA) shows that f_w is not a full square.

On the other hand, $0 \in \sigma(f_w)$ (see, e.g., [JLO, Lemma 2.2]). It is shown in [JLO] that $H_0(\mathbb{F}_q) = \emptyset$ for every q such that

- $q = p^{2r+1}, \quad r \geq 0$;
- $p \neq 5$;
- $p^2 \not\equiv 1 (\mathrm{mod} 16)$;
- $p^2 \not\equiv 1 (\mathrm{mod} 5)$.

Thus, for these q, the morphisms $P_{w,q}$ are dominant and equidistributed whereas the elements with zero trace are not in the range of $P_{w,q}$.

Remark 3.6. Note that many positive words $w = x^a y^b$, $a > 0$, $b > 0$, satisfy the assumptions of Corollary 3.4 and are equidistributed, see [BG] for details (and [LST] for generalizations to simple groups of higher Lie rank).

Proof of Proposition 2.15. We may assume that q is odd. Consider the commutative diagram

$$\begin{array}{ccc} G_q \times G_q & \xrightarrow{P_{w,q}} & G_q \\ \downarrow \rho' & \searrow \varkappa & \downarrow \rho \\ \hat{G}_q \times \hat{G}_q & \xrightarrow{\hat{P}_{w,q}} & \hat{G}_q \end{array}$$

where ρ and ρ' are natural projections, and $P_{w,q}$ and $\hat{P}_{w,q}$ correspond to the map $(x,y) \to w(x,y)$ on $G_q \times G_q$ and on $\hat{G}_q \times \hat{G}_q$, respectively.

Suppose w is p-equidistributed with respect to $\{G_q\}$ so that for $q > q_0$ we have inequalities (1) with parameters A, B, α, β. Define $\hat{S}_q := \rho(S_q)$, $\hat{T}_q := \hat{G}_q \setminus \hat{S}_q$.

For any element $\hat{g} \in \hat{G}_q$ the set $\rho^{-1}(\hat{g})$ contains precisely two elements g_1, g_2 of G_q. Therefore,

- $\#\hat{S}_q = \#S_q/2 = \#G_q \varepsilon_p(q)/2 = \#\hat{G}_q \varepsilon_p(q)$;
- $W_{\hat{g},\hat{G}_q} = \rho'(W_{g_1,G_q} \cup W_{g_2,G_q})$;
- $\#W_{\hat{g},\hat{G}_q} = (\#W_{g_1,G_q} + \#W_{g_2,G_q})/4$;
- for every $\hat{g} \in \hat{T}_q$ we have

$$\#W_{\hat{g},\hat{G}_q} = \frac{\#W_{g_1,G_q} + \#W_{g_2,G_q}}{4} = \#G_q \frac{1 + \varepsilon_p(q)}{2} = \#\hat{G}_q(1 + \varepsilon_p(q)).$$

Hence, w is p-equidistributed on $\{\hat{G}\}_q$ with the same parameters as on $\{G_q\}$. $\square$

Remark 3.7. In [GS] there is a discussion on relationship between two close properties of word maps on finite groups: be equidistributed and preserve the uniform measure. In our context, the proof of Theorem 2.13 allows us to formulate this relationship explicitly.

Corollary 3.8. *Assume that a word w has an almost noncomposite trace polynomial f_w of degree d. Let $q > 4(50d^4)^2$, and let $\varepsilon(d,q) = 3(100d^4 + 1)q^{-1/2}$. Let*

$G = \mathrm{SL}(2, q)$ or $G = \mathrm{PSL}(2, q)$. Then the word map $w\colon G \times G \to G$ is $\varepsilon(d, q)$-measure-preserving in the sense of [GS].

Proof. According to (6), the word map w is $\varepsilon(d, q)/3$-equidistributed, in the sense of Definition 2.1, and hence $\varepsilon(d, q)$-measure-preserving, by [GS, Proposition 3.2]. □

Corollary 2.14 will be proved in Section 4.

4. Composite trace polynomials

Our goal in this section is to describe words in two variables whose trace polynomial is composite. A full description could provide an answer, in the case of SL(2) and words in two variables, to the following basic question, which should apparently be attributed to Larsen and Shalev:

Question 4.1. Is it true that a word $w \in F_d$ is equidistributed on a Chevalley group G (of fixed type) if and only if w is not a proper power of another word?

Although our results (summarized in Table 1) are not conclusive, they give a strong evidence in favour of an affirmative answer to Question 4.1 in our case. Before explaining the table, we give some necessary preliminaries.

Throughout this section $D_n(x)$ stands for the n^{th} Dickson polynomial (see Example 2.11). It is well known (see, e.g., [LMT, (2.2)]) that this polynomial satisfies $D_n(x+1/x) = x^n+1/x^n$ and is completely determined by this functional equation.

For the sake of convenience, we define $D_{-n}(x) = D_n(x)$ and $D_0(x) \equiv 2$. We repeatedly use the decomposition $D_{nm}(x) = D_n(D_m(x))$.

Notation 4.2. We always assume that $w(x, y)$ is written in the form

$$w = x^{a_1}y^{b_1}\dots x^{a_r}y^{b_r} \tag{7}$$

and is reduced (all integers a_i, b_j are nonzero). We call the integer r the *complexity* of w.

If $\mathbb{F}$ is a field and $f_w \in \mathbb{Z}[s, u, t]$ is the trace polynomial of w, we keep the same notation for the polynomial $f_w \in \mathbb{F}[s, u, t]$ obtained after changing scalars to $\mathbb{Z}\otimes\mathbb{F}$.

We denote

- $A = A(w) := \sum\limits_{i=1}^{r} a_i$, $B = B(w) := \sum\limits_{i=1}^{r} b_i$;
- $\bar{A} = \bar{A}(w) := \sum\limits_{i=1}^{r} |a_i|$, $\bar{B} = \bar{B}(w) := \sum\limits_{i=1}^{r} |b_i|$;
- for a polynomial $p(x_1, \dots, x_n)$ we denote by $\deg_{x_i} p$ the degree of p with respect to the variable x_i.

Definition 4.3. Let $w = x^{a_1}y^{b_1}\dots x^{a_r}y^{b_r}$ and $v = x^{c_1}y^{d_1}\dots x^{c_{r'}}y^{d_{r'}}$ be reduced words written in form (7). We say that they are *trace-similar*, and denote this by $w \approx v$, if $r = r'$, the array $\{|a_i|\}$ is a rearrangement of $\{|c_i|\}$, and the array $\{|b_i|\}$ is a rearrangement of $\{|d_i|\}$.

Proposition 4.4. [Ho] *If reduced words written in form* (7) *have the same trace polynomial, then they are trace-similar.*

Example 4.5. The words $w = xy$ and $v = xy^{-1}$ are trace-similar but have different trace polynomials: $\mathrm{tr}(w) = u$, $\mathrm{tr}(v) = st - u$. Moreover, the value sets of the trace polynomials of trace-similar words may differ: let, say, $w = (xy)^2$ and $v = [x, y]$;

the words w and v are trace-similar but P_v is surjective on $SL(2,q)$ whereas P_w is not if q is odd.

The words $x^2y^{-1}xy$ and x^2yxy^{-1} are trace-similar, have the same trace polynomial but are not conjugate in F_2 [Ho].

We can now explain Table 1. It gives conditions under which one can conclude that if the trace polynomial f_w is composite then w is a proper power of another word (or is trace-similar to such a power). These conditions depend on relations between the degree n of the polynomial h appearing in the decomposition of f_w, the complexity r of the word w (these relations are put in the first column of the table), and the characteristic p of the ground field (which is put in the first row). The entries of the table contain conclusions on w and references to the corresponding assertions.

$f_w = h \circ q$ $\deg h = n$, compl.$= r$	$\mathbb{F} = \mathbb{Q}$	$\mathbb{F} = \mathbb{F}_p$, $p > r$	$\mathbb{F} = \mathbb{F}_p$, $p > r/2$, $p \neq r$
$n < r$	$w \approx v(x,y)^n$ Prop. 4.11	?	?
$n = r$ and ($A \neq 0$ or $B \neq 0$)	$w = (x^\alpha y^\beta)^r$ Cor. 4.16 Prop. 4.10	$w = (x^\alpha y^\beta)^r$ Prop. 4.19 Prop. 4.10	$(w \approx v^r) \Rightarrow (w = (x^\alpha y^\beta)^r)$ Prop. 4.14 Prop. 4.10
$n = r$, r prime	$w = (x^\alpha y^\beta)^r$ Cor. 4.16 Prop. 4.17 Prop. 4.10	$w = (x^\alpha y^\beta)^r$ Cor. 4.18	?

TABLE 1. Words with composite trace polynomial

Proposition 4.6. [BG] *Let w be a reduced word written in form* (7)*, and let $w_i = x^{a_i}y^{b_i}$. Then for the trace polynomials we have $f_{w_i}(s,u,t) = ug_{a_i,b_i}(s,t) + h_{a_i,b_i}(s,t)$, $\deg_s g_{a_i,b_i} = |a_i| - 1$, $\deg_t g_{a_i,b_i} = |b_i| - 1$. Moreover, if $\mathbb{F}$ is of characteristic zero or big enough, then $g_{a_i,b_i}(s,t) \not\equiv 0$ and*

$$f_w(s,u,t) = \sum_{k=0}^{r} u^k G_k(s,t) \text{ where } G_r(s,t) = \prod_{i=1}^{r} g_{a_i,b_i}(s,t). \tag{8}$$

In particular, $\deg_s G_r = \bar{A} - r$, $\deg_t G_r = \bar{B} - r$.

Proposition 4.7. [Ri], [Tu], [GC] *Let $\mathbb{F}$ be either $\overline{\mathbb{Q}}$ or $\overline{\mathbb{F}}_q$, and let n be a positive integer. If $p = char(\mathbb{F}) > 0$, assume that $(n,p) = 1$. Suppose that $D_n(x)$ is $\mathbb{F}$-composite, $D_n(x) = h(g(x))$. Then $h(x) = D_m(x-c)$ and $g(x) = D_k(x) + c$, where $km = n$ and $c \in \mathbb{F}$.*

Remark 4.8. The statement of Proposition 4.7 remains valid if p divides n. Indeed, suppose that $n = kp^s$, $(k,p) = 1$, $s \geq 1$. Write h in the form $h(y) = (h_1(y))^{p^t}$, where $h_1' \not\equiv 0$ (t may be zero). Denote $r = \deg h$, $r_1 = \deg h_1$, then $r = r_1 p^t \mid kp^s$, hence $t \leq s$.

Since $D_n(x) = D_k(x)^{p^s} = (h_1(g(x)))^{p^t}$, we have $D_k(x)^{p^{s-t}} = D_k(x^{p^{s-t}}) = \varepsilon h_1(g(x))$, where $\varepsilon^{p^t} = 1$. If $s > t$, then the derivative of the left-hand side is identically zero, and since $h_1' \not\equiv 0$, we have $g(x) = g_1(x^{p^{s-t}})$. Let $z = x^{p^{s-t}}$. Then $D_k(z) = \varepsilon h_1(g(z))$. Since $(k,p) = 1$, by [GC] we have $\varepsilon h_1(z) = D_{r_1}(z-c)$, $g(z) = D_{k/r_1}(z) + c$. Therefore, $h(z) = (h_1(y))^{p^t} = (D_{r_1}(z-c))^{p^t} = D_{r_1 p^t}(z-c)$. If $s = t$, then $D_k(x) = \varepsilon h_1(g(x))$, and we are under hypotheses of Proposition 4.7.

Remark 4.9. We may and will assume (see [Tu]) that $h(2) = 2$ which corresponds to $c = 0$.

Further on we assume that $\mathbb{F}$ is either $\mathbb{Q}$ or $\mathbb{F}_p$ (or the respective algebraic closure, if needed).

Let $q(s,u,t) = uG(s,t) + H(s,t)$. Assume that

$$q(s,2,s) = 2G(s,s) + H(s,s) = g_1(s) + c \tag{9}$$

and

$$q(s,s^2-2,s) = (s^2-2)G(s,s) + H(s,s) = g_2(s) + c. \tag{10}$$

Then, if $s \neq \pm 2$, we have

$$G(s,s) = \frac{g_2(s) - g_1(s)}{s^2-4}, \tag{11}$$

$$H(s,s) = \frac{(s^2-2)g_1(s) - 2g_2(s)}{s^2-4} + c. \tag{12}$$

Indeed, computing $G(s,s)$, $H(s,s)$ from (9) and (10), we obtain (11) and (12).

In particular, let $w'(x,y) = x^a y^b$ and $f_{w'}(s,u,t) = u g_{a,b}(s,t) + h_{a,b}(s,t)$. Then we have

$$\operatorname{tr} x^a x^{-b} = 2g_{a,b}(s,s) + h_{a,b}(s,s) = D_{a-b}(s),$$
$$\operatorname{tr} x^a x^b = (s^2-2)g_{a,b}(s,s) + h_{a,b}(s,s) = D_{a+b}(s),$$

and, according to (11), (12), for $s \neq \pm 2$ we obtain

$$g_{a,b}(s,s) = \frac{D_{a+b}(s) - D_{a-b}(s)}{s^2-4},$$
$$h_{a,b}(s,s) = \frac{(s^2-2)D_{a-b}(s) - 2D_{a+b}(s)}{s^2-4}.$$

Put $s = x + x^{-1}$, then

$$\begin{aligned} g_{a,b}(s,s) &= \frac{(x^{a+b} + x^{-(a+b)}) - (x^{a-b} + x^{-(a-b)})}{(x-x^{-1})^2} \\ &= \frac{(x^a - x^{-a})(x^b - x^{-b})}{(x-x^{-1})^2}. \end{aligned} \tag{13}$$

Proposition 4.10. *With Notation* 4.2*, assume that either* $A \neq 0$ *or* $B \neq 0$*. Suppose that* $f_w(s,u,t) = h(q(s,u,t))$ *where* $q \in \mathbb{F}[s,u,t]$ *and* $h \in \mathbb{F}[z]$, $\deg h \geq 2$*. Then* $h = D_d(z)$ *with* $d \geq 2$ *dividing both* A *and* B*.*

Proof. Putting $y = \mathrm{id}$, $x = \mathrm{id}$, $x = y^{-1}$, and $x = y$, we get, respectively (taking into account that $\operatorname{tr}(g^{-1}) = \operatorname{tr}(g)$):

$$f_w(s,s,2) = h(q(s,s,2)) = D_A(s),$$
$$f_w(2,t,t) = h(q(2,t,t)) = D_B(t),$$
$$f_w(s,2,s) = h(q(s,2,s)) = D_{A-B}(s),$$
$$f_w(s,s^2-2,s) = h(q(s,s^2-2,s)) = D_{A+B}(s).$$

These decompositions, together with Proposition 4.7, Remark 4.8 and the condition $\deg h \geq 2$, imply that there is a common divisor $d \geq 2$ of all the nonzero numbers from the list A, B, $A-B$, $A+B$ such that $h(z) = D_d(z)$. □

Proposition 4.11. *With Notation 4.2, suppose that $f_w(s,u,t)$ is $\mathbb{F}$-composite, $f_w(s,u,t) = h(q(s,u,t))$, where $q \in \mathbb{F}[s,u,t]$ and $h(x) = \mu x^n + \dots$ is a polynomial in one variable of degree n, $\mu \neq 0$. Then $r = nm$. Moreover, if the characteristic of $\mathbb{F}$ is 0 or big enough, $w(x,y)$ is trace-similar to a word $v(x,y)^n$ where the complexity of v is m.*

Proof. Let $q(s,u,t) = \sum\limits_{k=0}^{m} u^k H_k(s,t)$. Then

$$f_w(s,u,t) = h(q(s,u,t)) = \mu u^{mn} H_m^n(s,t) + \Phi(s,u,t)$$

where $\deg_u \Phi(s,u,t) < mn$. Hence $r = nm$ and

$$\mu H_m^n(s,t) = G_r(s,t) = \prod_{i=1}^{r} g_{a_i,b_i}(s,t)$$

(we use the notation of Proposition 4.6, in particular, formula (8)).

Therefore, by (13), we have

$$\mu H_m^n(s,s) = \pm \frac{\prod\limits_{i=1}^{r} (x^{|a_i|} - x^{-|a_i|})(x^{|b_i|} - x^{-|b_i|})}{(x-x^{-1})^{2r}}.$$

Let $p = \mathrm{char}(F)$. If $p > 0$, write $|a_i| = \tilde{a}_i p^{\alpha_i}$, $|b_j| = \tilde{b}_j p^{\beta_j}$, with $(\tilde{a}_i, p) = (\tilde{b}_j, p) = 1$. If $p = 0$, set $|a_i| = \tilde{a}_i$, $|b_j| = \tilde{b}_j$.

Choose an integer $N > \max\{|\tilde{b}_i|\}$ such that $(N,p) = 1$, and consider the word $w_N = w(x^N, y)$. Then

$$f_{w_N} = f_w(D_N(s), \delta(s,u,t), t) = h(q(D_N(s), \delta(s,u,t), t))$$

where $\delta(s,u,t) = \mathrm{tr}(x^N y) = u g_{N,1}(s,t) + h_{N,1}(s,t)$. Thus

$$q(D_N(s), \delta(s,u,t), t) =: q_1(s,u,t) = \sum_{k=0}^{m} u^k F_k(s,t)$$

and

$$f_{w_N}(s,u,t) = h(q_1(s,u,t)) = \mu u^{mn} F_m^n(s,t) + \Phi_1(u,s,t)$$

where $\deg_u \Phi_1(u,s,t) < mn$. Hence, since the words w and w_N have the same complexity r, we have

$$\mu F_m^n(s,s) = \pm \frac{\prod\limits_{i=1}^{r} (x^{N|a_i|} - x^{-N|a_i|})(x^{|b_i|} - x^{-|b_i|})}{(x-x^{-1})^{2r}}. \tag{14}$$

Fix an integer $i \in \{1, \dots, r\}$. Let $x_0 \neq 1$ denote a simple root of the equation $x^{N\tilde{a}_i} - 1 = 0$. If p is odd, the order of zero $o(x_0)$ of the product in the right-hand side of (14) is equal to the number $\sum_{k \geq 0} n_i(k) p^k$ where $n_i(k)$ denotes the number of appearances of $|\tilde{a}_i| p^k$ in the list $|a_1|, \dots, |a_r|$. On the other hand, $o(x_0) = n\varkappa_i$ where $\varkappa_i$ is the order of the zero of F_m at the point $x_0 + 1/x_0$. The same is true for $p = 0$ (if we set $0^0 = 1$).

Assume that $p > \max\{|a_i|, |b_j|, 1 \leq i,j \leq r\}$, or $p = 0$. Then $n_i(k) = 0$ for $k > 0$. This means that there are $n\varkappa_i$ appearances of each $|a_i|$ in the list $|a_1|, \dots, |a_r|$.

In a similar way, looking at the word $w(x, y^M)$ for M big enough and prime to p, we conclude that there are precisely $n\tau_i$ appearances of each $|b_i|$ in the list $|b_1|, \dots, |b_r|$ where τ_i is the order of the corresponding root of F_m. Moreover, $\sum \varkappa_i = \sum \tau_i = m$.

Define a word

$$v(x,y) = x^{t_1} y^{k_1} \dots x^{t_m} y^{k_m}$$

of complexity m in such a way that among the $|t_i|$ there will be $\varkappa_i$ of the $|a_i|$ and among the $|k_i|$ there will be τ_i of the $|b_i|$. By construction, $v^n(x,y)$ is trace-similar to $w(x,y)$ which completes the proof. □

Remark 4.12. In contrast with Proposition 4.10, in Proposition 4.11 we do not exclude the case $A = B = 0$.

In some particular cases, Proposition 4.11 provides even more information.

Proposition 4.13. *With the notation and assumptions of Proposition* 4.11*, assume also that*

- $n = r$;
- $A \neq 0$ *or* $B \neq 0$;
- *char* $\mathbb{F} = p > 0$;
- $p > r$.

Then $w(x,y) = (x^\alpha y^\beta)^r$ *where* $\alpha = A/r$, $\beta = B/r$.

Proof. First note that under the hypotheses of the proposition, the assumptions of Proposition 4.10 are also satisfied. In particular, both A and B are divisible by r and hence α and β are integers. We also have

$$f_w = D_r(q(s,u,t)),\ q(s,u,t) = G(s,t)u + H(s,t),$$
$$q(s,2,s) = D_{\alpha-\beta}(s),\ q(s,s^2-2,s) = D_{\alpha+\beta}(s).$$

Hence, similarly to (13), we have:

$$G(s,s) = \pm \frac{(x^{|\alpha|} - x^{-|\alpha|})(x^{|\beta|} - x^{-|\beta|})}{(x - x^{-1})^2}.$$

It follows that for any N, M we have

$$\left(\frac{(x^{N|\alpha|} - x^{-N|\alpha|})(x^{M|\beta|} - x^{-M|\beta|})}{(x - x^{-1})^2} \right)^r = \pm \frac{\prod_{i=1}^r (x^{N|a_i|} - x^{-N|a_i|})(x^{M|b_i|} - x^{-M|b_i|})}{(x - x^{-1})^{2r}}.$$

Hence,

$$(x^{|\alpha|} - x^{-|\alpha|})^r = \pm \prod_{i=1}^r (x^{|a_i|} - x^{-|a_i|}),$$
$$(x^{|\beta|} - x^{-|\beta|})^r = \pm \prod_{i=1}^r (x^{|b_i|} - x^{-|b_i|}).$$

Comparing the degrees of the corresponding polynomials, we get

$$|A| = |\alpha| r = \sum_{i=1}^r |a_i| = \bar{A},$$
$$|B| = |\beta| r = \sum_{j=1}^r |b_j| = \bar{B}.$$

Hence, all the a_i are of the same sign, and so are all the b_j. Let $|\alpha| = \tilde{\alpha}p^{\tau}$, $|\beta| = \tilde{\beta}p^{\varkappa}$. Comparing simple roots of the polynomials, we get

$$|a_i| = \tilde{\alpha}p^{k_i},\ |b_j| = \tilde{\beta}p^{s_i}$$

for every $1 \le i \le r$ and every $1 \le j \le r$. Moreover,

$$\bar{A} = \tilde{\alpha}p^{\tau}r = \tilde{\alpha}\sum_{k\ge 0} n(k)p^k, \tag{15}$$

where $n(k)$ denotes the number of appearances of $|\tilde{\alpha}|p^k$ in the list $|a_1|, \dots, |a_r|$, and

$$\bar{B} = \tilde{\beta}p^{\varkappa}r = \tilde{\beta}\sum_{k\ge 0} m(k)p^k$$

where $m(k)$ denotes the number of appearances of $|\tilde{\beta}|p^k$ in the list $|b_1|, \dots, |b_r|$.

Consider formula (15). Let $K = \max\{k \mid n(k) \ne 0\}$. Suppose that $\tau > 0$. Then

$$p^{\tau}r = \sum_{k=0}^{K} n(k)p^k \le \left(\sum_{k=0}^{K} n(k)\right) p^K = rp^K.$$

Thus $\tau \le K$. It follows that $p^{\tau} \mid \sum_{k=0}^{\tau-1} n(k)p^k$, and hence the latter sum equals sp^{τ} for some integer s. On the other hand,

$$sp^{\tau} = \sum_{k=0}^{\tau-1} n(k)p^k \le \left(\sum_{k=0}^{\tau-1} n(k)\right) p^{\tau-1} \le rp^{\tau-1} < p^{\tau}.$$

Contradiction shows that $s = 0$, and $n(k) = 0$ for $k < \tau$. Dividing (15) by p^{τ}, we get

$$\sum_{k=\tau}^{K} n(k)p^{k-\tau} = r.$$

This equality remains true in the case $\tau = 0$. On the other hand, by the definition of $n(k)$, we have

$$\sum_{k=\tau}^{K} n(k) = \sum_{k=0}^{K} n(k) = r.$$

Hence $n(k) = 0$ for $k \ne \tau$, i.e., $K = \tau$, $n(\tau) = r$, $|a_i| = \tilde{\alpha}p^{\tau} = \alpha$. Since the a_i are of the same sign, they are all equal. In a similar way, we conclude that all $b_j = \tilde{\beta}p^{\varkappa} = \beta$ are equal. Hence $w(x,y) = (x^{\alpha}y^{\beta})^r$. □

Proposition 4.14. *Let $w(x,y) = x^a y^b \dots$ be a reduced word of complexity r such that $f_w(s,u,t) = D_r(q(s,u,t))$, $q \in \mathbb{F}[s,u,t]$, over $\mathbb{F} = \mathbb{Q}$ or some $\mathbb{F} = \mathbb{F}_p$ with $p > r/2$, $p \ne r$. If $w(x,y)$ is trace-similar to $(x^a y^b)^r$, then $w(x,y) = (x^a y^b)^r$.*

Proof. By assumption, w is the product of syllables $x^{\pm a}y^{\pm b}$.

Assume that by cyclic permutation and exchanging roles of x and y one can modify w to a word $v = v_1 \dots v_r$, $v_k = x^{\pm a}y^{\pm b}$, $k = 1, \dots, r$, which contains repeated syllables, i.e., such that for some $i < j$ we have $v_i = v_j$. Then we consider the word

$$\tilde{v} = v_i \dots v_j \dots v_r v_1 \dots v_{i-1}.$$

The word $\tilde{v}$ will be called a *convenient* form of w. Note that either $f_w(s,u,t) = f_{\tilde{v}}(s,u,t)$ or $f_w(s,u,t) = f_{\tilde{v}}(t,u,s)$. If this procedure is impossible, we say that w is already in a convenient form. First consider the case where $a = b = 1$ and $\mathbb{F} = \mathbb{Q}$.

Lemma 4.15. *Let $w(x,y) = xy \dots x^{\pm 1}y^{\pm 1} = w_1 \dots w_r$ be a word in a convenient form, where w_i are syllables of the form $x^{\pm 1}y^{\pm 1}$. Let $\mathbb{F} = \mathbb{Q}$.*

Let $u = \mathrm{tr}(xy), s = \mathrm{tr}(x), t = \mathrm{tr}(y)$. Then

$$f_w(s,u,t) = \epsilon u^r - \epsilon mstu^{r-1} + \dots + g(s,t)$$

is a polynomial of degree r with respect to u such that

- *the coefficient at u^r is $\epsilon = \pm 1$;*
- *m is a non-negative integer, $m \le r/2$, and $m = 0$ if and only if $w = (xy)^r$;*
- *the coefficient at u^{r-1} is ϵmst;*
- *the coefficient $f_w(s,0,t)$ at u^0 is a polynomial g in s,t of total degree strictly less than $2r$.*

It is important here that we defined u as the trace of the first syllable.

Proof of Lemma 4.15. First consider the case when there are no repeated syllables.

r=1: $\mathrm{tr}(xy) = u$.

r=2:
- $\mathrm{tr}(xyxy^{-1}) = -u^2 + ust - t^2 + 2$,
- $\mathrm{tr}(xyx^{-1}y^{-1}) = u^2 - ust + t^2 + s^2 - 2$,
- $\mathrm{tr}(xyx^{-1}y) = \mathrm{tr}(yxyx^{-1}) = -u^2 + ust - s^2 + 2$.

r=3:
- the words $a_1 = x\mathbf{y}\mathbf{x}^{-1}\mathbf{y}\mathbf{x}^{-1}y^{-1}$, $a_2 = \mathbf{x}\mathbf{y}\mathbf{x}y^{-1}x^{-1}\mathbf{y}$ and $a_3 = \mathbf{x}yx^{-1}\mathbf{y}^{-1}\mathbf{x}\mathbf{y}^{-1}$ are not in a convenient form;
- for $a_4 = xyxy^{-1}x^{-1}y^{-1}$, we have

$$\begin{aligned}\mathrm{tr}(a_4) = f_{a_4}(s,u,t) &= (ust - u^2 - t^2 + 2)u - \mathrm{tr}(x^3y)\\ &= (ust - u^2 - t^2 + 2)u - u(s^2 - 2) + (st - u)\\ &= -u^3 + stu^2 + u(3 - t^2 - s^2) + st;\end{aligned}$$

- the word $a_5 = xyx^{-1}y^{-1}x^{-1}y$ may be modified to a_4 by cyclic permutation and exchanging roles of x and y, thus $\mathrm{tr}(a_5) = \mathrm{tr}(a_4)$;
- $a_6 = xyx^{-1}yxy^{-1}$ may be modified to a_4 by cyclic permutation and changing roles of y and y^{-1}, thus

$$\mathrm{tr}(a_6) = f_{a_4}(s, st - u, t) = u^3 - u^2st + u(s^2t^2 + t^2 + s^2 - 3) + st(4 - t^2 - s^2).$$

r=4:
- $b_1 = \mathbf{x}\mathbf{y}\mathbf{x}y^{-1}x^{-1}y^{-1}x^{-1}\mathbf{y}$ and $b_2 = x\mathbf{y}\mathbf{x}^{-1}\mathbf{y}\mathbf{x}^{-1}y^{-1}xy^{-1}$ are not in a convenient form;
- for $b_3 = xyxy^{-1}x^{-1}yx^{-1}y^{-1}$ we have

$$\begin{aligned}\mathrm{tr}(b_3) &= (ust - u^2 - t^2 + 2)^2 - \mathrm{tr}(x^2yx^2y^{-1}) = (ust - u^2 - t^2 + 2)^2\\ &- [(us - t)(s^2 - 2)t - (us - t)^2 - t^2 + 2]\\ &= u^4 - 2u^3st + u^2h_1 + uh_2 + (t^2 - 2)^2 + t^2(s^2 - 2) + 2t^2 - 2,\end{aligned}$$

where h_1, h_2 are polynomials in s, t;
- $b_4 = xyx^{-1}yxy^{-1}x^{-1}y^{-1}$ may be modified to b_3 by cyclic permutation, and substituting x by y^{-1} and y by x^{-1};
- $b_5 = xyx^{-1}y^{-1}xy^{-1}x^{-1}y$ may be modified to b_3 by cyclic permutation, and substituting x by y and y by x;
- $b_6 = xyx^{-1}y^{-1}x^{-1}yxy^{-1}$ may be modified to b_3 by cyclic permutation, and substituting x by x^{-1} and y by y^{-1}.

Note that these substitutions do not change u, and the coefficient m is not zero in convenient words.

Any word of complexity ≥ 5 must have repeated syllables. The case with repeated syllables will be proved by induction on the complexity r. Assume that for all words in a convenient form of complexity $k < r$ the statement of the lemma is valid.

Consider $w(x,y) = w_1 \dots w_r$ where $w_1 = xy$, $w_i = x^{\pm 1}y^{\pm 1}$, $i = 2, \dots, r$, $w_{j+1} = w_1$, $0 < j \leq r-1$. Thus $w = v_1 v_2$ where $v_1 = w_1 \dots w_j$, $v_2 = w_{j+1} \dots w_r$. Denote $v_3 = v_1 v_2^{-1}$, it is of complexity $r-2$ since its first syllable is xy and the last is $(xy)^{-1}$. By induction hypothesis,

$$\operatorname{tr}(v_1) = \epsilon_1 u^j - \epsilon_1 m_1 stu^{j-1} + \dots + g_1, \deg g_1 < 2j;$$
$$\operatorname{tr}(v_2) = \epsilon_2 u^{r-j} - \epsilon_2 m_2 stu^{r-j-1} + \dots + g_2, \deg g_2 < 2(r-j).$$

The word v_3 may not be in a convenient form. This means that $u = \operatorname{tr}(xy)$ may not be the trace of the first syllable of v_3. Anyway,

$$\operatorname{tr}(v_3) = \epsilon_3 \hat{u}^{r-2} - \epsilon_3 m_3 st\hat{u}^{r-3} + \dots + g_3, \deg g_3 < 2(r-2),$$

where $\hat{u}$ is either u or $st - u$. In both cases its degree with respect to u is at most $r-2$ and the coefficient at u^0 is of total degree at most $2(r-2)$. Therefore

$$\begin{aligned}\operatorname{tr}(w) &= \operatorname{tr}(v_1)\operatorname{tr}(v_2) - \operatorname{tr}(v_3) \\ &= \epsilon_1\epsilon_2 u^r - \epsilon_1\epsilon_2 st(m_1+m_2)u^{r-1} + \dots + g_1 g_2 - g_3.\end{aligned}$$

Here the total degree of the polynomial $g_1 g_2 - g_3$, which is the coefficient at u^0, is less than $2j + 2(r-j) = 2r$. Moreover, $m_1 + m_2$ may be zero only if $m_1 = m_2 = 0$, which means, by induction hypothesis, that $v_1 = w_1^j$, $v_2 = w_1^{r-j}$, so $w = w_1^r$. □

We continue the proof of Proposition 4.14: assume that $\mathbb{F} = \mathbb{Q}$ or $\mathbb{F}_p$, $p > r$. Assume that $w(x,y) = w_1 \dots w_r$, where $w_1 = x^a y^b$, $w_i = x^{\pm a} y^{\pm b}$, is written in a convenient form, and $f_w(s,u,t) = D_r(q(s,u,t))$. We denote $z = x^a, v = y^b$, i.e., $w(x,y) = \tilde{w}(z,v)$, and $\tilde{w}$ is a word of the type considered in Lemma 4.15. Let $\tilde{s} = D_a(s)$, $\tilde{t} = D_b(t)$, and $\tilde{u} = \operatorname{tr}(x^a y^b) = u g_{a,b}(s,t) + h_{a,b}(s,t)$, where $g_{a,b}, h_{a,b}$ are polynomials in s,t and $g_{a,b} \not\equiv 0$ (see Proposition 4.6). Since the polynomial $q(s,u,t)$ is of degree 1 with respect to u, we have $q(s,u,t) = \alpha(s,t)\tilde{u} + \beta(s,t)$, with rational coefficients α and β. According to Lemma 4.15, we have

$$\begin{aligned}f_w(s,u,t) &= \epsilon\tilde{u}^r - \epsilon m\tilde{s}\tilde{t}\tilde{u}^{r-1} + \dots + g(\tilde{s},\tilde{t}) = q^r - rq^{r-2} + \dots \\ &= (\alpha(s,t)\tilde{u} + \beta(s,t))^r - r(\alpha(s,t)\tilde{u} + \beta(s,t))^{r-2} + \dots\end{aligned} \tag{16}$$

Moreover, if $m \neq 0$ then $m \not\equiv 0 \pmod p$, since $m \leq r/2 < p$. It follows that

$$\alpha(s,t) = \alpha = \text{const},\ \alpha^r = \epsilon,\ \text{and}\ \beta(s,t) = -\frac{\epsilon m\tilde{s}\tilde{t}}{r\alpha^{r-1}} = -\frac{m\alpha\tilde{s}\tilde{t}}{r}$$

(division is legitimate because $p \neq r$). Substituting $q = \alpha\tilde{u} - m\alpha\tilde{s}\tilde{t}/r$ into (16), we get

$$\begin{aligned}f_w(s,u,t) &= \epsilon\tilde{u}^r - \epsilon m\tilde{s}\tilde{t}\tilde{u}^{r-1} + \dots + g(\tilde{s},\tilde{t}) \\ &= \left(\alpha\tilde{u} - m\alpha\tilde{s}\tilde{t}/r\right)^r - r\left(\alpha\tilde{u} - m\alpha\tilde{s}\tilde{t}/r\right)^{r-2} + \dots\end{aligned}$$

Thus, the coefficient at $(\tilde{u})^0$ is a polynomial in $\tilde{s}\tilde{t}$ of total degree r, hence it is a polynomial in $\tilde{s}, \tilde{t}$ of total degree $2r$, which implies, by Lemma 4.15, that $\beta \equiv 0$ and $\tilde{w} = (zv)^r$. □

Corollary 4.16. *Let $w(x,y) = x^a y^b \dots$ be a reduced word of complexity r such that $f_w(s,u,t) = D_r(q(s,u,t))$ over $\mathbb{Q}$. Then $w(x,y) = (x^a y^b)^r$.*

Proof. According to Proposition 4.11, every such w is trace-similar to $(x^a y^b)^r$. It remains to apply Proposition 4.14. □

Proposition 4.17. *With Notation* 4.2*, if r is prime and w is not p-equidistributed, then $r \neq p$ and at least one of A and B is nonzero.*

Proof. We maintain the notation of Proposition 4.11. Suppose that w is not p-equidistributed. Then its trace polynomial is $\mathbb{F}_p$-composite, $f_w = h(q(s,u,t))$, and h is not p-special. Since $\deg_u f_w = r$, h is not linear, and $\deg h$ divides r, we have $\deg h = r$. In the notation of Proposition 4.11, this means that $n = r$ and $q(s,u,t) = uG(s,t) + H(s,t)$. Consider two cases.

Case 1. $A = B = 0$. Then

$$f_w(s,2,s) = h(q(s,2,s)) = D_{A-B}(s) \equiv 2,$$

$$f_w(s,s^2-2,s) = h(q(s,s^2-2,s)) = D_{A+B}(s) \equiv 2.$$

Thus, by (9), (10), we have $q(s,2,s) \equiv c_1 \in \mathbb{F}$, $q(s,s^2-2,s) \equiv c_2 \in \mathbb{F}$. Since $G(s,s)$ is a polynomial, from (11) for $s \neq \pm 2$ it follows that $c_1 = c_2$, $G(s,s) \equiv 0$, $H(s,s) = \text{const}$. This would mean that for at least one of the syllables we have $g_{a_i,b_i}(s,s) \equiv 0$, which is impossible for big powers of p (see [BG, Lemma 2.3]). It follows that this case does not occur.

Case 2. At least one of A and B is not 0 and $r = p$. In this case, by Proposition 4.10 we have $h(z) = D_r(z)$ is a permutation polynomial, thus h is p-special, contrary to the assumption on h. □

Corollary 4.18. *Let $w(x,y) = x^a y^b \dots$ be a reduced word of prime complexity r. If $p > r$ and w is not p-equidistributed, then $w = v(x,y)^r$.*

Proof. If $w(x,y)$ is not p-equidistributed, then according to Theorem 2.12, Proposition 4.10, and Proposition 4.17, we have $f_w = D_r(q(s,u,t))$. By Proposition 4.17, either $A \neq 0$ or $B \neq 0$. By Proposition 4.13, we have $w = (x^\alpha y^\beta)^r$ where $\alpha = A/r$, $\beta = B/r$. □

Corollary 4.19. *The word $w(x,y) = x^a y^b x^c y^d$ is either equidistributed or equal to $(x^a y^b)^2$.*

Proof. Suppose that w is not equidistributed. Then for some prime p its trace polynomial f_w is $\mathbb{F}_p$-composite, $f_w = h(q(s,u,t))$, and h is not p-special. By Proposition 4.17, $w \neq x^a y^b x^{-a} y^{-b}$ and $p > 2$. Then, by Proposition 4.13, $a = c$, $b = d$, and $w(x,y) = (x^a y^b)^2$. □

5. Generic words

In this section, we address the following question: picking up a "generic" word w, should we expect that it is equidistributed? There is a large body of literature dedicated to the notion of genericity, and there are several different approaches to this notion. We mostly follow the setting adopted in [KS].

Definition 5.1. (cf. [KS]) Denote by $\mathcal{R}$ some set of reduced words $w \in F_2$ written in form (7). For a word of complexity r, let $\ell(w) = \sum_{i=1}^{r}(|a_i| + |b_i|)$ denote the length of w. Let $S \subseteq \mathcal{R}$. Set

$$\rho(n,S) = \#\{w \in S : \ell(w) \leq n\},$$

$$\mu(n,S) = \frac{\rho(n,S)}{\rho(n,\mathcal{R})}.$$

We say that S is

- *generic* if $\lim_{n\to\infty} \mu(n,S) = 1$,
- *exponentially generic* if it is generic and the convergence is exponentially fast,
- *negligible* if this limit equals 0,
- *exponentially negligible* if it is negligible and the convergence is exponentially fast.

Evidently, S is (exponentially) generic if and only if the complement $\mathcal{R} \setminus S$ is (exponentially) negligible.

Proposition 5.2. *Let $\mathcal{R}$ be the set of words w of* prime *complexity. Then the set S of words $w \in \mathcal{R}$, such that the corresponding morphism $\mathbb{P}_w\colon \mathrm{SL}_{2,\mathbb{Z}} \times \mathrm{SL}_{2,\mathbb{Z}} \to \mathrm{SL}_{2,\mathbb{Z}}$ is p-equidistributed for all but finitely many primes p, is exponentially generic in $\mathcal{R}$.*

Proof. Let $w \in \mathcal{R}$. Suppose that $w \notin S$, i.e., there exist infinitely many primes p such that the word morphism P_w is *not* p-equidistributed. Denote by $\mathcal{P}$ the set of all such primes. By Corollary 4.18, $w = (x^a y^b)^r$.

It remains to refer to [AO] where it is proven that the property of a word to be a proper power of another word is exponentially negligible. Hence S is exponentially generic in $\mathcal{R}$. □

Remark 5.3. We believe that with some more effort, one can significantly strengthen Proposition 5.2, in particular, by dropping the primality restriction on the complexity. We leave this to experts in word combinatorics.

6. Concluding remarks

It is tempting to generalize our results in the following directions:

(i) extend them from words in two letters to words in d letters, $d > 2$;
(ii) keep $d = 2$ but consider arbitrary finite Chevalley groups;
(iii) combine (i) and (ii).

Whereas in case (i) one can still hope to use trace polynomials, which exist for any d, to produce criteria for equidistribution, cases (ii) and (iii) require some new terms for formulating such criteria and new tools for proving them.

Regardless of getting such criteria, it would be interesting to compare, in the general case, the properties of having large image and being equidistributed, in the spirit of Corollary 3.4. We dare to formulate the following conjecture.

Conjecture 6.1. *For a fixed p, let G_q be a family of Chevalley groups of fixed Lie type over $\mathbb{F}_q$ ($q = p^n$ varies). For a fixed word $w \in F_d$, $d \geq 2$, let $P_q = P_{w,q}\colon (G_q)^d \to G_q$ be the corresponding map. Suppose that*
($*$) for all n big enough the image of P_q contains all regular semisimple elements of G_q.

Then the family $\{P_q\}$ is almost p-equidistributed.

It is a challenging task to describe the words w satisfying condition (*) in Conjecture 6.1 (cf. the discussion in [LST] after Theorem 5.3.2). Certainly, words of the form $w = v^k$, $k \geq 2$, do not satisfy this condition. We do not know any non-power word for which (*) does not hold.

One can try yet another direction: consider equidistribution problems for matrix *algebras* and for polynomials more general than word polynomials (see Introduction). Even the case of 2×2-matrices is completely open.

Acknowledgements. The authors were supported in part by the Minerva Foundation through the Emmy Noether Research Institute for Mathematics. Kunyavskiĭ was supported in part by grant 1207/12 of the Israel Science Foundation. A part of the work was done during the visit of the second author to the MPIM (Bonn). Support of these institutions is gratefully appreciated.

We thank S. Garion, I. Kapovich, M. Larsen, A. Shalev, and Yu. Zarhin for helpful discussions and G. L. Mullen for providing reference [GC].

References

[AP] I. V. Arzhantsev, A. P. Petravchuk, Closed polynomials and saturated subalgebras of polynomial algebras, Ukrain. Mat. Zh. 59 (2007) 1587–1593 = Ukrainian Math. J. 59 (2007) 1783–1790.

[AO] G. N. Arzhantseva, A. Yu. Ol'shanskii, The class of groups all of whose subgroups with lesser number of generators are free is generic, Mat. Zametki 59 (1996) 489–496; English transl. in: Math. Notes 59 (1996) 350–355.

[BG] T. Bandman, S. Garion, Surjectivity and equidistribution of the word $x^a y^b$ on $PSL(2,q)$ and $SL(2,q)$, Intern. J. Algebra Computation 22 (2012), no. 2, 1250017, 33 pp.

[BGG] T. Bandman, S. Garion, F. Grunewald, On the surjectivity of Engel words on $PSL(2,q)$, Groups Geom. Dyn. 6 (2012) 409–439.

[BGKP] T. Bandman, N. Gordeev, B. Kunyavskiĭ, E. Plotkin, Equations in simple Lie algebras, J. Algebra 355 (2012) 67–79.

[BGK] T. Bandman, F. Grunewald, B. Kunyavskiĭ (with an appendix by N. Jones), Geometry and arithmetic of verbal dynamical systems on simple groups, Groups Geom. Dyn. 4 (2010) 607–655.

[BDN] A. Bodin, P. Dèbes, S. Najib, Indecomposable polynomials and their spectrum, Acta Arith. 139 (2009) 79–100.

[Bo] A. Borel, On free subgroups of semisimple groups, Enseign. Math. **29** (1983) 151–164; reproduced in Œuvres - Collected Papers, vol. IV, Springer-Verlag, Berlin–Heidelberg, 2001, pp. 41–54.

[GS] S. Garion, A. Shalev, Commutator maps, measure preservation, and T-systems, Trans. Amer. Math. Soc. 361 (2009) 4631–4651.

[GL] S. R. Ghorpade, G. Lachaud, Étale cohomology, Lefschetz theorems and number of points of singular varieties over finite fields, Moscow Math. J. 2 (2002) 589–631; 9 (2009) 431–438.

[GC] H. Gomez-Calderon, On the decomposition of $x^d + a_l x^l + \cdots + a_1 x + a_0$, Internat. J. Math. & Math. Sci. 23 (2000) 777–781.

[Ho] R. D. Horowitz, Characters of free groups represented in the two-dimensional special linear group, Comm. Pure Appl. Math. 25 (1972) 635–649.

[JLO] S. Jambor, M. W. Liebeck, E. A. O'Brien, Some word maps that are non-surjective on infinitely many finite simple groups, `arXiv:1205.1952`.

[KBMR] A. Kanel-Belov, S. Malev, L. Rowen, The images of non-commutative polynomials evaluated on 2×2 matrices, Proc. Amer. Math. Soc. 140 (2012) 465–478.

[KS] I. Kapovich, P. Schupp, Random quotients of the modular group are rigid and essentially incompressible, J. reine angew. Math. 628 (2009) 91–119.

[KN] M. Kassabov, N. Nikolov, Words with few values in finite simple groups, Quart. J. Math., to appear.

[La] M. Larsen, Word maps have large image, Israel J. Math. 139 (2004) 149–156.

[LP] M. Larsen, R. Pink, Finite subgroups of algebraic groups, J. Amer. Math. Soc. 24 (2011) 1105–1158.

[LS1] M. Larsen, A. Shalev, Word maps and Waring type problems, J. Amer. Math. Soc. 22 (2009) 437–466.

[LS2] M. Larsen, A. Shalev, Fibers of word maps and some applications, J. Algebra 354 (2012) 36–48.

[LST] M. Larsen, A. Shalev, P. H. Tiep, Waring problem for finite simple groups, Ann. Math. 174 (2011) 1885–1950.

[Le1] M. Levy, Word maps with small image in simple groups, `arXiv:1206.1206`.

[Le2] M. Levy, Word maps with small image in almost simple groups and quasisimple groups, `arXiv:301.7188`.

[LMT] R. Lidl, G. L. Mullen, G. Turnwald, Dickson Polynomials, Pitman Monographs and Surveys in Pure Appl. Math., vol. 65, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1993.

[LN] R. Lidl, H. Niederreiter, Finite Fields, Encycl. Math. Appl., vol. 20, Addison-Wesley Publ. Company, Ma., 1983.

[Lu] A. Lubotzky, Images of word maps in finite simple groups, `arXiv:1211.6575`.

[MS] A. G. Myasnikov, V. Shpilrain, Automorphic orbits in free groups, J. Algebra 269 (2003) 18–27.

[Na] S. Najib, Une généralisation de l'inégalité de Stein–Lorenzini, J. Algebra 292 (2005) 566–573.

[Pu] D. Puder, Primitive words, free factors and measure preservation, `arXiv:1104.3991`.

[PP] D. Puder, O. Parzanchevski, Measure preserving words are primitive, `arXiv:1202.3269`.

[Ri] J. F. Ritt, Prime and composite polynomials, Trans. Amer. Math. Soc. 23 (1922) 51–66.

[Sh1] A. Shalev, Commutators, words, conjugacy classes and character methods, Turkish J. Math. 31 (2007) 131–148.

[Sh2] A. Shalev, Word maps, conjugacy classes, and a non-commutative Waring-type theorem, Ann. Math. 170 (2009) 1383–1416.

[Tu] G. Turnwald, On Schur's conjecture, J. Austral. Math. Soc. Ser. A 58 (1995) 312–357.

[Wa] D. Wan, A p-adic lifting and its application to permutation polynomials, in: "Finite Fields, Coding Theory and Advances in Communications and Computing" (G. L. Mullen, P. J.–S. Shiue, eds.), Lecture Notes Pure Appl. Math., vol. 141, Marcel Dekker, New York, 1993, pp. 209–216.

[WSC] D. Wan, P. J.-S. Shiue, C. Chen, Value sets of polynomials over finite fields, Proc. Amer. Math. Soc. 119 (1993) 711–717.

Department of Mathematics, Bar-Ilan University, 5290002 Ramat Gan, Israel
E-mail address: `bandman@macs.biu.ac.il`

E-mail address: `kunyav@macs.biu.ac.il`