VERSITA

Central European Journal of **Mathematics**

Equations in simple matrix groups:
algebra, geometry, arithmetic, dynamics

Tatiana Bandman[1*], Shelly Garion[2†], Boris Kunyavskiĭ[1‡]

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

2 Fachbereich Mathematik und Informatik, Universität Münster, Einsteinstrasse 62, Münster, 48149, Germany

Abstract: We present a survey of results on word equations in simple groups, as well as their analogues and generalizations, which were obtained over the past decade using various methods: group-theoretic and coming from algebraic and arithmetic geometry, number theory, dynamical systems and computer algebra. Our focus is on interrelations of these machineries which led to numerous spectacular achievements, including solutions of several long-standing problems.

MSC: 14G05, 14G15, 20D06, 20D10, 20G40, 37P05, 37P25, 37P35, 37P55

Keywords: Matrix groups • Finite simple groups • Special linear group • Word map • Trace map • Arithmetic dynamics • Periodic points • Finite fields • Lang–Weil estimate

© *Versita Sp. z o.o.*

...
И снова скальд чужую песню сложит
И как свою ее произнесет.
...
Once more a skald will other's song compose
And utter it as of his own gift.

Осип Мандельштам, 1914
Osip Mandelstam, 1914
(translated by Rafael Shusterovich)

* E-mail: bandman@macs.biu.ac.il
† E-mail: shelly.garion@uni-muenster.de
‡ E-mail: kunyav@macs.biu.ac.il

Springer

1. Introduction

Matrix equations, which in the most general form can be written as

$$F(A_1, \ldots, A_m, X_1, \ldots, X_d) = 0,$$

where $A_1, \ldots, A_m$ are some fixed matrices, $X_1, \ldots, X_d$ are unknowns, F is an associative noncommutative polynomial, and the solutions must belong to a certain class of matrices, constitute a vast research domain, with spectacular applications well beyond algebra, say, in areas such as differential equations and mathematical physics. Noncommutativity brings lots of counter-intuitive phenomena, which can be observed even when asking deceptively simple questions about solvability of the equation or about the number of its solutions, even within the class of innocent looking quadratic matrix equations, see, e.g., [25, 35, 39, 42–44, 116]. (Of course, the reader coming from differential equations and familiar, say, with Riccati matrix equations will not be too much surprised by the difficulty of the question.)

To simplify the problem, one can limit oneself to considering equations of the form

$$F(X_1, \ldots, X_d) = A, \tag{1}$$

where A is the only allowed constant matrix, and only scalar constants are permitted to appear as coefficients of the polynomial F. Even in this limited form, the question about solvability of (1) is far from being settled though it has been extensively studied since the 1970's when Kaplansky asked about the existence of polynomials whose value sets consist of the scalar matrices. Such polynomials (called central) were discovered by Formanek [36] and Razmyslov [106]; naturally, if F is central and the matrix A is not scalar then equation (1) has no solutions. The same happens in the case where the value set of F consists of matrices with zero trace (say, when F is a Lie polynomial) and $\operatorname{tr} A \neq 0$ (and, of course, in the trivial case where F is identically zero on the algebra $\mathsf{M}(n, K)$ of $n \times n$-matrices with coefficients from the field K). However, there are subtler obstacles to the solvability of (1), see [62] and references therein. Even a very special case where F is a Lie polynomial is not yet settled though some cases where (1) is solvable as well as some obstacles to solvability were discussed in [10].

Further simplification is essential for the present survey: we focus our attention on the case where the solutions to the equations under consideration must belong to a certain *multiplicative group* of matrices. This naturally leads to the next modification: instead of considering polynomials F as in (1), which can be viewed as elements of the free associative algebra $K\langle X_1, \ldots, X_d \rangle$, we consider monomials $w(x_1, x_1^{-1}, \ldots, x_d, x_d^{-1})$, which can be regarded as elements of the free d-generated group $\mathcal{F}_d = \langle x_1, \ldots, x_d \rangle$. Thus our main object is the following equation (where we shorten our previous notation for w):

$$w(x_1, \ldots, x_d) = g. \tag{2}$$

Here g is a fixed element of a group G, and we are looking for solutions among d-tuples $(g_1, \ldots, g_d)$ of elements of G. By $w(g_1, \ldots, g_d)$ we understand the element of G obtained after substituting g_i instead of x_i and performing all multiplications and inversions in G. We define $w(G) = \{w(g_1, \ldots, g_d) : g_1, \ldots, g_d \in G\}$ as the set of values of w in G. Although equation (2) (henceforth called a word equation) makes sense in the case where G is not necessarily a group of matrices, we make emphasis on the matrix case. The reason is two-fold. First, main applications we are going to discuss refer to the cases where G has nice natural matrix representations. Second, our main goal is to stress the strength of algebraic-geometric methods in treating problems related to word equations. More specifically, our viewpoint can be described as follows: let us regard all matrix entries of x_i as indeterminates, then, after clearing denominators arising because of determinants, we reduce equation (2) to a system of dn^2 polynomial equations in n^2 variables over the ground field K (all matrices are assumed to be of size $n \times n$). Thus we managed to go over from a noncommutative problem to a commutative one, but the price paid is high: the resulting system may be huge and inaccessible to computer algebra systems. However, in certain cases such an approach may turn out to be fruitful, and several instances of successful implementation of this idea will follow.

Once equation (2) is proclaimed as the main object of our investigation, the first natural question one should ask oneself is the following: what is the most natural class of groups G to start with? A strong hint is given by a theorem of

Borel [17] (see also [68]), which states that if G is the group of rational points of a connected semisimple algebraic group defined over an infinite field, then for any $w \neq 1$ equation (2) is solvable when its right-hand side is, roughly speaking, a "typical" element $g \in G$; see subsection 5.1 for a precise statement and some details. Thus the most substantial part of the present survey is devoted to the case where G is a finite simple group. Although in this case Borel's theorem does not give any direct indication which elements should be thought of as "typical", we shall try to convince the reader that there are several ways out which make some sense. In particular, Larsen [68] showed that for every nontrivial word w and $\epsilon > 0$ there exists a number $C(w, \epsilon)$ such that if G is a finite simple group with $\#G > C(w, \epsilon)$ then $\#w(G) \geq \#G^{1-\epsilon}$, see subsection 5.2.

One should note that there is not much hope for positive results of Borel's flavour for groups G which are too far from those singled out in his theorem; see the discussion in [10, Section 5] and relevant references therein (to which one should add a recent paper by Myasnikov and Nikolaev [94]). Therefore, as mentioned above, we restrict ourselves to considering the case where G runs within a family of finite non-abelian simple groups. Moreover, since most of the problems will be of asymptotic nature, we shall usually ignore the sporadic groups. We give more attention to the case of the family $SL(2, q)$ (sometimes extending it to all groups of Lie rank 1, thus adding the family of Suzuki groups $Sz(q)$). This case is, on the one hand, typical enough, and sufficiently rich to have interesting applications. On the other hand, it allows one to use some additional efficient tools (such as trace maps) and get more precise results. Details can be found below, in Sections 3, 4 and 6.

One can ask various questions on equation (2). First, one can fix a concrete group or a family of groups G (say, $G = SL(2, q)$), a concrete element g of G (say, $g = 1$) and a concrete word w, and ask whether equation (2) is solvable (or has a non-identity solution, if $g = 1$ is chosen). Even in such a restrictive setting, some spectacular applications can be obtained, in particular, to a long-standing problem of characterization of finite solvable groups by recursive identities, see Section 3. A variation of this approach, where one proves solvability of a countable system of equations $w_n = 1$ (with G fixed as above), or, particularly, of the system $w_m = w_n$ with w_i obtained after the ith iteration from certain initial data, can be reformulated in the language of periodic points. This viewpoint was (implicitly) taken in [20], where another sequence of words characterizing finite solvable groups was constructed, and made explicit in [13]. In the latter paper this machinery has been further developed which led naturally to some new concepts in arithmetical dynamics. Another impressive application of arithmetical dynamics to a hard group-theoretic problem has been demonstrated by Borisov and Sapir [18, 19], see subsection 4.2.

Further, one can consider more general questions. First, still fixing some "interesting" word w, one may vary $g \in G$ and ask about the solvability of (2) for all g, or, at least, for "most of g". In the language of word maps, where we denote by

$$w \colon G^d \to G$$

the map given by $(g_1, \ldots, g_d) \mapsto w(g_1, \ldots, g_d)$, this means that we are asking whether the word map is surjective or, at least, has a "large" image. Ore's conjecture [102] on representing every element of a finite simple group as a commutator is a fabulous instance of such a setting. Due to immense work spread over more than 50 years, it is now known that the commutator word $w = [x, y] \in \mathcal{F}_2$ satisfies $w(G) = G$ for any finite non-abelian simple group G (see [79], the references therein and subsection 5.2.2 below). It was therefore conjectured by Shalev that a similar result also holds for iterated commutator (so-called Engel words). The study of this case started in [9] (for $G = SL(2, q)$, see subsection 6.1). The words of the form $w = x^a y^b \in \mathcal{F}_2$ have also attracted special interest. Larsen, Shalev and Tiep [73] proved that any such word is surjective on sufficiently large finite simple groups. By further recent results of Guralnick and Malle [53] and of Liebeck, O'Brien, Shalev and Tiep [81], some words of the form $x^b y^b$ are known to be surjective on *all* finite simple groups. The particular case of surjectivity of these words on $SL(2, q)$ was studied in [8] (see subsection 6.1).

A ramification of this sort of questions, where one asks whether any element $g \in G$ (or "most" of them) can be represented as a product of at most k values of w (k is a fixed natural number), has been christened "a Waring-type problem" by Shalev (by analogy to a celebrated problem on representing a natural number as a sum of k powers), and a number of impressive results has been obtained. The most conclusive one, due to Larsen, Shalev and Tiep, says that for every nontrivial word w there exists a constant $C(w)$ such that if G is a finite simple group satisfying $\#G > C(w)$ then $w^2(G) = G$, see [71, 73, 74, 114] and subsection 5.2.1 below.

Note that a naive question whether $w(G) = G$ for any nontrivial word w and all sufficiently large finite simple non-abelian groups G is clearly answered in the negative: indeed, it is easy to see that if G is a finite group and m is an

integer which is not relatively prime to the order of G then for the word $w = x_1^m$ one has $w(G) \neq G$. Hence, if $v \in \mathcal{F}_d$ is any word, then the word map corresponding to $w = v^m$ cannot be surjective. A natural question, suggested by Shalev, whether these words are generally the only exceptions for surjectivity of word maps in finite non–abelian simple groups, was recently answered in the negative by Jambor, Liebeck and O'Brien [60] (see Section 7).

Yet another type of problems, first introduced in [41], arises when one asks about the behaviour of the fibres of the word map rather than about its image. Namely, Shalev asked [113, Problem 2.10] whether (the cardinalities of) these fibres are equidistributed (or close to be equidistributed) when g varies in some "large" subset of the image of w and G runs over some family of finite groups. It was proved in [41] that the word $w = x^2 y^2 \in \mathcal{F}_2$, the commutator word $w = [x, y] \in \mathcal{F}_2$, as well as the words $w = [x_1, \dots, x_d] \in \mathcal{F}_d$, d-fold commutators in any arrangement of brackets, are almost equidistributed on the family of finite simple non-abelian groups. The case $G = \mathsf{SL}(2, q)$ was studied in some more detail in [8, 9, 14], see Section 6.

We do not pretend that the present overview of word maps is comprehensive. As one can see from the title, shamelessly plagiarized from Manin's book [91], it was conceived for emphasizing the role of various tools lying outside group theory, particularly, of algebraic-geometric and arithmetic-geometric nature, in solving hard group-theoretic problems (note, however, that the role of the "algebra–geometry–arithmetic" triad in Manin's setting was somewhat different: his focus was on systematic use of algebraic, mainly Galois-cohomological, methods in relating arithmetic phenomena, such as counter-examples to local-global principles, to geometric ones, such as non-rationality; our viewpoint is different: a typical target is a group-theoretic question whereas algebraic and arithmetic geometry provide an efficient machinery). The reader inclined to purely algebraic approaches to word equations is referred to an excellent exposition given in the monograph by Segal [111] and the more recent survey by Nikolov [97]. Vast literature on equations and system of equations in free (and close to free) groups is left aside; see, e.g., [24], particularly the introduction, for a comprehensive bibliographical survey of this theory.

2. Main tools

In this section we shall briefly describe some common machinery used while treating word matrix equations, leaving details, which may vary from one problem to another, for subsequent relevant sections (note, in particular, that the problems discussed in Section 5 require lots of other techniques).

2.1. Crucial assumptions

Our general setting is as follows.

(I) We view a matrix group G as an algebraic variety $G \subset \mathbb{A}^N_{x_1, \dots, x_N}$ embedded into an affine space, where the entries of a matrix $v \in G$ play role of its coordinates $(x_1(v), \dots, x_N(v)) = (v_1, \dots, v_N)$ (though this is not the case if G is a Suzuki group, G then has a "large" subset meeting this condition; see Section 3).

(II) The group product $(v_1, v_2) \mapsto v = v_1 v_2$ is a polynomial morphism $G \times G \to G$.

Changing N if necessary, we can guarantee that the inversion will also be a polynomial map; in most cases, however, this is not needed because usually we work with matrices of determinant 1.

Let $d \geq 2$ be an integer, and let $w = w(x_1, \dots, x_d)$ be a nontrivial element of the free group $\mathcal{F}_d = \langle x_1, \dots, x_d \rangle$ on d letters x_i. This means that it is a reduced word in x_i and x_i^{-1} with nonzero exponents. Then, given a group H, the word w defines a map $f_w \colon H^d \to H$,

$$f_w(h_1, \dots, h_d) = w(h_1, \dots, h_d).$$

When no confusion may arise, we will shorten f_w to w.

Provided (I) and (II), we may interpret an endomorphism of a group as a polynomial morphism $G \to G$ and a word $w(x_1, \dots, x_d)$ as a morphism $G^d \to G$. Similarly, for a word map $w \colon G^d \to G$ and an element $g \in G$, a solution of the equation $w(x_1, \dots, x_d) = g$ corresponds to an affine subvariety $S(w, g) \subset G^d$. If G is a *finite group* (and this is the

case in most applications considered in the present survey), we can choose the ground field to be a *finite field* $\mathbb{F}$. If the problem under consideration requires a study of a family of finite matrix groups G_q, each defined over its own ground field $\mathbb{F}_q$, it is convenient to view G_q as a fibre of a group $\mathbb{Z}$-scheme $\mathcal{G}$. Then solutions to the equation $w(x_1, \ldots, x_d) = g$ are described by points $S(w, g)(\mathbb{F}_q)$. This leads to the following local–global approach: instead of solving equations in an infinite family of groups G_q, one has to study a single $\mathbb{Z}$-scheme $\mathcal{S}(w, g)$ and then look at its fibres at every q.

2.2. Lang–Weil inequality

Once the initial group-theoretic problem is reduced to an arithmetic geometry problem of proving the existence of a rational point on a variety defined over a finite field, a natural tool to use is the Lang–Weil inequality [67]. It says that if X is an n-dimensional *absolutely irreducible* variety over $\mathbb{F}_q$, then asymptotically (i.e., for q large enough), the number of its rational points $\# X(\mathbb{F}_q)$ does not differ too much from the number $\# \mathbb{P}^n(\mathbb{F}_q)$ of points of the n-dimensional projective space (which is $q^n + q^{n-1} + \cdots + q + 1$). Namely, the difference is $O\left(q^{n-1/2}\right) \leq C_1(X) q^{n-1/2} + C_2(X) q^{n-1}$. If one can make this inequality *effective*, i.e., compute or at least estimate C_1 and C_2, one can guarantee the existence of an $\mathbb{F}_q$-point on X for q large enough. Note that such a computation is related to deep topological properties of X and requires some information on its Betti numbers (in suitable cohomology), see [45] for a nice survey of various ramifications and generalizations of the Lang–Weil inequality. In the one-dimensional case the classical Weil estimate gives an expression of the remainder term via the genus of the curve.

One should not forget another difficulty related to checking absolute irreducibility of X, which may be a highly nontrivial task in concrete examples. Most of arising problems comprise computational aspects, and usually require some advanced computer algebra to overcome them, see Section 3.

2.3. Trace map

In the case where $G = \mathrm{SL}(2)$, apart from the general techniques described in the previous section, we have at our disposal another powerful tool going back to classical works of Vogt, Fricke and Klein [37, 38, 124], cited here from the paper [56] (see also [47, 89, 90] for a nice exposition of these results).

Let $\mathcal{F}_d = \langle x_1, \ldots, x_d \rangle$ denote the free group on d generators. For $G = \mathrm{SL}(2, k)$ (k is any commutative ring with 1) and for any $u \in \mathcal{F}_d$ denote by $\mathrm{tr}\, u \colon G^d \to G$ the trace character, $(g_1, \ldots, g_d) \mapsto \mathrm{tr}(u(g_1, \ldots, g_d))$.

Theorem 2.1.
If u is an arbitrary element of $\mathcal{F}_d$, then the character of u can be expressed as a polynomial

$$\mathrm{tr}\, u = P(t_1, \ldots, t_d, t_{12}, \ldots, t_{12\ldots d})$$

with integer coefficients in the $2^d - 1$ characters $t_{i_1 i_2 \ldots i_v} = \mathrm{tr}(x_{i_1} x_{i_2} \ldots x_{i_v})$, $1 \leq v \leq d$, $1 \leq i_1 < i_2 < \ldots < i_v \leq d$.

Let $G = \mathrm{SL}(2, q)$, and let $\pi \colon G^d \to \mathbb{A}^{2^d - 1}$ be defined by

$$\pi(g_1, \ldots, g_d) = (t_1, \ldots, t_d, t_{12}, \ldots, t_{12\ldots d})$$

in the notation of Theorem 2.1.

Let $Z_d = \pi(G^d) \subset \mathbb{A}^{2^d - 1}$. Let $w \colon G^d \to G$ be a word map. It follows from Theorem 2.1 that for every d there exists a polynomial map $\psi \colon \mathbb{A}^{2^d - 1} \to \mathbb{A}^1$ such that the following diagram commutes:

$$
\begin{array}{ccc}
G^d & \xrightarrow{\ w\ } & G \\
\downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \mathrm{tr}} \\
Z_d(\mathbb{F}_q) & \xrightarrow{\ \psi\ } & \mathbb{A}^1.
\end{array}
$$

Moreover, for small d we have a more precise information: one can take $Z_2 = \mathbb{A}^3$ and $Z_3 \subset \mathbb{A}^7$ an explicitly given hypersurface. This diagram allows one to reduce the study of the image and fibres of w to the corresponding problems for ψ, which may be much simpler, see subsections 4.1.3 and 4.1.4.

3.   Equations in groups of Lie rank one and characterization of finite solvable groups

As mentioned in the introduction, even a solution of only *one* word equation in a *small* family of finite groups may lead to spectacular consequences. Here is an instance of such a phenomenon.

We consider a problem of characterizing various classes of groups by identical laws. Say, G is abelian if and only if $[x, y] = 1$ for all $x, y \in G$; a finite group G is nilpotent if and only if there exists $n = n(G)$ such that

$$e_n(x, y) = \underbrace{\big[[x, y], \ldots, y\big]}_{n \text{ times}} \equiv 1$$

(Zorn's theorem [128]); what about finite *solvable* groups?

There are various characterizations of finite solvable groups (some of them are discussed in [52]), among which there are those using identities. For instance, a finite group G satisfying a "short" identity of some special form must be solvable [23] (we thank the second referee for providing this reference). We want to obtain a criterion as close as possible to Zorn's theorem.

Two-variable sequences

The purpose is to establish a characterization of the solvable groups in the class of finite groups by an inductively defined, Engel–like sequence of two–variable identities. This problem has a long history and admits many counterparts and generalizations; the interested reader is referred to [52] for a survey.

Definition 3.1.
Let $w \in \mathcal{F}_2$ and $f \in \mathcal{F}_3$.

 (i) We say that a two–variable sequence $\{v_n(x, y)\}$ is defined by the recursive law (w, f) if

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

 (ii) We say that the sequence is k–valent, $k = 1, 2, 3$, if $f = f(x, y, z)$ depends on k variables among which z must appear.

 (iii) We say that the sequence is Engel–like if $f(x, y, 1) = 1$ in $\mathcal{F}_3$.

 (iv) We say that the sequence $\{v_n\}$ characterizes finite solvable groups if the following holds: A finite group G is solvable if and only if there exists $n \in \mathbb{N}$ such that for all $(x, y) \in G \times G$, $v_n(x, y) = 1$.

Remark 3.2.
If $\{v_n\}$ is an Engel–like sequence, the condition $v_n(x, y) = 1$ implies that $v_m(x, y) = 1$ for all $m \geq n$.

Example 3.3.
The original Engel sequence corresponds to $w(x, y) = [x, y]$, $f(x, y, z) = [z, y]$. It is 2–valent and Engel–like in the sense of the previous definition but, of course, does not characterize finite solvable groups.

The first example of a sequence satisfying the conditions of Definition 3.1 was found in [11]. Let the commutator be defined as $[a, b] = aba^{-1}b^{-1}$. Set

$$w(x, y) = x^{-2}y^{-1}x, \qquad f(x, y, z) = \left[xzx^{-1}, yzy^{-1}\right].$$

The corresponding sequence is

$$u_1(x, y) = x^{-2}y^{-1}x, \qquad u_{n+1}(x, y) = \left[xu_n(x, y)x^{-1}, yu_n(x, y)y^{-1}\right]. \tag{3}$$

Theorem 3.4 ([11, 12]).
A finite group G is solvable if and only if for some n we have $u_n(x, y) = 1$ for all $x, y \in G$.

This section is a brief exposition of the proof of this theorem. It involves algebraic geometry, arithmetic geometry, algebra and heavy MAGMA and SINGULAR computations.

Remark 3.5.
The role of the first word is very important: for example, if one changes the first word to a more natural one: $u_1(x, y) = w(x, y) = [x, y]$, keeping the same recursive law f, we do not know whether or not the obtained sequence characterizes finite solvable groups. Moreover, neither do we know this for the prototypical sequence arising from $u_1(x, y) = [x, y]$, $f(x, y, z) = \left[[z, x], [z, y]\right]$, which characterizes finite-dimensional solvable Lie algebras (if $[\cdot, \cdot]$ is understood as Lie bracket) [51].

More 3-valent sequences $\{u_n\}$ for which Theorem 3.4 holds were produced in [107]. We conjecture after long computer experiments that for most (if not for any) choices of f satisfying obvious necessary conditions, there is an initial word u_1 such that Theorem 3.4 holds; see the discussion in subsection 4.1.4 and Conjecture 7.1 below.

An example of a 2-valent sequence which still characterizes finite solvable groups was given in [20],

$$s_1(x, y) = x, \qquad s_{n+1}(x, y) = \left[ys_n(x, y)y^{-1}, s_n(x, y)^{-1}\right]. \tag{4}$$

Theorem 3.6 ([20]).
A finite group G is solvable if and only if for some n we have $s_n(x, y) = 1$ for all $x, y \in G$.

Whereas the proof of Theorem 3.4 is based on solvability of a single word equation in a family of simple groups (see the next section), the proof of Theorem 3.6 requires solvability of a system of countably many equations and has an evident dynamical flavour, see subsection 4.1.

Equations in $\mathrm{SL}(2)$ and Suzuki groups

In this section we sketch a proof of Theorem 3.4. Clearly in every solvable group the identities $u_n(x, y) \equiv 1$ are satisfied from a certain $n \in \mathbb{N}$ onward. By a standard argument, using minimal counter-examples, the nontrivial direction of Theorem 3.4 follows from

Theorem 3.7.
Let G be a finite non-abelian simple group. Then there are $x, y \in G$ such that

$$u_1(x, y) = u_2(x, y), \qquad u_1(x, y) \neq 1. \tag{5}$$

Moreover, the same argument shows that we only need to prove Theorem 3.7 for the groups G in J. Thompson's list of the minimal simple non-solvable groups [119]. For simplicity we slightly extend it:

(a) $G = \mathrm{PSL}(2, q)$ where $q \geq 4$ ($q = p^m$, p prime),

(b) $G = \mathrm{Sz}\left(2^{2m+1}\right)$, $m \in \mathbb{N}$,

(c) $G = \mathrm{PSL}(3, 3)$,

because this does not make our proof more complicated. Here $\mathrm{Sz}\left(2^{2m+1}\right)$, $m \in \mathbb{N}$, denotes the Suzuki groups (see [59, XI.3]).

For small groups from this list it is an easy computer exercise to verify Theorem 3.7. There are for example altogether 44928 suitable pairs x, y in the group $\mathrm{PSL}(3, 3)$.

The general idea of the proof is as described in Section 2. For a group G in the above list, using a matrix representation over $\mathbb{F}_q$, we interpret solutions of the equation $u_1(x, y) = u_2(x, y)$ as $\mathbb{F}_q$-rational points of an algebraic variety. Estimates of Lang–Weil type for the number of rational points on a variety defined over a finite field guarantee in appropriate circumstances the existence of such points for big q. Of course we are faced here with the extra difficulty of having to ensure that $u_1(x, y) \neq 1$ holds. This is achieved by taking x, y from appropriate Zariski-open subsets only.

The case $G = \mathrm{PSL}(2, q)$

Let us look for a solution to (5) among the matrices of the form

$$x(v) = \begin{bmatrix} t & -1 \\ 1 & 0 \end{bmatrix}, \qquad y(v) = \begin{bmatrix} 1 & b \\ c & 1 + bc \end{bmatrix},$$

where $v = (t, b, c) \in \mathbb{A}^3$.

SINGULAR computations show that the equation $u_1(x(v), y(v)) = u_2(x(v), y(v))$ defines an algebraic *curve* $C \subset \mathbb{A}^3$. Note that a priori one should expect $\dim C = 0$ because C is defined by three equations in three variables. It is this dimension jump that forced a somewhat peculiar choice of the initial word $u_1(x, y)$.

It turns out that

- C does not depend on $\mathbb{F}_q$;

- C is absolutely irreducible *for any* p;

- $p_a(\overline{C}) = 10$, $\deg \overline{C} = 12$ (where $\overline{C}$ denotes the projective closure of C, p_a is the arithmetic genus and deg is the degree).

(The second and third statements were established by computer calculations; the proof of absolute irreducibility is technically the most difficult part; the interested readers are referred to the original papers.)

From Weil's estimate for the number of rational points over a finite field it follows that $C(\mathbb{F}_q) \geq q + 1 - 2p_a\sqrt{q} - d$, hence for $q > 593$ there are enough rational points on C to prove Theorem 3.7. Solutions for (5) in the groups $G = \mathrm{PSL}(2, q)$, $q \leq 593$, were found by computer.

The case of Suzuki groups

In the proof of Theorem 3.7 the case of Suzuki groups $G = \mathrm{Sz}(q)$, $q = 2^{2m+1}$, is the most difficult one. The main reason is that although $\mathrm{Sz}(q)$ is contained in $\mathrm{GL}(4, q)$, it is not a fibre of a $\mathbb{Z}$-scheme.

In fact the group $\mathrm{Sz}(q)$ is defined with the help of a field automorphism of $\mathbb{F}_q$ ("square root of Frobenius", see, e.g., [58, Chapter 20] for a precise definition), and hence the standard matrix representation for $\mathrm{Sz}(q)$, obtained in the original paper of Suzuki [117], contains entries depending on q. We shall describe now how our problem can still be treated by methods of algebraic geometry.

This time one looks for a solution of (5) among the matrices parameterized by points of an 8-dimensional space: for $v = (a, b, c, d, a_0, b_0, c_0, d_0) \in \mathbb{A}^8$, let

$$x(v) = \begin{bmatrix} a^2 a_0 + ab + b_0 & b & a & 1 \\ a a_0 + b & a_0 & 1 & 0 \\ a & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \qquad y(v) = \begin{bmatrix} c^2 c_0 + cd + d_0 & d & c & 1 \\ c c_0 + d & c_0 & 1 & 0 \\ c & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Moreover, $x(v) \in \mathrm{Sz}(q)$ if and only if

$$a_0 = a^{2^{m+1}}, \qquad b_0 = b^{2^{m+1}}. \tag{6}$$

Similarly, $y(v) \in \mathrm{Sz}(q)$ if and only if

$$c_0 = c^{2^{m+1}}, \qquad d_0 = d^{2^{m+1}}. \tag{7}$$

Define $\Sigma = \{v : u_1(x(v), y(v)) = u_2(x(v), y(v))\} \subset \mathbb{A}^8$. From relations (6) and (7) it follows that

$$\Sigma(\mathbb{F}_q) = \Sigma \cap \left\{ a_0 = a^{2^m}, b_0 = b^{2^m}, c_0 = c^{2^m}, d_0 = d^{2^m} \right\}.$$

In order to prove that the set $\Sigma(\mathbb{F}_q)$ is not empty, it is represented as the set of fixed points of an automorphism of $\mathbb{A}^8$. Namely, consider an automorphism $\alpha \colon \mathbb{A}^8 \to \mathbb{A}^8$ defined by

$$\alpha(a, b, c, d, a_0, b_0, c_0, d_0) = (a_0^2, b_0^2, c_0^2, d_0^2, a, b, c, d).$$

Then α^2 is the Frobenius automorphism. Moreover, from (6) and (7) it follows that $x(v), y(v) \in \mathrm{Sz}(q)$ if and only if $\alpha^m(v) = v$, i.e., if and only if v is a fixed point of the mth iteration of α. Computations show that there exists a subset $U \subset \Sigma$ such that

(i) U is a smooth, absolutely irreducible, affine variety,

(ii) $\dim U = 2$,

(iii) U is α-invariant.

Denote by $b^i(U) = \dim H^i_{\mathrm{ét}}(U, \overline{\mathbb{Q}}_\ell)$ the ℓ-adic Betti number, $\ell \neq 2$. Then

Proposition 3.8.
$b^1(U) \leq 675$ and $b^2(U) \leq 2^{22}$.

The estimates contained in Proposition 3.8 are derived from results of Adolphson–Sperber [2] and Ghorpade–Lachaud [46] allowing one to bound the Euler characteristic of an affine variety in terms of the number of variables, the number of defining polynomials and their degrees. Note that since U is affine, we have $b^3(U) = b^4(U) = 0$. Since U is nonsingular, the ordinary and compact Betti numbers of U are related by the Poincaré duality, and we have $b_c^i(U) = b^{4-i}(U)$.

In order to estimate the number $\#\,\mathrm{Fix}(U, n)$ of α^n-fixed points in U, applied was the *Lefschetz Trace Formula*:

$$\#\,\mathrm{Fix}(U, n) = \sum_{i=0}^{4} (-1)^i \,\mathrm{tr}\left(\alpha^n : H_c^i(U, \overline{\mathbb{Q}}_\ell) \right) \tag{8}$$

where $\mathrm{Fix}(U, n)$ is the set of fixed points of α^n acting on U.

(A remark for the reader interested in algebraic-geometric details: technical difficulties arising from the fact that U is not projective and α is not Frobenius, were overcome, roughly, as follows: α can be extended to an endomorphism of $\mathbb{P}^8$ having no fixed points in $\overline{U} \setminus U$, where $\overline{U}$ is the projective closure of U. Then one can use Deligne's conjecture stating that a formula of Lefschetz type holds after composing α with a sufficiently high power of Frobenius, which, in our case, means a high odd power of α. Recall that Deligne's conjecture has been proved by Fujiwara [40]; see [123] for simplifications and generalizations.)

From formula (8) and Deligne's estimates for the eigenvalues of the endomorphism induced by α on the étale cohomology, an inequality of Lang–Weil type follows:

$$\left| \#\,\mathrm{Fix}(U, n) - 2^n \right| \leq b^1(U) 2^{3n/4} + b^2(U) 2^{n/2}.$$

An easy estimate shows that $\#\,\mathrm{Fix}(U, n) \neq 0$ for $n > 48$. The cases $n < 48$ were checked with the help of MAGMA.

4. Arithmetic dynamics

In this section we present two group-theoretic problems for which the language of arithmetic dynamics appears to be an adequate one. The first one, discussed in subsection 4.1, arose from attempts to understand the proof of Bray, J. Wilson, R. Wilson [20], who exhibited another sequence characterizing finite solvable groups (see Section 3), find an explanation of the phenomenon that a sequence possesses this property, and produce more such sequences. This was essentially done in [13], where the interested reader can find more elaborated constructions in arithmetic dynamics which are beyond the scope of the present survey.

Another instance is related to the work of Borisov and Sapir [18, 19], see subsection 4.2, where somewhat similar philosophy led to an answer to another long-standing group-theoretic question (this time, from the theory of infinite groups).

4.1. Verbal dynamical systems

Given a two-variable sequence $\{v_n(x, y)\}$ defined by a recursive law (w, f) (see Definition 3.1), one can ask whether or not it characterizes finite solvable groups. An obvious necessary condition is that $\{v_n\}$ must descend along the derived series, and we always assume that this condition holds. We also assume that $\{v_n\}$ does not contain the identity word. Then the only condition to check is the following one:

$$G \text{ is not solvable} \quad \Longrightarrow \quad \text{for all } n \text{ there exists } (x, y) \in G \times G \text{ such that } v_n(x, y) \neq 1.$$

This condition may be further reduced (see Section 3) to the following property.

Property 4.1.
Let G be one of the groups $\mathrm{PSL}(2, q)$, $q \neq 2, 3$, $\mathrm{Sz}\left(2^{2m+1}\right)$, or $\mathrm{PSL}(3, 3)$. Then for all n there exists $(x, y) \in G \times G$ such that $v_n(x, y) \neq 1$.

Since it is very easy to check the single case $\mathrm{PSL}(3, 3)$, we assume throughout below that the property holds for this group.

We will say that a sequence is *very good* if Property 4.1 holds for all groups listed therein and is *good* if it holds at least for all $\mathrm{PSL}(2, q)$, $q \neq 2, 3$. In this section we want to describe good sequences.

For a further simplification, the following observation is crucial: if $\{v_n\}$ is an Engel-like sequence then Property 4.1 holds in a group G as soon as there exist n, $m > n$ and $(x, y) \in G \times G$ such that

$$v_n(x, y) = v_m(x, y) \neq 1.$$

This simple reformulation allows one to replace the proof of solvability of a concrete equation (say, $v_1(x, y) = v_2(x, y) \neq 1$) in a family of groups, as was done in Section 3, by the proof of existence of a *preperiodic point* (different from identity) of a certain dynamical system generated by the recursive law f.

Recall that a sequence $v_n(x, y) = \left(w(x, y), f(x, y, z)\right)$, generated by the first word w and recursive law f, is formed by the rule

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

In such a situation, one can define, for any group G, a self-map $G \times G \times G \to G \times G \times G$ by adding "tautological" variables. We arrive at the notion of *verbal dynamical system* $(\mathcal{G}, \widetilde{\phi}, V)$ consisting of the following data:

- a group scheme $\mathcal{G}$ (in our case $\mathcal{G} = \mathrm{SL}(2, \mathbb{Z})$);
- a morphism $\phi \colon \mathcal{G} \times \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ induced by the word $f(x, y, z)$;

- an endomorphism $\widetilde{\phi} \colon \mathcal{G} \times \mathcal{G} \times \mathcal{G} \to \mathcal{G} \times \mathcal{G} \times \mathcal{G}$ defined by $\widetilde{\phi}(x, y, z) = (x, y, \phi(x, y, z))$;

- a forbidden set V (in our case $V = \mathcal{G} \times \mathcal{G} \times \{\mathrm{id}\}$);

- an initial word $w \colon \mathcal{G} \times \mathcal{G} \to \mathcal{G}$.

Conversely, given such data, we reconstruct our iterative sequence $\{v_n(x, y)\}$.

For each field $\mathbb{F}_q$ we consider the fibre $\left(\mathrm{SL}(2, q), \widetilde{\phi}_q, V(\mathbb{F}_q)\right)$. The sequence $\{v_n\}$ is good if for every q there are $\widetilde{\phi}_q$-preperiodic points outside $V(\mathbb{F}_q)$.

Consider three possible types of f.

4.1.1. 1-*valent law*

The only 1-valent law is $f(v) = v^r$. Then $v_{n+1} = v_1^{r^n}$. This sequence is not going down along the derived series and hence does not meet the necessary condition for characterizing solvable groups.

4.1.2. PSL(2), 2-*valent law*

An example of a 2-valent law is sequence (4). In such a case our general setting can be simplified. Namely, if $f = f(y, z)$ does not depend on x, we can restrict our verbal dynamical system to the form $\widetilde{\phi} \colon \mathcal{G} \times \mathcal{G} \to \mathcal{G} \times \mathcal{G}$, $(y, z) \mapsto (y, f(y, z))$, with the forbidden set $\mathcal{G} \times \{\mathrm{id}\}$ and initial word $w(x, y) = x$.

4.1.3. *Traces*

In order to further simplify the dynamical system, one can use the trace map (see subsection 2.3). In the special case $d = 2$, Theorem 2.1 may be formulated as follows.

Theorem 4.2.
Let $G = \mathrm{SL}(2, \mathbb{Z})$. Define $\pi \colon G \times G \to \mathbb{A}^3$ by $\pi(x, y) = (\mathrm{tr}\, x, \mathrm{tr}(xy), \mathrm{tr}\, y)$. Then for a word map $\phi \colon G \times G \to G$, there is a polynomial in three variables $P_\phi(s, u, t)$ such that $\mathrm{tr}\, \phi(x, y) = P_\phi(\mathrm{tr}\, x, \mathrm{tr}(xy), \mathrm{tr}\, y)$.

Denote $s = \mathrm{tr}\, x$, $u = \mathrm{tr}(xy)$, $t = \mathrm{tr}\, y$. Let $f_1(s, u, t) = \mathrm{tr}\, \phi(x, y)$, $f_2(s, u, t) = \mathrm{tr}(\phi(x, y)y)$, and $\psi = (f_1(s, u, t), f_2(s, u, t), t)$. According to Theorem 4.2, we have the following factorization of the verbal dynamical system $(\mathcal{G}, \widetilde{\phi}, V)$, i.e., the following commutative diagram:

$$
\begin{array}{ccc}
\mathcal{G} \times \mathcal{G} & \xrightarrow{\ \widetilde{\phi}\ } & \mathcal{G} \times \mathcal{G} \\
\ \downarrow{\scriptstyle \pi} & & \ \downarrow{\scriptstyle \pi} \\
\mathbb{A}^3_{s,u,t} & \xrightarrow{\ \psi\ } & \mathbb{A}^3_{s,u,t}.
\end{array}
\tag{9}
$$

In diagram (9),

- π is defined over $\mathbb{Z}$;

- π is surjective for all $\mathbb{F}_q$ (see, e.g., [87] or [13]);

- the set Σ of fixed points of ψ, being defined by the system $f_1(s, u, t) = s$, $f_2(s, u, t) = u$, has a positive dimension.

Respectively, for every q we have a commutative diagram:

$$
\begin{array}{ccc}
\mathrm{SL}(2, q) \times \mathrm{SL}(2, q) & \xrightarrow{\ \widetilde{\phi}_q\ } & \mathrm{SL}(2, q) \times \mathrm{SL}(2, q) \\
\ \downarrow{\scriptstyle \pi} & & \ \downarrow{\scriptstyle \pi} \\
\mathbb{A}^3_{s,u,t}(\mathbb{F}_q) & \xrightarrow{\ \psi_q\ } & \mathbb{A}^3_{s,u,t}(\mathbb{F}_q).
\end{array}
\tag{10}
$$

Assume that $\Sigma(\overline{\mathbb{F}}_q)$ contains an irreducible over $\overline{\mathbb{F}}_q$ curve C of genus g and degree d that intersects $\pi(V)$ at k points. Then by Weil's inequality

$$\#(C \setminus \pi(V))(\mathbb{F}_q) \geq (q+1) - 2g\sqrt{q} - d - k,$$

and $(C \setminus \pi(V))(\mathbb{F}_q) \neq \emptyset$ for q big enough, $q \geq q_0(g, d, k)$. Let $a \in (C \setminus \pi(V))(\mathbb{F}_q)$ and $F_a = \pi^{-1}(a) \subset \mathrm{SL}(2, q) \times \mathrm{SL}(2, q) \setminus V_q$. Then F_a is a $\widetilde{\phi}_q$-invariant finite set and thus contains a nontrivial $\widetilde{\phi}_q$-preperiodic point.

We can now formulate a procedure for checking that a 2-valent sequence $\{v_n\}$ is good. Namely, assume that $\{v_n\}$ is defined by a word (law) $f(y, z)$. The process consists of the following steps:

 (i) compute the trace map ψ of the endomorphism $\widetilde{\phi}$;

 (ii) compute the set Σ of fixed points of ψ;

(iii) find an affine curve $S \subset \Sigma$ such that

   (iii$_1$) S is defined over $\mathbb{Z}$;

   (iii$_2$) S is irreducible over $\overline{\mathbb{Q}}$;

   (iii$_3$) S is not contained in $\pi(V)$.

If the process succeeds, then one has to check finitely many remaining cases, namely,

   • for finitely many primes $p_1, \ldots, p_s$ one has to check whether C is irreducible over $\overline{\mathbb{F}}_{p_i}$ (cases of bad reduction);

   • for all small fields $\mathbb{F}_q$ with $q < q_0(g, d, k)$ one has to find a nontrivial $\widetilde{\phi}$-preperiodic point in $\mathrm{SL}(2, q) \times \mathrm{SL}(2, q)$.

All these steps can be performed by computer.

Example 4.3.
The process was performed for the sequence s_n. The curve Σ contains a line in $\pi(V)$ and two curves, each of genus 1 with three punctures. Thus, each of these two curves has a point over every field $\mathbb{F}_q$, $q > 9$.

Example 4.4.
Using this process, it was proved that the 2-valent sequence $r_1(x, y) = x$, $r_{n+1}(x, y) = \left[y^2 x y^{-2}, x^{-1}\right]$ is good as well. Moreover, it was checked for the Suzuki groups too. Thus it characterizes finite solvable groups.

4.1.4. 3-valent sequences

An example of a 3-valent sequence characterizing finite solvable groups is sequence (3). Consider any 3-valent sequence $v_n = (w, f)$. As explained above, it gives rise to a morphism $\phi \colon \mathcal{G}^3 \to \mathcal{G}$ and an endomorphism $\widetilde{\phi} \colon \mathcal{G}^3 \to \mathcal{G}^3$, $\widetilde{\phi}(x, y, z) = (x, y, f(x, y, z))$. The sequence is good if for every q there exists $m = m(q)$ such that there is a solution in $\mathrm{SL}(2, q) \times \mathrm{SL}(2, q)$ to the equation

$$v_1(x, y) = v_m(x, y) \neq 1.$$

This means that there is a pair $(x, y) \in \mathrm{SL}(2, q) \times \mathrm{SL}(2, q)$ such that

$$\widetilde{\phi}^m(x, y, w(x, y)) = (x, y, w(x, y)) \notin V_q,$$

where $V_q = \mathrm{SL}(2, q) \times \mathrm{SL}(2, q) \times \{\mathrm{id}\}$ denotes the forbidden set.

As in Theorem 2.1, we express the trace of $\phi(x, y, z)$ as a polynomial in seven variables $a_1 = \mathrm{tr}\, x$, $a_2 = \mathrm{tr}\, y$, $a_3 = \mathrm{tr}\, z$, $a_{12} = \mathrm{tr}(xy)$, $a_{13} = \mathrm{tr}(xz)$, $a_{23} = \mathrm{tr}(yz)$, $a_{123} = \mathrm{tr}(xyz)$. These variables are dependent (see, e.g., [89] or formulas (2.3)–(2.5) in [56]):

$$a_{123}^2 - a_{123}(a_{12}a_3 + a_{13}a_2 + a_{23}a_1 - a_1 a_2 a_3) + \left(a_1^2 + a_2^2 + a_3^2 + a_{12}^2 + a_{13}^2 + a_{23}^2 - a_1 a_2 a_{12} - a_1 a_3 a_{13} - a_2 a_3 a_{23} + a_{12} a_{13} a_{23} - 4\right) = 0.$$

Let $\overline{a} = (a_1, a_2, a_3, a_{12}, a_{13}, a_{23}, a_{123}) \in \mathbb{A}^7$. Let $\pi(x, u, y) = \overline{a} \in Z$ be the trace projection. Let $Z = \pi(\mathcal{G}^3) \subset \mathbb{A}^7$. Then we have a commutative diagram

$$
\begin{array}{ccc}
\widetilde{G} \times \widetilde{G} \times \widetilde{G} & \xrightarrow{\ \widetilde{\varphi}\ } & \widetilde{G} \times \widetilde{G} \times \widetilde{G} \\
\Big\downarrow{\scriptstyle \pi} & & \Big\downarrow{\scriptstyle \pi} \\
Z & \xrightarrow{\ \psi\ } & Z
\end{array}
\tag{11}
$$

where $\psi(\overline{a}) = (a_1, a_2, l_1(\overline{a}), a_{12}, l_2(\overline{a}), l_3(\overline{a}), l_4(\overline{a}))$,

$$l_1 = \operatorname{tr}\varphi(x, y, z), \qquad l_2 = \operatorname{tr}(\varphi(x, y, z)x), \qquad l_3 = \operatorname{tr}(\varphi(x, y, z)y), \qquad l_4 = \operatorname{tr}(\varphi(x, y, z)xy).$$

In [13] it is proved that Z is an irreducible hypersurface over any algebraically closed field and that π is surjective for every $\mathbb{F}_q$. The dimension of the variety $F(\varphi) \subset Z$ of fixed points of ψ is at least 3. The additional condition $z = w(x, y)$ defines a 3-dimensional affine subset $W(w) \subset Z$.

The sequence is good if $\big((F(\varphi) \cap W(w)) \setminus V\big)(\mathbb{F}_q) \neq \emptyset$ for all q big enough. Since $\dim Z = 6$, $\dim F(\varphi) = 3$, $\dim W(w) = 3$, one should expect that $F(\varphi) \cap W(w)$ is zero-dimensional. However, it turns out that for the sequence u_n defined by (3) it is an absolutely irreducible curve! Thus we can formulate a sufficient condition for a 3-valent sequence to be good.

Theorem 4.5.
Let w, f define a sequence $\{v_n(x, y)\}$. Let $F(\varphi)$ be the variety of fixed points of the trace map ψ of the corresponding endomorphism $\widetilde{\varphi}$ (see diagram (11)), and let $W(w)$ be defined by w. Let $V = \{a_2 = 2, a_1 = a_{12}, a_3 = a_{23}, a_{13} = a_{123}\}$. Assume that $F(\varphi) \cap W(w)$ contains a positive dimensional, absolutely irreducible $\mathbb{Q}$-subvariety $\Phi \not\subset V$. Then there is q_0 such that for every $q > q_0$ there exists a $\widetilde{\phi}$-preperiodic point in $\mathrm{SL}(2, q)^3 \setminus V_q$.

Once again, this theorem provides a finite process, which may be performed by computer, determining whether a sequence is good.

It is a conceptual challenge to understand whether or not the property of a sequence to characterize finite solvable groups is generic in some reasonable sense. We suspect that this question can be answered in the affirmative, in the sense that for almost every law $f(x, y, z)$, satisfying necessary conditions, there exists a first word $w(x, y)$ such that the resulting sequence is as required. Of course, one has to make more precise what is meant by "almost every law". See Conjecture 7.1 below.

4.2. Mapping tori of endomorphisms of free groups

Below we describe in brief another spectacular application of arithmetic dynamics to group theory, following [18].

Given a group G with generators $x_1, \ldots, x_d$ and a relation set R, and its injective endomorphism ϕ taking x_i to a word $w_i \in \mathcal{F}_d$, $i = 1, \ldots, d$, the mapping torus T of ϕ is defined as the group extension of G obtained by adding a generator t subject to the relations $tx_it^{-1} = w_i$. One can ask whether T is residually finite (this property means that the intersection of the subgroups of finite index of T is trivial). We refer to [64] and [18] for the history and context of this problem.

The following theorem answers this question in affirmative in the case where the set R is empty, i.e., G is a free group.

Theorem 4.6 ([18]).
The mapping torus of any injective endomorphism of a free group is residually finite.

Here are the main steps of the proof.

- By the definition of the mapping torus of an injective endomorphism, it is enough to prove that for any word $w \in \mathcal{F}_d$ and any positive integer a there is a homomorphism $h \colon T \to H$ to a finite group H such that

$$h(t^a w t^{-a}) = h\big(\phi^{(a)}(w)\big) \neq \mathrm{id}. \tag{12}$$

- Any homomorphism h_g of $\mathcal{F}_d$ to $\mathrm{SL}(2)$ is defined by a point $g = (g_1, \ldots, g_d) \in \mathrm{SL}(2)^d$. Then

$$h_g(x_i) = g_i, \qquad h_g(w(x_1, \ldots, x_d)) = w(g_1, \ldots, g_d).$$

- Define $\Phi \colon \mathrm{SL}(2)^d \to \mathrm{SL}(2)^d$ by $\Phi(s_1, \ldots, s_d) = (w_1(s_1, \ldots, s_d), \ldots, w_d(s_1, \ldots, s_d))$. Then

$$h_g\big(\phi^{(a)}(w)\big) = w\big(\Phi^{(a)}(g_1, \ldots, g_d)\big). \tag{13}$$

- Take a field $\mathbb{F}_q$ big enough and consider the endomorphism $\Phi_q \colon \mathrm{SL}(2, q)^d \to \mathrm{SL}(2, q)^d$. Then (13) is still valid.

- Thus any point $g \in \mathrm{SL}(2, q)^d$ such that
  - it is periodic for Φ_q,
  - $\pi_w(g) = w(g) = w(g_1, \ldots, g_d) \neq 1$,

  will have property (12).

- Let $V \subset \mathrm{SL}(2)^d$ be the Zariski closure of $\Phi^{(4d)}(\mathrm{SL}(2)^d)$. Then

  - V is Φ-invariant,
  - $\Phi\!\restriction_V \colon V \to V$ is dominant,
  - $\pi_w(V) \neq \{\mathrm{id}, -\mathrm{id}\}$.

- Let $Z = V \setminus \pi_w^{-1}(\{\mathrm{id}, -\mathrm{id}\})$ and fix $\mathbb{F}_q$. By a theorem of Hrushovski [57, Corollary 1.2], there is an extension $\mathbb{F}_{q_1}$ of $\mathbb{F}_q$ and a Φ_{q_1}-periodic point $g \in Z(\mathbb{F}_{q_1})$.

In order to avoid Hrushovski's theorem, the authors prove its particular case for endomorphisms of affine space. That is why they consider $\mathrm{SL}(2)$ as a subscheme of the scheme M_2 of all 2×2 matrices. Note, however, that using Hrushovski's theorem, they obtain a stronger result where residual finiteness is established for the mapping torus of any endomorphism of any finitely generated linear group. Moreover, a further refinement of their method in [19], involving, in particular, search of periodic points of self-maps defined over p-adic fields, allows one to get more precise information on the structure of the mapping tori.

Remark 4.7.
Note that in the framework of the dynamical approach presented above it is permitted in Theorem 4.6 to look for periodic points defined over some extension of the original ground field. This is a subtle but important difference from the method described in subsection 4.1 where such an extension is forbidden, which prevents from using Hrushovski's theorem.

5. Word maps: image and fibres

In this section we focus on the case of finite simple groups asking the following questions:

- How big is the image of a word map?
- What are the sizes of the fibres of a word map?

These questions are interrelated: one can prove that the image is large by estimating the sizes of fibres. We start, however, with recalling a seminal result by Borel [17] where a general answer to the first question was obtained for connected semisimple algebraic groups.

5.1. Borel's theorem

Theorem 5.1 ([17]).
Let G be a connected semisimple algebraic group defined over a field K. Let $f_w \colon G^d \to G$ be the map associated to a nontrivial element w of the free group on $d \geq 2$ letters. Then f_w is dominant, i.e., its image is Zariski dense in G.

Here are the main steps of the proof.

   (I) Since dominance is preserved by field extension, one may assume that K is algebraically closed of arbitrary transcendence degree.

  (II) One can show that the assertion of the theorem does not depend on the choice of G within its isogeny class, so one may assume G is simply connected. One can also easily reduce to the case where G is simple.

 (III) There is a d–tuple $(g_1, \ldots, g_d) \in G^d$ such that $f_w(g_1, \ldots, g_d) \neq 1$ because a simple group has no identical relations, in view of a theorem of Platonov [103] stating that a linear group with an identical relation must be solvable–by–finite (this can also be deduced from the Tits alternative [122]).

 (IV) First consider the case $G = SL_n$.

   - The following observation, going back to the unitary trick of Weyl and used in [26], is crucial: one can find a subfield $L \subset K$ and a division L-algebra D so that the group $G = SL_n(K)$ contains its anisotropic form $H = SL_1(D)$, the group of elements of D of reduced norm 1, as a dense subset.
   - One proves that
     $$\text{if}\quad \mathrm{id} \neq h \in H, \quad \text{then}\quad 1 \text{ is not an eigenvalue of } h. \tag{14}$$
   - Let $\overline{\mathrm{im}\, f_w} = Z$ be the closure of $\mathrm{im}\, f_w$. One should prove that $Z = SL(n, K)$. Since H is dense in G,
     $$\{0\} \neq f_w(H^d) \subset Z \cap H. \tag{15}$$
   - Use induction on n. If $n = 2$ and $Z \neq G$, then $\dim Z \leq 2$ and Z is a union of a finite number of conjugacy classes of (non-identity) semisimple elements and the set U consisting of unipotents. Since $\mathrm{id} \in Z$ and Z is irreducible, it is contained in U, which contradicts (14). Hence, for $n = 2$ the statement is valid.
   - Assume that the statement is proved for $n \leq m - 1$. Fix a maximal torus $T \subset SL(m, K)$. Let $T' \subset T$ be the union of subtori consisting of elements with at least one eigenvalue 1. T' is a hypersurface in T. By the induction hypothesis, $Z \supseteq T'$. On the other hand, by (15), there is a d-tuple $(h_1, \ldots, h_d) \in H^d$ such that $f_w(h_1, \ldots, h_d) \in T \setminus T'$. Since $SL(m, K)^d$ is an irreducible variety, Z should be irreducible as well. Therefore, if it contains a hypersurface T' and at least one point outside T', it contains T. Thus, $Z = SL(m, K)$ since the conjugates of T are dense in $SL(m, K)$ and $\mathrm{im}\, f_w$ is invariant under conjugation.

  (V) Any simple group of rank r not isogenous to SL_n contains a subgroup of the same rank r which is isogenous to a direct product of groups SL_{n_i}, and the assertion of the theorem follows from the previous step.

Remark 5.2.
See [61] for an alternative proof based on Amitsur's theorem on generic division rings [3].

5.2. The image of the word map on finite simple groups

From now on G is a finite simple group, $w(x_1, \ldots, x_d) \in \mathcal{F}_d$ is a nontrivial word, and we shorten our previous notation so that $w \colon G^d \to G$ denotes the corresponding word map. It turns out that one can provide an analogue of Borel's theorem. Although dominance does not make any sense in this context, it was shown by Larsen [68] that the image of the word map is large.

Theorem 5.3 ([68]).

For every nontrivial word w and any $\epsilon > 0$ there exists $N = N(w, \epsilon)$ such that if G is a finite simple group of order greater than N, then

$$\# w(G) \geq \# G^{1-\epsilon}.$$

The original proof of this theorem heavily relied on the techniques of Larsen–Pink [69] for estimating the sizes of the fibres of w in the case where G is of Lie type. The case $G = A_n$ always requires separate consideration, and sporadic groups can be ignored whenever one restricts attention to asymptotic questions.

Later on, this theorem was reproved by Larsen and Shalev [71] in the case of groups of Lie type using more traditional methods of arithmetic geometry, such as Lefschetz's trace formula (not including, however, Suzuki and Ree groups, successfully treated in [69]). In the same paper [71], Larsen and Shalev considered some other variations on estimating the size of $w(G)$. Namely, if G is a group of Lie type (different from A_r or ${}^2A_r$) of *fixed Lie rank r*, they obtained an estimate of the form

$$\# w(G) > cr^{-1} \# G.$$

Here c is a positive absolute constant and G is of order greater than $N(w)$. For the alternating groups the estimate is slightly weaker.

As there is no hope to establish surjectivity of the map w for arbitrary words (power maps provide easy counter-examples), one can try to say something more about $w(G)$. The following terminology, reminiscent of classical number theory, was introduced by Shalev.

5.2.1.   Waring type properties [66, 71, 73, 74, 114]

Waring's problem deals with expressing every natural number as a sum of $f(k)$ k^{th} powers for some suitable function f. Noncommutative analogues of this problem were investigated during the past years, answering the following questions: Can one write any element of a finite simple group as a product of $f(k)$ k^{th} powers? (See [93, 109].) Can this result be extended by replacing the word x^k with an arbitrary nontrivial word w? Can these results be improved by replacing the function f with a global (small) constant? See [73] and the references therein, and the most recent improvements in [55].

A prototypical result here is due to Liebeck and Shalev [83]: if G is a finite simple group and w is not trivial on G, then there is a constant $c = c(w)$ such that every element in G is the product of c values of w on G. Using lots of various methods and ingenious techniques, Shalev showed in [114] that for every nontrivial word w there exists a constant $N(w)$ such that if G is a finite simple group of order greater than $N(w)$ then $w^3(G) = G$. Two alternative proofs of this result were recently found. The first, due to Nikolov and Pyber [98], is based on a recent result of Gowers [50], and the second, for finite simple groups of bounded Lie rank, due to Macpherson and Tent [88], relies on model theory. This result was substantially improved by Larsen, Shalev and Tiep in [70, 71, 73].

Theorem 5.4 ([73]).

For any nontrivial word w there exists a constant $N(w)$ such that for all finite non-abelian simple groups G of order greater than $N(w)$ we have $w^2(G) = G$.

The particular case of $w = x^k$ shows that this is the best possible Waring type result for powers.

Conjecture 5.5 (Shalev, [113, Conjectures 2.8 and 2.9]).

Let $w \neq 1$ be a word which is not a proper power of another word. Then there exists a number $N(w)$ such that, if G is either A_n or a finite simple group of Lie type of rank n, where $n > N(w)$, then $w(G) = G$.

Note that the assumption on the rank of G cannot be removed, in view of counter-examples presented in [60] (see the footnote to Conjecture 7.14).

A recent result of Kassabov and Nikolov [66] shows that the assumption in Theorem 5.4 that G is sufficiently large cannot be removed, even if we only require that $G = w^k(G)$ for a fixed k. Indeed, it is shown in [66] that for any integer k

there exist a word w and a finite simple group G, such that w is not an identity in G, but $G \neq w^k(G)$. This is done by constructing for any $n > 13$ a specific word $w \in \mathcal{F}_2$ such that $w(A_n)$ consists of the identity and all 3-cycles. The result follows since for $n > 2k+1$ there are elements in A_n which cannot be written as a product of less than $k+1$ 3-cycles.

Further examples of word maps on $\mathrm{SL}(2, 2^n)$ whose image is very small (consisting of the identity and a single conjugacy class) have been constructed in [75]. In a more recent preprint [85], Lubotzky proved (assuming the classification of finite simple groups) that *any* given subset of a finite simple group G which contains the identity and is invariant under $\mathrm{Aut}(G)$ can arise as the image of some word map. In [76], this result was extended to some almost simple and quasisimple groups.

The main tools in [71, 73, 114] involve representation theory, algebraic geometry and probabilistic methods. For any two nontrivial words w_1, w_2 the rough idea is to construct special conjugacy classes $C_1, C_2 \subset G$ satisfying

$$C_1 \subset w_1(G), \qquad C_2 \subset w_2(G), \tag{16}$$

$$C_1 C_2 \supseteq G \setminus \{1\}. \tag{17}$$

The proof of (17) relies on the following classical result of Frobenius. Let $C_1 = s_1^G$, $C_2 = s_2^G$. The number of ways to write a group element $g \in G$ as $g = x_1 x_2$, where $x_i \in C_i$, is given by

$$\frac{\# C_1 \# C_2}{\# G} \sum_{\chi \in \mathrm{Irr}(G)} \frac{\chi(s_1) \chi(s_2) \overline{\chi}(g)}{\chi(1)}, \tag{18}$$

where $\mathrm{Irr}(G)$ denotes the set of irreducible complex characters of G.

The case of the alternating groups was established by Larsen and Shalev in [71]. First, it was proved that for any $w \neq 1$ there exists $N(w)$ such that if $n \geq N(w)$ then the image $w(A_n)$ contains the conjugacy classes of permutations s_n with a few cycles (at most 23), thus implying (16). This is highly non-elementary, involving algebraic geometry and results from analytic number theory (such as weak versions of the Goldbach Conjecture). The idea is to embed groups of the form $\mathrm{SL}(2, p)$ and their products into A_n, basing on the fact that $\mathrm{SL}(2, p)$ embeds into A_{p+1}, and an element of order $(p-1)/2$ in $\mathrm{SL}(2, p)$ has two nontrivial cycles and two fixed points in this embedding, and then use the following property of word maps on $\mathrm{SL}(2, p)$.

Theorem 5.6 ([71, Theorem 4.1]).
For every nontrivial word w there exist constants M_w and m_w with the following property: for every prime $p > M_w$, such that $p-1$ is divisible neither by 4 nor by any prime $3 \leq l \leq m_w$, $w(\mathrm{SL}(2, p))$ contains an element of order $(p-1)/2$.

Inclusion (17) can now be obtained by combining (18) with the fact that all character values $\chi(s)$ of a permutation $s \in S_n$ can be bounded in terms of the number of cycles alone.

Theorem 5.7 ([71, Theorem 7.2]).
Let $s \in S_n$ be a permutation with k cycles (including 1-cycles). Then

$$|\chi(s)| \leq 2^{k-1} k!$$

for all irreducible characters χ of S_n.

Larsen and Shalev [71] also treated the case of finite simple groups of Lie type of bounded rank. Later on, Larsen, Shalev and Tiep [73] completed the proof for finite simple groups of Lie type of arbitrary rank.

For finite simple groups of Lie type, C_1, C_2 are the classes of suitable regular semisimple elements $s_1, s_2 \in G$ lying in maximal tori $T_1, T_2 \subset G$. The tori T_i are chosen so that if χ is an irreducible character of G such that $\chi(s_1)\chi(s_2) \neq 0$ then χ is unipotent; moreover, there is a small (in particular, bounded) number of such unipotent characters. These results are obtained using the machinery of Deligne–Lusztig, see, e.g., [27]. This implies that the number of nonzero summands in (18) is small, and moreover, the character ratios can be bounded.

Theorem 5.8 ([73, Theorem 1.2.1]).

If G is a finite quasisimple classical group over $\mathbb{F}_q$ and $g \in G$ is an element of support at least N, then

$$\frac{|\chi(g)|}{\chi(1)} < q^{-\sqrt{N}/481}$$

for all $1_G \neq \chi \in \mathrm{Irr}(G)$.

(Here the support of g is defined as the codimension of its largest eigenspace, see [73, Definition 4.1.1]. Recall that a *quasisimple* group G is a perfect group such that $G/Z(G)$ is simple.)

The proof of (16) is based on geometric tools, and in particular, on the Lang–Weil estimate, that allows one to establish a Chebotarev Density Theorem for word maps.

Theorem 5.9 ([73, Corollary 5.3.3]).

For every fixed nontrivial word w and fixed integer N, there exists $\delta > 0$ such that for every semisimple algebraic group $\mathbb{G}$ of dimension less than N over a finite field $\mathbb{F}_q$ and every maximal torus $\mathbb{T}$ of $\mathbb{G}$ defined over $\mathbb{F}_q$, we have

$$\# \left(\mathbb{T}(\mathbb{F}_q) \cap w(\mathbb{G}(\mathbb{F}_q)) \right) \geq \delta \# \mathbb{T}(\mathbb{F}_q).$$

Hence, if q is sufficiently large, there exist regular semisimple elements $s_i \in w_i(\mathbb{G}(\mathbb{F}_q))$ lying in any prescribed maximal torus $\mathbb{T}(\mathbb{F}_q)$. This itself is not enough, since the group G is of unbounded Lie rank. This obstacle is treated by embedding groups H of very small rank (such as SL_2) over large extension fields into G so that $s_i \in w_i(H)$ remains regular semisimple in G, and lies in the required maximal torus T_i of G. Clearly $s_i \in w_i(G)$ so that $w_i(G)$ contains the conjugacy class $C_i = s_i^G$, as required.

Similar Waring type results were obtained by Larsen, Shalev and Tiep in [74] for quasisimple groups.

Theorem 5.10 ([74]).

For a fixed nontrivial word w there exists a constant $N(w)$ such that if G is a finite quasisimple group of order greater than $N(w)$, then $w^3(G) = G$.

For various families of finite quasisimple groups, including covers of alternating groups, a stronger result was proved in [74], namely that $w^2(G) = G$. This was recently finalized by Guralnick and Tiep [55], who proved that for any nontrivial word w there exists $N = N(w)$ such that $w^2(G) \supseteq G \setminus Z(G)$ for all quasisimple groups G of order greater than N. Note, however, that in contrast with the case of simple groups studied in [73], the equality $w^2(G) = G$ may not hold for all large finite quasisimple groups G. The nontrivial central elements of finite quasisimple groups G provide the main obstructions (see subsections 5.2.3 and 6.1 for examples).

Further variations on the Waring theme, discussed in [71, 113], include considering products $w_1(G) \cdots w_k(G)$ and intersections $w_1(G) \cap \cdots \cap w_k(G)$ where w_i are *distinct* words. The latter case can be fit into the same framework by looking at the fibre product of the word maps $w_i \colon G^d \to G$ over G (a different approach was suggested in [98]). The reader is addressed to the original papers for details.

Some counterparts of Waring type properties discussed above can be formulated for maps of matrix algebras induced by associative noncommutative polynomials, see [61] for a survey.

Remark 5.11.

In [72], Larsen and Shalev obtained a general estimate for the size $\# N_w(g)$ of the fibres of word maps: for *all* $w \in \mathcal{F}^d$, $w \neq 1$, there exists $\varepsilon > 0$ such that for *all* finite simple groups G and *all* $g \in G$ we have $\# N_w(g) = O(\# G^{d-\varepsilon})$, where the implicit constant depends only on w. Naturally, these universal estimates are rough in comparison with the equidistribution results because they hold for *all* nontrivial words, including power words, which are far from being equidistributed, and also for *all* elements in the group, including 1, for which the fibre may be very large (as in the case of commutators).

Remark 5.12.

In a different spirit, estimates for the size of the fibres of word maps were used in [1, 99], where they yielded new criteria for distinguishing finite nilpotent and solvable groups. (We thank the first referee for pointing out the references mentioned above.)

5.2.2. Commutators [34, 53, 79, 80, 102, 120]

In this and next sections we consider the image of word maps for some special words w. First note that for any primitive word w (this means that w is a part of a free generating set of $\mathcal{F}_d$), as well as for any word of the form $w = x_1^{e_1} \ldots x_d^{e_d} f$, where the e_i are coprime and f belongs to the derived group $\mathcal{F}_d'$, the induced map $w \colon G^d \to G$ is surjective for an arbitrary group G (see, e.g., [111, 3.1.1]). The commutator word is the first nontrivial instance of the surjectivity problem.

Theorem 5.13 (Ore's Conjecture).

If G is a finite non-abelian simple group, then every element of G is a commutator.

In other words, for the commutator word $w = [x, y] \in \mathcal{F}_2$, one has $w(G) = G$ for any finite non-abelian simple group G. This statement was originally posed in 1951 and proved by Ore himself for the alternating groups [102]. During the years, this conjecture was proved for various families of finite simple groups (see [79] and the references therein). R. Thompson [120] established it for the linear groups $\mathrm{PSL}(n, q)$, later Ellers and Gordeev [34] proved the conjecture for all finite simple groups of Lie type defined over a field with more than 8 elements, and recently an impressive full stop was put by Liebeck, O'Brien, Shalev and Tiep [79] who completed the proof for all finite simple groups.

The original proofs of Ore [102] and R. Thompson [120] were obtained by explicitly finding pairs of permutations (respectively, matrices) whose commutator corresponded to some representative in any given conjugacy class.

In order to complete the proof of Ore's Conjecture, Liebeck, O'Brien, Shalev and Tiep used in [79] the following classical criterion dating back to Frobenius, that the number of ways to write an element g in a finite group G as a commutator is

$$\# G \sum_{\chi \in \mathrm{Irr}(G)} \frac{\chi(g)}{\chi(1)}. \tag{19}$$

Roughly speaking, it was shown that if g is an element with a small centralizer, then $\chi(g)/\chi(1)$ is small for $\chi \neq 1$, and the main contribution to the character sum (19) comes from the trivial character $\chi = 1$. Hence, this sum is positive, so elements with small centralizers are commutators. This is based on the Deligne–Lusztig theory, and also on the theory of dual pairs and Weil characters of classical groups [54, 121]. For elements whose centralizers are not small, the strategy is to reduce to groups of Lie type of lower dimension and use induction. Namely, if a certain element has a Jordan decomposition into several Jordan blocks, and if it is possible to express each block as a commutator in the smaller classical group, then clearly the original element is itself a commutator.

Computer calculations (using MAGMA) played a significant role in the proof of [79]. Since the proof uses induction, it was necessary to establish various base cases. The conjecture was proved directly for many of these base cases by constructing the character table of the relevant group. For various other groups certain elements with prescribed Jordan forms as commutators were explicitly constructed. Similar methods were used in the subsequent paper [80], in which it was shown that with a few (small) exceptions, every element of a finite quasisimple group is a commutator, and moreover, any such element is a product of two commutators.

Ellers and Gordeev [34] have proved, for the finite simple groups of Lie type over fields with more than eight elements, a stronger conjecture, known as *Thompson's Conjecture*.

Conjecture 5.14 (Thompson's Conjecture).

Every finite simple group G has a conjugacy class C such that $C^2 = G$.

Observe that Thompson's Conjecture immediately implies Ore's Conjecture. Indeed, if $C^2 = G$ then $1 \in C^2$ so $C^{-1} = C$ and $G = CC^{-1}$. Hence, for any $g \in G$ there exist $x \in C$ and $h \in G$ such that $g = x^h x^{-1} = [h, x]$, as required.

Thompson's conjecture was verified for many families of finite simple groups, including the alternating groups and the sporadic groups, see the introduction of [34], but nevertheless it is still very much open today.

The proof of Ellers and Gordeev is based on the following generalization of the Gauss decomposition of matrices.

Theorem 5.15 ([31–33]).
Let G be a Chevalley group, and let Γ be a group generated by G and a cyclic group $\langle \sigma \rangle$ which normalizes G in Γ and acts as a diagonal automorphism on G (perhaps trivially).

Let $\gamma = \sigma g \in \Gamma \setminus Z(\Gamma)$. If h is any fixed element in the group H, then there is $\tau \in G$ such that $\tau \gamma \tau^{-1} = \sigma u_1 h u_2$, where $u_1 \in U^-$ and $u_2 \in U$.

Here H, U and U^- are the subgroups of G defined by

$$H = \langle h_\alpha : \alpha \in \Pi \rangle, \qquad U = \langle X_\alpha : \alpha \in \Phi^+ \rangle, \qquad U^- = \langle X_\alpha : \alpha \in \Phi^- \rangle,$$

where Φ is the root system corresponding to G and Π denotes the simple roots of Φ. Recall that the Chevalley group G is generated by the root subgroups X_α, $\alpha \in \Phi$.

As a consequence of Theorem 5.15, one easily gets a statement in the spirit of inclusion (17).

Corollary 5.16.
If $h_1, h_2 \in H$ are regular semisimple elements in G from a maximal split torus and C_1, C_2 are the respective conjugacy classes, then $C_1 C_2 \supseteq G \setminus Z(G)$.

Indeed, by [31, Proposition 1], for fixed h_1, h_2 any $u_1 \in U^-$ and $u_2 \in U$ can be represented as

$$u_1 = v_1 h_1 v_1^{-1} h_1^{-1} \qquad \text{and} \qquad u_2 = h_2^{-1} v_2 h_2 v_2^{-1},$$

for some $v_1 \in U^-$ and $v_2 \in U$. Thus by Theorem 5.15, for any noncentral conjugacy class $C \subset G$ one can find a representative $c \in C$ such that

$$c = u_1 h_1 h_2 u_2 = (v_1 h_1 v_1^{-1} h_1^{-1}) h_1 h_2 (h_2^{-1} v_2 h_2 v_2^{-1}) = (v_1 h_1 v_1^{-1})(v_2 h_2 v_2^{-1}).$$

Corollary 5.16 immediately implies Ore's conjecture for any simple group G containing a regular semisimple element h in a maximal split torus, and Thompson's Conjecture if this element is in addition real, i.e., if h and h^{-1} are conjugate. In [34] a careful analysis is done to show that such desired elements actually exist in groups of Lie type over fields with more than eight elements.

In addition, Guralnick and Malle [53] have extended the aforementioned result (17) from [73] and proved the following variant of Thompson's Conjecture.

Theorem 5.17 ([53, Theorem 1.4]).
If G is a finite non-abelian simple group, then there exist conjugacy classes C_1, C_2 in G with $G \setminus \{1\} = C_1 C_2$. Moreover, aside from $G = \mathrm{PSL}(2, 7)$ or $\mathrm{PSL}(2, 17)$, one can assume that each C_i consists of elements of order prime to 6.

Similarly to [73], the proof in [53] also relies on estimating the character sum (18) using the Deligne–Lusztig theory, or for some small rank groups it is computed directly from known character tables, to show that triples (x_1, x_2, g) of elements from specified conjugacy classes C_i exist in a given group G. The conjugacy classes C_i are chosen so that only few irreducible characters vanish simultaneously on these classes. These triples moreover generate G, since the conjugacy classes C_i were chosen so that their elements are contained in few maximal subgroups of G.

After considering the commutator word, it is natural to go over to the Engel words, defined recursively by

$$e_1(x, y) = [x, y] = xyx^{-1}y^{-1}, \qquad e_n(x, y) = [e_{n-1}, y],$$

and the corresponding Engel word maps $e_n \colon G \times G \to G$. The following conjecture is naturally raised.

Conjecture 5.18 (Shalev).
Let $n \in \mathbb{N}$, then the nth Engel word map is surjective for any finite simple non-abelian group G.

See Section 6 for the cases $G = \mathrm{SL}(2, q)$ and $\mathrm{PSL}(2, q)$.

5.2.3. Two-power words [53, 73, 74, 81]

It follows from Theorem 5.4, due to Larsen, Shalev and Tiep [73], that any two-power word is surjective on sufficiently large finite simple groups (see [73, Theorem 1.1.1 and Corollary 1.1.3]). More precisely,

Theorem 5.19 ([73]).
Let a, b be two nonzero integers. Then there exists a number $N = N(a, b)$ such that if G is a finite non-abelian simple group of order at least N, then any element in G can be written as $x^a y^b$ for some $x, y \in G$.

Furthermore, by recent results of Liebeck, O'Brien, Shalev and Tiep [81] and of Guralnick and Malle [53] (see Theorem 5.17), some words of the form $x^b y^b$ are known to be surjective on *all* finite simple groups.

Theorem 5.20 ([53, Corollary 1.5]).
Let G be a finite non-abelian simple group and let b be either a prime power or a power of 6. Then any element in G can be written as $x^b y^b$ for some $x, y \in G$.

Note that in general, the word $x^b y^b$ is not necessarily surjective on *all* finite simple groups. Indeed, if b is a multiple of the exponent of G then necessarily $x^b y^b = 1$ for every $x, y \in G$.

In addition, by another recent work of Larsen, Shalev and Tiep [74], if G is a finite quasisimple group, then the word $w = x^2 y^2$ is surjective on G. On the other hand, if $b > 2$ then the word $w = x^b y^b$ is *not* surjective on infinitely many finite quasisimple groups.

5.3. The fibres of the word map

Recall that estimating the sizes of the fibres of the word map appeared as an integral part of estimating its image. In this section we address a subtler problem trying to distinguish words for which the fibres of the corresponding word map are of the same size, at least approximately. First note that for certain words w all fibres of the word map $w \colon G^d \to G$ are exactly of the same size *for any finite group G*. According to recent results of [104, 105], this holds only for primitive words. Primitive words are asymptotically very rare (exponentially negligible, in the terminology of [65]): if we count them among all words of fixed length, their proportion tends to 0 exponentially fast (see, e.g., [95]). Another viewpoint at the set of primitive elements of $\mathcal{F}_d$ is that this set is closed in the profinite topology of $\mathcal{F}_d$ [105].

We are interested in weaker equidistribution properties which hold for more general words.

5.3.1.   Equidistribution and measure-preservation [41, 80]

For a word $w = w(x_1, \ldots, x_d) \in \mathcal{F}_d$, a finite group G and some $g \in G$, we denote

$$N_w(g) = \{(g_1, \ldots, g_d) \in G^d : w(g_1, \ldots, g_d) = g\}.$$

Definition 5.21 ([41]).
A word map $w : G^d \to G$ is *almost equidistributed* for a family of finite groups $\mathcal{G}$ if any group $G \in \mathcal{G}$ contains a subset $S = S_G \subseteq G$ with the following properties:

  (i) $\# S = \# G(1 - \epsilon(G))$,
 (ii) $\# N_w(g) = (\# G)^{d-1}(1 + \epsilon(G))$ uniformly for all $g \in S$,

where $\epsilon(G) \to 0$ whenever $\# G \to \infty$.

Theorem 5.22 ([41, Theorem 1.5]).
The commutator word $w = [x, y] \in \mathcal{F}_2$ is almost equidistributed for the family of finite simple groups.

Note that we cannot require in this theorem that $S = G$. Indeed, it is well known (and follows from (19) above) that for $w = [x, y]$ we have $N_w(1) = k(G)\# G$ where $k(G)$ is the number of conjugacy classes in G. Since $k(G) \to \infty$ as $\# G \to \infty$ we see that the fibre above $g = 1$ is large and does not satisfy condition (ii).

Two proofs are given in [41] for this theorem. The first is probabilistic whereas in the second the subsets S are explicitly constructed.

Let $P = P^G$ be the commutator distribution on G, namely $P(g) = N_w(g)/\# G^2$, and let $U = U^G$ be the uniform distribution on G (so $U(g) = 1/\# G$ for all $g \in G$). The probabilistic proof bounds the L_1-distance

$$\|P - U\|_1 = \sum_{g \in G} |P(g) - U(g)|$$

between the probability measures above. Using the Frobenius formula (19) and the Cauchy–Schwarz inequality, it is deduced in [41, Proposition 1.1] that

$$\|P^G - U^G\|_1 \leq (\zeta^G(2) - 1)^{1/2}, \qquad \text{where} \quad \zeta^G(s) = \sum_{\chi \in \mathrm{Irr}(G)} \chi^{-s}(1)$$

is the so-called *Witten zeta function*.

Now Theorem 5.22 follows from results of Liebeck and Shalev [84], who showed that if G is simple, then

$$\zeta^G(s) \to 1 \qquad \text{as} \quad \# G \to \infty \quad \text{provided} \quad s > 1.$$

This proof also provides some estimation of the function ϵ. Namely, $\epsilon(A_n) = O(n^{-1/2})$, and if G is of Lie type of rank r over a field with q elements then $\epsilon(G) = O(q^{-r/4})$.

The second, constructive proof describes the subsets S explicitly. If $G = \mathrm{PSL}(2, q)$ then Theorem 5.22 follows from (19) directly using the well-known character table of G. If $G \neq \mathrm{PSL}(2, q)$ varies among groups of Lie type of bounded Lie rank then $S = S_G$ is chosen as the set of regular semisimple elements of G. If G varies among groups of Lie type with unbounded Lie rank, then $S = S_G$ contains elements whose centralizer is not very large. If $G = A_n$ then $S = S_G$ is chosen as the set of permutations in A_n with at most $\sqrt{n}$ fixed points. This yields better lower bounds on the cardinality of S. For example, in the constructive proof for A_n we obtain $\# S \geq (1 - 2/[\sqrt{n}]!) \# A_n$, which is much better than the lower bound $(1 - O(n^{-1/2})) \# A_n$ given by the probabilistic proof.

Similarly, it was shown in [80, Proposition 3] that the commutator word is also almost equidistributed on the family of finite quasisimple groups.

In [41, Section 7] it is proved that the property to be almost equidistributed behaves well under direct products and compositions, implying that the words $w = [x_1, \ldots, x_d] \in \mathcal{F}_d$, d-fold commutators in any arrangement of brackets, are almost equidistributed within the family of finite simple groups. Similar methods are used in [41, Theorem 7.1] to show that the word $w = x^2 y^2$ is almost equidistributed on finite simple groups, and Larsen and Shalev have recently obtained similar results in more general contexts.

By [41, Section 3], any "almost equidistributed" word map $w \colon G^d \to G$ (see Definition 5.21) is also "almost measure preserving" in the following sense.

Definition 5.23 ([41]).

A word map $w \colon G^d \to G$ is *almost measure preserving* for a family of finite groups $\mathcal{G}$ if every group $G \in \mathcal{G}$ satisfies the following conditions:

 (i) for every subset $Y \subseteq G$ we have $\# w^{-1}(Y)/\# G^d = \# Y/\# G + \mathrm{o}(1)$;

 (ii) for every subset $X \subseteq G^d$ we have $\# w(X)/\# G \geq \# X/\# G^d - \mathrm{o}(1)$;

 (iii) in particular, if $X \subseteq G^d$ and $\# X/\# G^d = 1 - \mathrm{o}(1)$, then almost every element $g \in G$ can be written as $g = w(g_1, \ldots, g_d)$ where $(g_1, \ldots, g_d) \in X$;

here $\mathrm{o}(1)$ denotes a function depending only on G which tends to zero as $\# G \to \infty$.

This allows one to deduce in [41, Corollary 1.6] that the commutator map is almost measure preserving on the family of finite simple groups. Since almost all pairs of elements of a finite simple group are generating pairs (see [28, 63, 82]), the probability that some $g \in G$ can be represented as a commutator $g = [x, y]$, where x, y generate G, tends to 1 as $\# G \to \infty$, by [41, Theorem 1.7].

It seems to be not so easy to extend equidistribution results from commutators to more general words. See, however, the next section where this is done in the particular cases $G = \mathrm{SL}(2, q)$ and $\mathrm{PSL}(2, q)$.

6. Word maps on $\mathrm{SL}(2, q)$ and $\mathrm{PSL}(2, q)$

6.1. Surjectivity for Engel words and some positive words

In [9] the particular case of Engel words in the groups $\mathrm{PSL}(2, q)$ and $\mathrm{SL}(2, q)$ was analyzed using the trace map method, in an attempt to prove Conjecture 5.18 for $\mathrm{PSL}(2, q)$. The main idea was to check the surjectivity of the trace map on $\mathbb{A}^3(\mathbb{F}_q)$ instead of the surjectivity of the word map on $\mathrm{SL}(2, q)$.

For any $x, y \in G = \mathrm{SL}(2, q)$ denote $s = \mathrm{tr}\, x$, $t = \mathrm{tr}\, y$ and $u = \mathrm{tr}(xy)$, and define a morphism $\pi \colon G \times G \to \mathbb{A}^3_{s,u,t}$ by $\pi(x, y) = (s, u, t)$. Then the following diagram commutes:

$$
\begin{array}{ccc}
G \times G & \xrightarrow{\ \phi^n\ } & G \times G \\
\downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\
\mathbb{A}^3_{s,u,t} & \xrightarrow{\ \psi^n\ } & \mathbb{A}^3_{s,u,t}.
\end{array}
$$

In this diagram, the maps ϕ^n and ψ^n are defined recursively as follows.

 • $\phi(x, y) = ([x, y], y) \Rightarrow \psi(s, u, t) = (s_1, t, t)$, $s_1 = s^2 + u^2 + t^2 - sut - 2$;

 • $\phi^2(x, y) = ([[x, y], y], y) \Rightarrow \psi^2(s, u, t) = \psi(s_1, t, t) = (s_2, t, t)$, $s_2 = r(s_1, t) = s_1^2 + 2t^2 - s_1 t^2 - 2$; $\ldots$

- $\phi^n(x, y) = (e_n(x, y), y) \Rightarrow \psi^n(s, u, t) = \psi^{n-1}(s_1, u, t) = \ldots = \psi(s_{n-1}, t, t) = (s_n, t, t), \; s_n = r(s_{n-1}, t) = r(r(s_{n-2}, t), t) = \ldots = r^{(n-1)}(s_1, t),$

where $r(s, t) = s^2 + 2t^2 - st^2 - 2$.

Let $\pm 2 \neq a \in \mathbb{F}_q$. It is proven in [9] that a matrix $z \in SL(2, q)$ with $\operatorname{tr} z = a$ can be written as $z = e_{n+1}(x, y)$ for some $x, y \in SL(2, q)$ if and only if there is a solution $(s_1, t) \in \mathbb{F}_q^2$ of the equation $r^{(n)}(s_1, t) = a$, that is if and only if the curve $C_{n,a}$ defined over $\mathbb{F}_q$ by the equation $r^{(n)}(s_1, t) = a$ has a rational point.

In [9] it was shown that the curve $C_{n,a}$ is absolutely irreducible over any finite field $\mathbb{F}_q$. In addition, its genus satisfies the inequality $g(C_{n,a}) \leq 2^n(n-1) + 1$, and it has at most $\delta = 5 \cdot 2^n$ punctures. By Weil's inequality, for $a \neq \pm 2$ and $q > 2^{2n+3}(n-1)^2$ we have $C_{n,a}(\mathbb{F}_q) \neq \emptyset$, as required.

This implies the following results, obtained in [9].

Theorem 6.1.
The nth Engel word map is surjective on $SL(2, q) \setminus \{-\operatorname{id}\}$ (and hence on $PSL(2, q)$) provided that $q \geq q_0(n)$ is sufficiently large.

On the other hand,

Proposition 6.2.
There is an infinite family of finite fields $\mathbb{F}_q$ such that if $n \geq n_0(q)$ is large enough, then the nth Engel word map is not surjective on $SL(2, q) \setminus \{-\operatorname{id}\}$.

Proposition 6.3.
For every odd prime power q there is $n_0 = n_0(q)$ such that $e_n(x, y) \neq -\operatorname{id}$ for every $n > n_0$ and every $x, y \in SL(2, q)$.

Indeed, $z = e_{n+1}(x, y) = -\operatorname{id}$ implies that there is a solution to the equation $r^{(n)}(s_1, 0) = -2$, and then there is $c \in \mathbb{F}_{q^2}$ such that $c^{2^n} = -1$.

In certain cases, e_n is *always* surjective on $PSL(2, q)$.

Proposition 6.4.
e_n is surjective on $SL(2, 2^e) = PSL(2, 2^e)$.

Indeed, if $q = 2^e$, take $t = 0$ and then $r(s, 0) = s^2$, so $r^n(s, 0) = s^{2^n}$ is an isomorphism, implying that e_n is surjective on $PSL(2, q)$.

Proposition 6.5.
e_n is surjective on $PSL(2, q)$ if $\sqrt{2} \in \mathbb{F}_q$, $\sqrt{-1} \notin \mathbb{F}_q$.

Proposition 6.6.
If $n \leq 4$ then e_n is surjective for all groups $PSL(2, q)$.

The last result is a consequence of Theorem 6.1 and MAGMA calculations.

In [8] there is a precise description of the positive integers a, b and prime powers q for which the word map $w(x, y) = x^a y^b$ is surjective on the group $PSL(2, q)$ (and $SL(2, q)$). The proof is based on the investigation of the trace map of positive words. The key result is that $\operatorname{tr}(x^a y^b)$ is a *linear* polynomial in u, namely,

$$\operatorname{tr}(x^a y^b) = u \cdot f_{a,b}(s, t) + h_{a,b}(s, t), \qquad \text{where} \quad f_{a,b}(s, t), h_{a,b}(s, t) \in \mathbb{F}_q[s, t].$$

Thus, if neither a nor b is divisible by the exponent of $\mathrm{PSL}(2, q)$, then any element in $\mathbb{F}_q$ can be written as $\mathrm{tr}(x^a y^b)$ for some $x, y \in \mathrm{SL}(2, q)$. This immediately implies that in this case, any *semisimple* element (namely, $z \in \mathrm{SL}(2, q)$ with $\mathrm{tr}\, z \neq \pm 2$) can be written as $z = x^a y^b$ for some $x, y \in \mathrm{SL}(2, q)$. However, when z is *unipotent* (namely, $z \neq \pm\,\mathrm{id}$ and $\mathrm{tr}\, z = \pm 2$) one has to be more careful, and a detailed analysis is needed. Indeed, it may happen that neither a nor b is divisible by the exponent of $\mathrm{PSL}(2, q)$, but nevertheless the image of the word map $w = x^a y^b$ does not contain any unipotent. For example, the word $w = x^{42} y^{42}$ is *not* surjective on $\mathrm{PSL}(2, 7)$ and $\mathrm{PSL}(2, 8)$.

In addition, it was determined when $-\,\mathrm{id}$ can be written as $x^a y^b$ for some $x, y \in \mathrm{SL}(2, q)$. In particular, if $q \equiv \pm 3 \bmod 8$, then $x^4 y^4 \neq -\,\mathrm{id}$ for every $x, y \in \mathrm{SL}(2, q)$ (the same result was obtained independently in [74]).

These results demonstrate, in particular, the difference between word maps in simple and quasisimple groups (see also the previous discussion in subsections 5.2.1, 5.2.2 and 5.2.3).

6.2. Criteria for equidistribution

In this section we describe some results on equidistribution of solutions of word equations of the form $w(x, y) = g$ in the family of finite groups $\mathrm{SL}(2, q)$ which were obtained in [14]. A criterion for equidistribution in terms of the trace polynomial of w is given in Theorem 6.18 below. This allows one to get an explicit description of certain classes of words possessing the equidistribution property and show that this property is generic within these classes. This result can be viewed, on the one hand, as a refinement (in the SL_2-case) of equidistribution theorems of [69] and [73] on general words w and general Chevalley groups G, and, on the other hand, as a generalization of equidistribution theorems for some particular words: [41] (commutator words on any finite simple G), [9] (Engel words on SL_2), [8] (words of the form $w = x^a y^b$ on SL_2). Acting in the spirit of [41], we deduce a criterion for $w \colon \mathrm{SL}_2 \times \mathrm{SL}_2 \to \mathrm{SL}_2$ to be *almost measure-preserving*. It turns out that "good" (equidistributed, measure-preserving) words are essentially those whose trace polynomial cannot be represented as a composition of two other polynomials.

Here are precise definitions and results. We will follow the approach to equidistribution adopted in [41] (see subsection 5.3.1 above).

Definition 6.7 (cf. [41, § 3] and subsection 5.3.1).
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) $\left| \# f^{-1}(y) - \# X / \# Y \right| < \varepsilon\, \# X / \# Y$ for all $y \in Y'$.

The 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 6.8 ([14]).
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 $\varepsilon_p \colon \mathbb{N} \to \mathbb{R}$ 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) $\left| \# P_y - \# X_q / \# Y_q \right| < \varepsilon_p(q)\, \# X_q / \# Y_q$ for all $y \in Y_q \setminus S_q$.

Remark 6.9.
Definition 6.8 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 6.7.

Definition 6.10 ([14]).

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 $Y_q = G_q = \mathrm{SL}(2, q)$, $X_q = (G_q)^2$ is a direct product of its two copies, and $P_q = P_{w,q} \colon (G_q)^2 \to G_q$ is the morphism induced by some fixed word $w \in \mathcal{F}_2$. Accordingly, we say that w is equidistributed (or p-equidistributed) if so is the family of maps $P_{w,q} \colon \mathrm{SL}(2, q) \times \mathrm{SL}(2, q) \to \mathrm{SL}(2, q)$ (or, in other words, if so is 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 $\mathbb{Z}$).

Recall some properties of polynomials.

Definition 6.11.

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 6.12 ([78, 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 $a \ne 0$ and k is a non-negative integer.

The following notions are essential for our criteria.

Definition 6.13.

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

Note that if E/F is a separable field extension, it is known [4, Theorem 1 and Proposition 1] that P is F-composite if and only if P is E-composite. In particular, working over perfect ground fields, we may always assume, if needed, that F is algebraically closed.

Definition 6.14.

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

- We say that P is *p-composite* if the reduced polynomial $P_p \in \mathbb{F}_p[x_1, \ldots, x_n]$ 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 6.13, $P_p = h \circ Q$ where $h \in \mathbb{F}_p[x]$ is a permutation polynomial of all finite extensions of $\mathbb{F}_p$.

Definition 6.15.

We say that a polynomial $P \in \mathbb{Z}[x_1, \ldots, x_n]$ 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 6.16.

If a polynomial $P \in \mathbb{Z}[x_1, \ldots, x_n]$ is $\mathbb{Q}$-noncomposite, it is p-noncomposite for all but finitely many primes p [16, 2.2.1]. If $P \in \mathbb{Z}[x_1, \ldots, x_n]$ is $\mathbb{Q}$-composite, it is very composite.

We can now formulate the main results of [14].

Theorem 6.17.

Let $w \in \mathcal{F}_2$. 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 6.18.
Let $w \in \mathcal{F}_2$. 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 6.19.
Suppose that for each p and all n big enough the image of $P_{w,p^n} \colon \mathrm{SL}(2,p^n) \times \mathrm{SL}(2,p^n) \to \mathrm{SL}(2,p^n)$ contains all noncentral semisimple elements of $\mathrm{SL}(2,p^n)$. Then w is equidistributed.

For a given word $w \in \mathcal{F}_2$, let us now consider the family of groups $\widehat{G}_q = \mathrm{PSL}(2,q)$ and the corresponding word maps $w_q \colon \widehat{G}_q \times \widehat{G}_q \to \widehat{G}_q$.

Proposition 6.20.
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 of maps $w_q \colon \widehat{G}_q \times \widehat{G}_q \to \widehat{G}_q$.

Here are the main ingredients of the proofs of Theorems 6.17 and 6.18:

- The diagram

$$
\begin{array}{ccc}
G_q \times G_q & \xrightarrow{\;w\;} & G_q \\
\downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \mathrm{tr}} \\
\mathbb{A}^3_{s,u,t}(\mathbb{F}_q) & \xrightarrow{\;f_w\;} & \mathbb{A}^1_z(\mathbb{F}_q)
\end{array}
\tag{20}
$$

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

- An explicit Lang–Weil estimate (Ghorpade–Lachaud [46]): if $H \subset \mathbb{A}^3_{\mathbb{F}_q}$ is an *absolutely irreducible* hypersurface of degree d, then

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

or, equivalently, $\# H(\mathbb{F}_q) = q^2(1+r_1)$ with

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

For $q > Cd^8$ this gives $|r_1| < 1/2$.

- A generalized Stein–Lorenzini inequality [96]: if $f_{w,p}$ is *p-noncomposite*, then the *spectrum* $\sigma(f_{w,p})$, 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, contains at most $d - 1$ points, where $d = \deg f_w$. The same is true for each $\sigma_q(f_w) = \sigma(f_{w,p}) \cap \mathbb{F}_q$. 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 satisfies the Ghorpade–Lachaud inequality.

- Estimates for fibres of the trace map:

Lemma 6.21.
Let $D(s,u,t) = (s^2 - 4)(t^2 - 4)(s^2 + t^2 + u^2 - ust - 4)$, and let $\Delta \subset \mathbb{A}^3_{s,u,t}$ be defined by the equation $D = 0$. Let $H \subset \mathbb{A}^3_{s,u,t}(\overline{\mathbb{F}}_p)$ be a hypersurface of degree d such that $H \not\subset \Delta$. Then for π from diagram (20) we have $\# \pi^{-1}(H)(\mathbb{F}_q) = \# H(\mathbb{F}_q) q^3(1 + r_2)$, where $|r_2| < Cd/q$.

- Estimates for the size of the value set of polynomials [125]: if R 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 $\mathbb{F}_q$ and omits at least $(q^m - 1)/d_1$ values of $\mathbb{F}_{q^m}$, where $d_1 = \deg R$.

6.2.1. Composite trace polynomials

The goal of this section is to describe words in two variables whose trace polynomial is composite. We send the curious reader to [14] for proofs.

Throughout this section $T_n(x)$ stands for the nth Chebyshev polynomial, and $D_n(x) = 2T_n(x/2)$ for the nth Dickson polynomial. It is well known (see, e.g., [77, (2.2)]) that this polynomial satisfies $D_n(x + 1/x) = x^n + 1/x^n$ and is completely determined by this functional equation. We always assume that $w(x, y)$ is written in the form

$$w = x^{a_1} y^{b_1} \cdots x^{a_r} y^{b_r} \tag{21}$$

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

Definition 6.22.
We say that two reduced words $w = x^{a_1} y^{b_1} \ldots x^{a_r} y^{b_r}$ and $v = x^{c_1} y^{d_1} \ldots x^{c_{r'}} y^{d_{r'}}$, written in form (21), are *trace-similar* if $r = r'$, the array $\{|a_i|\}$ is a rearrangement of $\{|c_i|\}$, and the array $\{|b_i|\}$ is a rearrangement of $\{|d_i|\}$.

Note that if reduced words $w = x^{a_1} y^{b_1} \cdots x^{a_r} y^{b_r}$ and $v = x^{c_1} y^{d_1} \cdots x^{c_{r'}} y^{d_{r'}}$, written in form (21), have the same trace polynomial, then w and v are trace-similar [56].

The following propositions are valid (see [14] for the proofs).

Proposition 6.23.
Let $w(x, y) = x^{a_1} y^{b_1} \cdots x^{a_r} y^{b_r}$, $A = \sum a_i$, $B = \sum b_i$. Assume that either $A \neq 0$ or $B \neq 0$. Assume that the trace polynomial $f_w(s, u, t)$ is $\mathbb{C}$-composite, $f_w(s, u, t) = h(q(s, u, t))$, where $q \in \mathbb{C}[s, u, t]$ and $h \in \mathbb{C}[z]$, $\deg h \geq 2$. Then $h = D_d(z)$ for some $d \geq 2$.

Proposition 6.24.
Let w be a reduced word of complexity r written in form (21). If its trace polynomial f_w is $\mathbb{C}$-composite, $f_w(s, u, t) = h(q(s, u, t))$ where $q \in \mathbb{C}[s, u, t]$ and $h(x) = \mu x^n + \cdots$ is a polynomial in one variable of degree n, then $r = nm$ and w is trace-similar to $v(x, y)^n$ where v is a word of complexity m.

Proposition 6.25.
Let $w(x, y) = x^a y^b \cdots$ be a reduced word of complexity n such that $f_w(s, u, t) = D_n(q(s, u, t))$ for some q. Then $w(x, y) = (x^a y^b)^n$.

Remark 6.26.
The statements of Propositions 6.23, 6.24 and 6.25 remain valid if we replace $\mathbb{C}$ by (the algebraic closure of) a sufficiently big prime field $\mathbb{F}_p$, and "composite" by "p-composite" ($p > p_0$ depending on w).

Here are some concrete cases where one can get more conclusive results.

Corollary 6.27.
Let $w(x, y) = x^a y^b \cdots$ be a reduced word of prime complexity r. If $p > r$ and w is not p-equidistributed, then $w = v^r(x, y)$.

Corollary 6.28.
The word $w(x, y) = x^a y^b x^c y^d$ is either equidistributed or equal to $(x^a y^b)^2$.

All facts mentioned above allow one to describe a class of words within which a "generic" word induces the map which is almost equidistributed. More precisely, we have the following proposition.

Proposition 6.29.
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 word 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 $\mathcal{R}$.

(According to the terminology of [65], this means that the proportion of words from S among all words from $\mathcal{R}$ of fixed length tends to 1 exponentially fast as the length tends to infinity.) This is proved by combining the results quoted above with the well-known fact (see, e.g., [5]) stating that the class of words which are proper powers of other words is exponentially negligible.

7. Concluding remarks and open problems

We conclude with a brief discussion of various ramifications, analogues and generalizations of results presented in this survey, focusing on open problems, the list of which does not pretend to be comprehensive and reflects the authors' taste. Most of them are borrowed from [7, 9, 10, 14, 52].

Engel-like words and solvability properties

In light of Theorems 3.4 and 3.6 and subsequent discussions in subsection 4.1.4, it is natural to ask whether or not the property of a given sequence to characterize finite solvable groups is generic. A possible way to express it is the following conjecture.

Conjecture 7.1.
Let $\mathcal{R}_0$ (resp. $\mathcal{R}$) denote the class of words $f(x, y, z) \in \mathcal{F}_3$ satisfying the following condition: there exists $w_0(x, y) \in \mathcal{F}_2$ (resp. w) such that the sequence $\{v_n^{(0)}(x, y)\}$ (resp. $\{v_n(x, y)\}$) generated by the first word w_0 (resp. w) and law f

  (i) *does not contain the identity word,*

 (ii) *is Engel-like,*

(iii) *descends along the derived series*

(*resp. conditions* (i)–(iii) *and the additional one:*

 (iv) *for any finite group G the following holds: G is solvable if and only if there is n such that $v_n(x, y) \equiv 1$ in G).*

Then the class $\mathcal{R}$ is generic within $\mathcal{R}_0$ (in the sense of [65], as in subsection 6.2).

We hope that algebraic–geometric approaches developed in subsection 4.1.4 could be useful in establishing this conjecture. One can also try the following counter-part as a testing ground: in Conjecture 7.1, replace throughout "solvable" with nilpotent and "derived series" with "lower central series".

It would be interesting to investigate further, in the spirit of [7, 51], the relationship between finite groups and finite-dimensional Lie algebras from the point of view of solvability properties. Namely, one can put forward the following (maybe over-optimistic) conjecture.

Conjecture 7.2.
Let $\{v_n(x, y)\}$ be an Engel-like sequence of words in the free Lie algebra $\mathcal{W}_2$ which characterizes finite-dimensional solvable Lie algebras defined over fields of arbitrary characteristics (i.e., a finite-dimensional Lie algebra $\mathfrak{g}$ defined over an arbitrary field K is solvable if and only if there is n such that $v_n(x, y) \equiv 0$ in $\mathfrak{g}$). Then the same sequence, regarded as a sequence of words in the free group $\mathcal{F}_2$ (i.e., viewing Lie bracket as commutator), characterizes finite solvable groups.

Note that the assumption on arbitrary characteristics is essential: in [7] there have been exhibited sequences characterizing finite-dimensional Lie algebras defined over fields of characteristic zero but not over fields of prime characteristic (and not characterizing finite solvable groups). The first intriguing example is the sequence defined by

$$v_1(x, y) = [x, y], \qquad v_{n+1}(v, y) = \big[[v_n(x, y), x], [v_n(x, y), y]\big].$$

This example resists the algebraic–geometric approach described in Section 3 for purely computational reasons: the arising equations lead to varieties which are out of range of SINGULAR and MAGMA. Perhaps experts in computer algebra, who are able to apply more sophisticated methods, will be more lucky.

It is desirable to use an Engel-like sequence for characterizing the solvable radical of a finite group in the same way as the original Engel sequence is used to characterize the nilpotent radical. A theorem of Baer [6] says that the nilpotent radical of a finite group G coincides with the set of elements $y \in G$ with the following property: for every $x \in G$ there is $n = n(x, y)$ such that $e_n(x, y) = 1$.

Problem 7.3.
Exhibit an Engel-like sequence $\{v_n(x, y)\}$ such that the solvable radical of any finite group G coincides with the collection of $y \in G$ with the property: for every $x \in G$ there is $n = n(x, y)$ such that $v_n(x, y) = 1$.

A recent result of Wilson [126], stating the existence of a two-variable countable set of words with the required property, gives a strong evidence that such an Engel-like sequence should exist but does not provide any candidate. Note that even the toy problem of characterizing the solvable radical of a finite-dimensional Lie algebra with the help of Engel-like sequences remains open in the case where the ground field is of positive characteristic; see [7, 52]. We hope that algebraic–geometric machinery in the spirit of Section 3 may turn out to be useful for achieving this goal.

Borel's theorem and around

Let G be a connected semisimple algebraic group defined over an infinite field k.

Problem 7.4.
What can be said about the "fine structure" of $w(G)$? In particular, describe w such that $w(G)$ contains the set of all semisimple elements of G.

A similar problem was discussed in [62] for associative noncommutative polynomials on associative matrix algebras, in the spirit of Kaplansky's problem. In such a setting, the resulting maps may well be non-dominant, and not only for obvious reasons mentioned in the introduction. In the paper cited above there have been described certain classes of polynomials which are not central and whose image contains elements with nonzero trace but the induced map is not dominant on 2×2 matrices. Here are some general questions remaining open.

Question 7.5 ([62]).
Let P be an associative, noncommutative, noncentral, *multilinear* polynomial in d variables whose image contains matrices with nonzero trace. Does P induce a dominant map $M(n, K)^d \to M(n, K)$? Does there exist P such that this map is not surjective?

Similar problems were discussed in [10] for Lie polynomials P on Chevalley Lie algebras $\mathfrak{g}$. It was shown that the induced map is dominant provided P is not identically zero on $\mathfrak{sl}(2)$. We do not know whether or not the latter assumption can be dropped.

Question 7.6.
Does there exist a Lie polynomial in d variables, not identically zero on a Chevalley Lie algebra $\mathfrak{g}$, such that the induced map $\mathfrak{g}^d \to \mathfrak{g}$ is not dominant?

It would be interesting to understand the situation with *infinite–dimensional* simple Lie algebras (as well as with finite–dimensional algebras of Cartan type over fields of positive characteristic). The first question, which does not seem too complicated, is the following one.

Question 7.7.
For which Lie algebras $\mathfrak{g}$ of Cartan type the map $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, $(x, y) \mapsto [x, y]$, is surjective?

If for some $\mathfrak{g}$ this question is answered in the affirmative, one can continue by looking at the Engel maps, as in [10].

Remark 7.8.
It would be interesting to consider a more general set-up when we have a polynomial map $P \colon L^d \to L^s$. (This means that we consider systems of equations rather than single equations.) In [48] some dominance results were obtained for the multiple commutator map $P \colon L \times L^d \to L^d$ given by the formula $P(X, X_1, \ldots, X_d) = ([X, X_1], \ldots, [X, X_d])$.

Remark 7.9.
In a similar spirit, one can consider generalized word maps $w \colon G^d \to G^s$ on simple groups. Apart from [48], see also a discussion of a particular case $w = (w_1, w_2) \colon G^2 \to G^2$ in [21, Problem 1].

Remark 7.10.
It would be interesting to find more classes of infinite simple (or close to simple) groups admitting some analogue of Borel's statement in the sense that the image of the word map is "large", at least for a generic word. Some such classes were discussed in the literature: infinite symmetric groups [86] (surjectivity of the commutator word was established by Ore [102]), groups of automorphisms of trees [92] and of random graphs [29].

We would suggest looking at Cremona groups, which share many common properties with linear algebraic groups and in which such notions as Zariski topology and dominance can be defined (see, e.g., [112]).

Problem 7.11.
Let $w \in \mathcal{F}_d$ be a nontrivial word. Is the corresponding word map $w \colon (\mathrm{Cr}(n, k))^d \to \mathrm{Cr}(n, k)$ dominant?

Of course, the first case to be considered is $n = 2$. Recall two recent spectacular results on $\mathrm{Cr}(2, k)$ answering long–standing questions on its simplicity. It turns out that this group is at the same time simple and non–simple: it does not contain nontrivial normal subgroups closed in the topology mentioned above [15] but contains lots of abstract normal subgroups [22].

Remark 7.12.
In light of the previous remark, one can note that analogues of Borel's theorem may look differently at the first glance when one considers other classes of groups. Say, a certain analogue of Borel's theorem for profinite groups is provided by the following deep theorem by Nikolov and Segal [100, 101]: let G be a finitely generated profinite group, let w be a non–commutator word, and let $\langle w(G) \rangle$ denote the corresponding verbal subgroup (i.e., the subgroup generated by the set $\{w(g)^{\pm 1}\}$, $g \in G$); then $\langle w(G) \rangle$ is open in G. As pointed out to us by the first referee, this is a stronger version of the positive solution to the restricted Burnside Problem [127] (and of course the proof depends on the latter).

Remark 7.13.
To prevent the reader from an overoptimistic view on Borel's theorem, one has to note that there may be a significant gap between dominance and surjectivity. Moreover, the image of the word map may be very large in the Zariski topology (according to Borel's theorem) but very small in some natural topology. See [118] where such word maps are constructed for real compact Lie groups. Note that the behaviour of Engel word maps on such groups is much better: they are all surjective [30].

Word maps on finite simple groups

Regarding the image of the word map, we can recall here Thompson's and Shalev's Conjectures 5.14, 5.5 and 5.18 discussed in subsection 5.2.2. In particular, one can ask whether the following variant of Shalev's Conjecture 5.5, for the family of groups $PSL(2, q)$, holds.

Conjecture 7.14 (Shalev[1]).
Assume that $w = w(x, y) \in \mathcal{F}_2$ is not of the form $v^m(x, y)$ for some $v = v(x, y) \in \mathcal{F}_2$ and $m > 1$. Then there exists a constant $q_0(w)$ such that if $q > q_0(w)$ then $w(G) = G$ for $G = PSL(2, q)$.

In particular, this conjecture holds for Engel words and for words of the form $x^a y^b$ (see subsection 6.1 and [8, 9]). One can therefore attempt to use the trace map method described in subsection 2.3 to find more words satisfying the above conjecture. In particular, the following questions are raised (see subsection 4.1.4 and [9, Section 8]).

Question 7.15 ([9]).
What are the words $w = w(x, y) \in \mathcal{F}_2$ for which the corresponding trace map $\psi(s, u, t) = (f_1(s, u, t), f_2(s, u, t), t)$ has one of the following properties:

  ($\star$) the set $\{f_1(s, u, t) = a\}$ is absolutely irreducible for almost all q and for every $a \in \mathbb{F}_q$?

  ($\star\star$) there exists a ψ-invariant plane A and the curves $\{\psi\!\restriction_A = a\}$ are absolutely irreducible for for almost all q and for a general $a \in \mathbb{F}_q$?

It is tempting to generalize the results on the image of some word maps of [8, 9], presented in Section 6.1, as well as the criteria for almost equidistribution of [14], presented in Section 6.2, 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 almost 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 6.19. We dare to formulate the following conjecture.

Conjecture 7.16.
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 $w \in \mathcal{F}_d$, $d \geq 2$, 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 7.16 (cf. the discussion in [73] 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.

---

[1] *After the first version of the present survey had been submitted, a counter-example to this conjecture was constructed in [60].*

Other interesting problems, arising in the context of measure preservation and primitivity, were raised by Puder and Parzanchevski [105]. Namely, in [105] it was shown that the property of a word w to be measure-preserving within the class of all finite groups can be detected on the family of all symmetric groups S_n. Are there other natural families (say, $PGL(n, q)$) that can be used as such detectors? They also ask whether their results on measure preservation can be extended to the class of compact Lie groups (with respect to the Haar measure).

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

Miscellaneous remarks and problems

Remark 7.17.
Recently, Shalev with his collaborators extended many results of Waring type from finite simple groups to some simple algebraic groups over p-adic integers. The case of simple Chevalley groups over rings of integers remains widely open, see [115] for some relevant questions and conjectures.

Remark 7.18.
A new dynamical viewpoint on the image of the word map was developed by Schul and Shalev [110]. They showed that the random walk on any finite simple group G, with respect to this image as a generating set, has mixing time 2.

Remark 7.19.
One could try to extend some of the results of this survey to the case of matrix groups or algebras over some sufficiently good ring. One has to be careful: say, in [108] there are examples of rings R such that not every element of $\mathfrak{sl}(n, R)$ is a commutator.

Remark 7.20.
One can ask questions of Borel's type for other classes of algebras (beyond groups, Lie algebras and associative algebras). The interested reader may refer to [49] for the case of values of commutators and associators on alternative and Jordan algebras.

Remark 7.21.
It would be interesting to extend (at least part of) the methods and results described in the present survey to matrix equations with constant matrix coefficients, as discussed in the introduction.

Acknowledgements

References

[1] Abért M., On the probability of satisfying a word in a group, J. Group Theory, 2006, 9(5), 685–694

[2] Adolphson A., Sperber S., On the degree of the L-functions associated with an exponential sum, Compositio Math., 1988, 68(2), 125–159

[3] Amitsur S.A., The T-ideals of the free ring, J. London Math. Soc., 1955, 30(4), 470–475

[4] Arzhantsev I.V., Petravchuk A.P., Closed polynomials and saturated subalgebras of polynomial algebras, Ukrainian Math. J., 2007, 59(12), 1783–1790

[5] Arzhantseva G.N., Ol'shanskii A.Yu., Generality of the class of groups in which subgroups with a lesser number of generators are free, Math. Notes, 1996, 59(3-4), 350–355

[6] Baer R., Engelsche Elemente Noetherscher Gruppen, Math. Ann., 1957, 133(3), 256–270

[7] Bandman T., Borovoi M., Grunewald F., Kunyavskiĭ B., Plotkin E., Engel-like characterization of radicals in finite dimensional Lie algebras and finite groups, Manuscripta Math., 2006, 119(4), 465–481

[8] Bandman T., Garion S., Surjectivity and equidistribution of the word $x^a y^b$ on PSL(2, q) and SL(2, q), Internat. J. Algebra Comput., 2012, 22(2), #1250017

[9] Bandman T., Garion S., Grunewald F., On the surjectivity of Engel words on PSL(2, q), Groups Geom. Dyn., 2012, 6(3), 409–439

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

[11] Bandman T., Greuel G.-M., Grunewald F., Kunyavskiĭ B., Pfister G., Plotkin E., Two-variable identities for finite solvable groups, C. R. Acad. Sci. Paris, 2003, 337(9), 581–586

[12] Bandman T., Greuel G.-M., Grunewald F., Kunyavskiĭ B., Pfister G., Plotkin E., Identities for finite solvable groups and equations in finite simple groups, Compos. Math., 2006, 142(3), 734–764

[13] Bandman T., Grunewald F., Kunyavskiĭ B., Geometry and arithmetic of verbal dynamical systems on simple groups, Groups Geom. Dyn., 2010, 4(4), 607–655

[14] Bandman T., Kunyavskiĭ B., Criteria for equidistribution of solutions of word equations in SL(2), J. Algebra, 2013, 382, 282–302

[15] Blanc J., Groupes de Cremona, connexité et simplicité, Ann. Sci. Éc. Norm. Supér., 2010, 43(2), 357–364

[16] Bodin A., Dèbes P., Najib S., Indecomposable polynomials and their spectrum, Acta Arith., 2009, 139(1), 79–100

[17] Borel A., On free subgroups of semisimple groups, Enseign. Math., 1983, 29(1-2), 151–164

[18] Borisov A., Sapir M., Polynomial maps over finite fields and residual finiteness of mapping tori of group endomorphisms, Invent. Math., 2005, 160(2), 341–356

[19] Borisov A., Sapir M., Polynomial maps over p-adics and redisual properties of mapping tori of group endomorphisms, Int. Math. Res. Not. IMRN, 2009, 16, 3002–3015

[20] Bray J.N., Wilson J.S., Wilson R.A., A characterization of finite soluble groups by laws in two variables, Bull. London Math. Soc., 2005, 37(2), 179–186

[21] Breuillard E., Green B., Guralnick R., Tao T., Strongly dense free subgroups of semisimple algebraic groups, Israel J. Math., 2012, 192(1), 347–379

[22] Cantat S., Lamy S., Normal subgroups in the Cremona group, Acta Math., 2013, 210(1), 31–94

[23] Cargo D.P., de Launey W., Liebeck M.W., Stafford R.M., Short two-variable identities for finite groups, J. Group Theory, 2008, 11(5), 675–690

[24] Casals-Ruiz M., Kazachkov I., On Systems of Equations over Free Partially Commutative Groups, Mem. Amer. Math. Soc., 212(999), American Mathematical Society, Providence, 2011

[25] Connes A., Schwarz A., Matrix Vieta theorem revisited, Lett. Math. Phys., 1997, 39(4), 349–353

[26] Deligne P., Sullivan D., Division algebras and the Hausdorff–Banach–Tarski paradox, Enseign. Math., 1983, 29(1–2), 145–150

[27] Digne F., Michel J., Representations of Finite Groups of Lie Type, London Math. Soc. Stud. Texts, 21, Cambridge University Press, Cambridge, 1991

[28] Dixon J.D., The probability of generating the symmetric group, Math. Z., 1969, 110(3), 199–205

[29] Droste M., Truss J.K., On representing words in the automorphism group of the random graph, J. Group Theory, 2006, 9(6), 815–836

[30] Elkasapy A., Thom A., About Goto's method showing surjectivity of word maps, preprint available at http://arxiv.org/abs/1207.5596

[31] Ellers E.W., Gordeev N., Gauss decomposition with prescribed semisimple part in classical Chevalley groups, Comm. Algebra, 1994, 22(14), 5935–5950

[32] Ellers E.W., Gordeev N., Gauss decomposition with prescribed semisimple part in Chevalley groups II, Exceptional cases, Comm. Algebra, 1995, 23(8), 3085–3098

[33] Ellers E.W., Gordeev N., Gauss decomposition with prescribed semisimple part in Chevalley groups III, Finite

twisted groups, Comm. Algebra, 1996, 24(14), 4447–4475

[34] Ellers E.W., Gordeev N., On the conjectures of J. Thompson and O. Ore, Trans. Amer. Math. Soc., 1998, 350(9), 3657–3671

[35] Etingof P., Gelfand I., Retakh V., Factorization of differential operators, quasideterminants, and nonabelian Toda field equations, Math. Res. Lett., 1997, 4(2-3), 413–425

[36] Formanek E., Central polynomials for matrix rings, J. Algebra, 1972, 23(1), 129–132

[37] Fricke R., Über die Theorie der automorphen Modulgruppen, Nachr. Akad. Wiss. Göttingen, 1896, 91–101

[38] Fricke R., Klein F., Vorlesungen über die Theorie der Automorphen Funktionen, 1 and 2, Teubner, Leipzig, 1897 and 1912

[39] Fuchs D., Schwarz A., Matrix Vieta theorem, In: Lie groups and Lie algebras: E.B. Dynkin's Seminar, Amer. Math. Soc. Transl. Ser. 2, 169, American Mathematical Society, Providence, 1995, 15–22

[40] Fujiwara K., Rigid geometry, Lefschetz–Verdier trace formula and Deligne's conjecture, Invent. Math., 1997, 127(3), 489–533

[41] Garion S., Shalev A., Commutator maps, measure preservation, and T-systems, Trans. Amer. Math. Soc., 2009, 361(9), 4631–4651

[42] Gelfand I., Retakh V., Noncommutative Vieta theorem and symmetric functions, In: The Gelfand Mathematical Seminars, 1993–1995, Gelfand Math. Sem., Birkhäuser, Boston, 1996, 93–100

[43] Gelfand I., Retakh V., Quasideterminants I, Selecta Math. (N.S.), 1997, 3(4), 517–546

[44] Gelfand S., On the number of solutions of a quadratic equation, In: Globus: General Mathematical Seminar, 1, Independent University of Moscow, Moscow, 2004, 124–133 (in Russian)

[45] Ghorpade S.R., Lachaud G., Number of solutions of equations over finite fields and a conjecture of Lang and Weil, In: Number Theory and Discrete Mathematics, Chandigarh, October 2–6, 2000, Trends Math., Birkhäuser, Basel, 2002, 269–291

[46] Ghorpade S.R., Lachaud G., Étale cohomology, Lefschetz theorems and number of points of singular varieties over finite fields, Mosc. Math. J., 2002, 2(3), 589–631; 2009, 9(2), 431–438

[47] Goldman W.M., An exposition of results of Fricke and Vogt, preprint available at http://arxiv.org/abs/math/0402103

[48] Gordeev N., Rehmann U., On multicommutators for simple algebraic groups, J. Algebra, 2001, 245(1), 275–296

[49] Gordon S.R., Associators in simple algebras, Pacific J. Math., 1974, 51(1), 131–141

[50] Gowers W.T., Quasirandom groups, Combin. Probab. Comput., 2008, 17(3), 363–387

[51] Grunewald F., Kunyavskiĭ B., Nikolova D., Plotkin E., Two-variable identities in groups and Lie algebras, J. Math. Sci. (N.Y.), 2003, 116(1), 2972–2981

[52] Grunewald F., Kunyavskiĭ B., Plotkin E., Characterization of solvable groups and solvable radical, Internat. J. Algebra Comput., 2013, 23(5), 1011–1062

[53] Guralnick R., Malle G., Products of conjugacy classes and fixed point spaces, J. Amer. Math. Soc., 2012, 25(1), 77–121

[54] Guralnick R.M., Tiep P.H., Cross characteristic representations of even characteristic symplectic groups, Trans. Amer. Math. Soc., 2004, 356(12), 4969–5023

[55] Guralnick R.M., Tiep P.H., The Waring problem for finite quasisimple groups. II, preprint available at http://arxiv.org/abs/1302.0333

[56] Horowitz R.D., Characters of free groups represented in the two-dimensional special linear group, Comm. Pure Appl. Math., 1972, 25(6), 635–649

[57] Hrushovski E., The elementary theory of the Frobenius automorphisms, preprint available at http://arxiv.org/abs/math.LO/0406514/

[58] Humphreys J.E., Modular Representations of Finite Groups of Lie Type, London Math. Soc. Lecture Note Ser., 326, Cambridge University Press, Cambridge, 2006

[59] Huppert B., Blackburn N., Finite Groups, III, Grundlehren Math. Wiss., 243, Springer, Berlin–Heidelberg–New York, 1982

[60] Jambor S., Liebeck M.W., O'Brien E.A., Some word maps that are non-surjective on infinitely many finite simple groups, Bull. Lond. Math. Soc., 2013, 45(5), 907–910

[61] Kanel-Belov A., Kunyavskiĭ B., Plotkin E., Word equations in simple groups and polynomial equations in simple algebras, Vestnik St. Petersburg Univ. Math., 2013, 46(1), 3–13

[62] Kanel-Belov A., Malev S., Rowen L., The images of non-commutative polynomials evaluated on 2×2 matrices,

Proc. Amer. Math. Soc., 2012, 140(2), 465–478

[63] Kantor W.M., Lubotzky A., The probability of generating a finite classical group, Geom. Dedicata, 1990, 36(1), 67–87

[64] Kapovich I., Mapping tori of endomorphisms of free groups, Comm. Algebra, 2000, 28(6), 2895–2917

[65] Kapovich I., Schupp P.E., Random quotients of the modular group are rigid and essentially incompressible, J. Reine Angew. Math., 2009, 628, 91–119

[66] Kassabov M., Nikolov N., Words with few values in finite simple groups, Quart. J. Math. (in press), DOI: 10.1093/qmath/has018

[67] Lang S., Weil A., Number of points of varieties in finite fields, Amer. J. Math., 1954, 76(4), 819–827

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

[69] Larsen M.J., Pink R., Finite subgroups of algebraic groups, J. Amer. Math. Soc., 2011, 24(4), 1105–1158

[70] Larsen M., Shalev A., Characters of symmetric groups: sharp bounds and applications, Invent. Math., 2008, 174(3), 645–687

[71] Larsen M., Shalev A., Word maps and Waring type problems, J. Amer. Math. Soc., 2009, 22(2), 437–466

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

[73] Larsen M., Shalev A., Tiep P.H., The Waring problem for finite simple groups, Ann. of Math., 2011, 174(3), 1885–1950

[74] Larsen M., Shalev A., Tiep P.H., Waring problem for finite quasisimple groups, Int. Math. Res. Not. (IMRN), 2013, 10, 2323–2348

[75] Levy M., Word maps with small image in simple groups, preprint available at http://arxiv.org/abs/1206.1206

[76] Levy M., Word maps with small image in almost simple groups and quasisimple groups, preprint available at http://arxiv.org/abs/1301.7188

[77] Lidl R., Mullen G.L., Turnwald G., Dickson Polynomials, Pitman Monogr. Surveys Pure Appl. Math., 65, Longman Scientific & Technical, Harlow, 1993

[78] Lidl R., Niederreiter H., Finite Fields, Encyclopedia Math. Appl., 20, Addison-Wesley, Reading, 1983

[79] Liebeck M.W., O'Brien E.A., Shalev A., Tiep P.H., The Ore conjecture, J. Eur. Math. Soc. (JEMS), 2010, 12(4), 939–1008

[80] Liebeck M.W., O'Brien E.A., Shalev A., Tiep P.H., Commutators in finite quasisimple groups, Bull. Lond. Math. Soc., 2011, 43(6), 1079–1092

[81] Liebeck M.W., O'Brien E.A., Shalev A., Tiep P.H., Products of squares in finite simple groups, Proc. Amer. Math. Soc., 2012, 140(1), 21–33

[82] Liebeck M.W., Shalev A., The probability of generating a finite simple group, Geom. Dedicata, 1995, 56(1), 103–113

[83] Liebeck M.W., Shalev A., Diameters of finite simple groups: sharp bounds and applications, Ann. of Math., 2001, 154(2), 383–406

[84] Liebeck M.W., Shalev A., Fuchsian groups, finite simple groups, and representation varieties, Invent. Math., 2005, 159(2), 317–367

[85] Lubotzky A., Images of word maps in finite simple groups, Glasg. Math. J. (in press), DOI: 10.1017/S0017089513000396

[86] Lyndon R.C., Words and infinite permutations, In: Mots, Lang. Raison. Calc., Hermès, Paris, 1990, 143–152

[87] Macbeath A.M., Generators of the linear fractional groups, In: Number Theory, Houston, 1967, American Mathematical Society, Providence, 1969, 14–32

[88] Macpherson D., Tent K., Pseudofinite groups with NIP theory and definability in finite simple groups, In: Groups and Model Theory, Mülheim an der Ruhr, May 30–June 3, 2011, Contemp. Math., 576, American Mathematical Society, Providence, 2012, 255–267

[89] Magnus W., Rings of Fricke characters and automorphisms groups of free groups, Math. Z., 1980, 170(1), 91–102

[90] Magnus W., The uses of 2 by 2 matrices in combinatorial group theory. A survey, Resultate Math., 1981, 4(2), 171–192

[91] Manin Yu.I., Cubic Forms, North-Holland Math. Library, 4, North-Holland, Amsterdam, 1986

[92] Maroli J.A., Representation of tree permutations by words, Proc. Amer. Math. Soc., 1990, 110(4), 859–869

[93] Martinez C., Zelmanov E., Products of powers in finite simple groups, Israel J. Math., 1996, 96(2), 469–479

[94] Myasnikov A., Nikolaev A., Verbal subgroups of hyperbolic groups have infinite width, preprint available at http://arxiv.org/abs/1107.3719

[95] Myasnikov A.G., Shpilrain V., Automorphic orbits in free groups, J. Algebra, 2003, 269(1), 18–27

[96] Najib S., Une généralisation de l'inégalité de Stein–Lorenzini, J. Algebra, 2005, 292(2), 566–573

[97] Nikolov N., Algebraic properties of profinite groups, preprint available at http://arxiv.org/abs/1108.5130

[98] Nikolov N., Pyber L., Product decompositions of quasirandom groups and a Jordan type theorem, J. Eur. Math. Soc. (JEMS), 2011, 13(4), 1063–1077

[99] Nikolov N., Segal D., A characterization of finite soluble groups, Bull. Lond. Math. Soc., 2007, 39(2), 209–213

[100] Nikolov N., Segal D., Powers in finite groups, Groups Geom. Dyn., 2011, 5(2), 501–507

[101] Nikolov N., Segal D., Generators and commutators in finite groups; abstract quotients of compact groups, Invent. Math., 2012, 190(3), 513–602

[102] Ore O., Some remarks on commutators, Proc. Amer. Math. Soc., 1951, 2(2), 307–314

[103] Platonov V.P., Linear groups with identical relations, Dokl. Akad. Nauk BSSR, 1967, 11, 581–582 (in Russian)

[104] Puder D., Primitive words, free factors and measure preservation, Israel J. Math., 2014 (in press), DOI: 10.1007/s11856-013-0055-2

[105] Puder D., Parzanchevski O., Measure preserving words are primitive, preprint available at http://arxiv.org/abs/1202.3269

[106] Razmyslov Yu.P., A certain problem of Kaplansky, Math. USSR Izv., 1973, 7(3), 479–496

[107] Ribnere E., Sequences of words characterizing finite solvable groups, Monatsh. Math., 2009, 157(4), 387–401

[108] Rosset M., Rosset S., Elements of trace zero that are not commutators, Comm. Algebra, 2000, 28(6), 3059–3072

[109] Saxl J., Wilson J.S., A note on powers in simple groups, Math. Proc. Cambridge Philos. Soc., 1997, 122(1), 91–94

[110] Schul G., Shalev A., Words and mixing times in finite simple groups, Groups Geom. Dyn., 2011, 5(2), 509–527

[111] Segal D., Words: Notes on Verbal Width in Groups, London Math. Soc. Lecture Note Ser., 361, Cambridge University Press, Cambridge, 2009

[112] Serre J.-P., Le groupe de Cremona et ses sous-groupes finis, In: Séminaire Bourbaki, 2008/2009 (997–1011), Astérisque, 2010, 332(1000), 75–100

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

[114] Shalev A., Word maps, conjugacy classes, and a noncommutative Waring-type theorem, Ann. of Math., 2009, 170(3), 1383–1416

[115] Shalev A., Applications of some zeta functions in group theory, In: Zeta Functions in Algebra and Geometry, Palma de Mallorca, May 3–7, 2010, Contemp. Math., 566, American Mathematical Society, Providence, 2012, 331–344

[116] Slusky M., Zeros of 2×2 matrix polynomials, Comm. Algebra, 2010, 38(11), 4212–4223

[117] Suzuki M., On a class of doubly transitive groups, Ann. of Math., 1962, 75(1), 105–145

[118] Thom A., Convergent sequences in discrete groups, Canad. Math. Bull., 2013, 56(2), 424–433

[119] Thompson J.G., Nonsolvable finite groups all of whose local subgroups are solvable, Bull. Amer. Math. Soc., 1968, 74(3), 383–437

[120] Thompson R.C., Commutators in the special and general linear groups, Trans. Amer. Math. Soc., 1961, 101(1), 16–33

[121] Tiep P.H., Zalesskii A.E., Some characterizations of the Weil representations of the symplectic and unitary groups, J. Algebra, 1997, 192(1), 130–165

[122] Tits J., Free subgroups in linear groups, J. Algebra, 1972, 20(2), 250–270

[123] Varshavsky Ya., Lefschetz–Verdier trace formula and a generalization of a theorem of Fujiwara, Geom. Funct. Anal., 2007, 17(1), 271–319

[124] Vogt H., Sur les invariants fondamentaux des equations différentielles linéaires du second ordre, Ann. Sci. École Norm. Supér., 1889, 6(Suppl.), 3–70

[125] Wan D., A p-adic lifting and its application to permutation polynomials, In: Finite Fields, Coding Theory, and Advances in Communications and Computing, Las Vegas, August 7–10, 1991, Lecture Notes in Pure and Appl. Math., 141, Marcel Dekker, New York, 1993, 209–216

[126] Wilson J.S., Characterization of the soluble radical by a sequence of words, J. Algebra, 2011, 326, 286–289

[127] Zelmanov E.I., On the restricted Burnside problem, In: Proceedings of the International Congress of Mathematicians, Kyoto, August 21–29, 1990, Mathematical Society of Japan, Tokyo, 1991, 395–402

[128] Zorn M., Nilpotency of finite groups, Bull. Amer. Math. Soc., 1936, 42(7), 485–486