ENGEL-LIKE IDENTITIES
CHARACTERISING FINITE SOLVABLE GROUPS

TATIANA BANDMAN, GERT-MARTIN GREUEL, FRITZ GRUNEWALD,
BORIS KUNYAVSKIĬ, GERHARD PFISTER, AND EUGENE PLOTKIN

Contents

1. Introduction

1.1. **Statement of the problem and main results.** In this paper we characterise solvable groups in the class of finite groups by identities in two variables. The starting point for this research is the following classical fact: the class of finite nilpotent groups is characterised by Engel identities. To be more precise, Zorn's theorem [Zo], [H, Satz III.6.3] says that a finite group G is nilpotent if and only if it satisfies one of the identities $e_n(x, y) = [y, x, x, \dots, x] = 1$ (here $[y, x] = yxy^{-1}x^{-1}$, $[y, x, x] = [[y, x], x]$, etc.).

Our goal is to obtain a similar characterisation of solvable groups in the class of finite groups. We say that a sequence of words $u_1, \dots, u_n, \dots$ is correct if $u_k \equiv 1$ in a group G implies $u_m \equiv 1$ in a group G for all $m > k$. We have found an explicit correct sequence of words $u_1(x, y), \dots, u_n(x, y), \dots$ such that a group G is solvable if and only if for some n the word u_n is an identity in G.

B. Plotkin suggested some Engel-like identities which could characterise finite solvable groups (see [PPT], [GKNP]). In the present paper we establish B. Plotkin's conjecture (in a slightly modified form).

Define

$$u_1(x, y) := x^{-2}y^{-1}x, \qquad \text{and inductively} \qquad u_{n+1}(x, y) := [\, x\, u_n(x, y)\, x^{-1},\ y\, u_n(x, y)\, y^{-1}\,]. \tag{1.1}$$

Note that sequence (1.1) is correct.

Our main result is

Theorem 1.1. *A finite group G is solvable if and only if for some n the identity $u_n(x, y) \equiv 1$ holds in G.*

Note two obvious properties of the initial word $w = x^{-2}y^{-1}x$: (1) if a group G satisfies the identity $w \equiv 1$, then $G = \{1\}$; (2) the words w and x generate the free group $F = \langle x, y\rangle$. Thus w can also be used as the initial term of a sequence characterising finite nilpotent groups, see Proposition 4.1. We shall discuss the choice of the initial word below. We conjecture after long computer experiments that Theorem 1.1 holds for any sequence formed like in (1.1) from any initial word not of the form $w = (x^{-1}y)^k$ ($k \in \mathbb{N}$).

Our results can be viewed as a natural development of the classical Thompson–Flavell theorem ([Th], [Fl]), stating that if G is a finite group in which every two elements generate a solvable subgroup, then G is solvable. Of course, Theorem 1.1 immediately implies this theorem (see Corollary 4.16 for an analogous statement in the pro-finite setting). As mentioned in [BW], the Thompson–Flavell theorem, together with [Br, Satz 2.12], implies that finite solvable groups can be characterised by a countable set of two-variable identities. (This fact also follows from Lemma 16.1 and Theorem 16.21 from [Ne] saying that an n-generator group G belongs to a variety V if and only if all n-variable identities from V are fulfilled in G.) However, this does not provide explicit two-variable identities for finite solvable groups. Furthermore, in the above cited paper, R. Brandl and J. S. Wilson construct a countable set of words $w_n(x, y)$ with the property that a finite group G is solvable if and only if for almost all n the identity $w_n(x, y) \equiv 1$ holds in G. Since in their construction there is no easily described relationship between terms of $w_n(x, y)$, they raise the question whether one can characterise finite solvable groups by sequences of identities fitting into a simple recursive definition.

Recently A. Lubotzky proved that for any integer $d \geq 2$ the free pro-solvable group $\hat{F}_d(S)$ can be defined by a *single* pro-finite relation [Lu, Prop. 3.4]. Using this proposition and Thompson's theorem, one can derive the existence of a needed sequence of identities characterising finite solvable groups (Lubotzky's result does not give, however, any candidate for such a sequence).

The sequence constructed in our Theorem 1.1 answers the question of Brandl–Wilson and fits very well into a pro-finite setting (see Subsection 4.2).

One can mention here some more cases where certain interesting classes of finite groups were characterised by two-variable commutator identities [Br], [BP], [BN], [Gu], [GH], [Ni1], [Ni2]; see [GKNP] or the above cited papers for more details.

Although Theorem 1.1 is a purely group-theoretic result, its proof involves surprisingly diverse methods of algebraic geometry, arithmetic geometry, group thery, and computer algebra (note, however, a paper of Bombieri [Bo] which served for us as an inspiring example of such an approach). We want to emphasise a special role played by problem oriented software (particularly, the packages Singular and MAGMA): not only proofs but even the precise statements of our results would hardly have been found without extensive computer experiments.

Clearly in every solvable group the identities $u_n(x, y) \equiv 1$ are satisfied from a certain $n \in \mathbb{N}$ onward. We shall deduce the non-trivial "if" part of the theorem from the following

Theorem 1.2. *Let G be one of the following groups:* (1) $G = \mathrm{PSL}(2, \mathbb{F}_q)$ *where* $q \geq 4$ $(q = p^n$, *p a prime),* (2) $G = \mathrm{Sz}(2^n)$, *$n \in \mathbb{N}$, $n \geq 3$ and odd,* (3) G=$\mathrm{PSL}(3, \mathbb{F}_3)$. *Then there are x, $y \in G$ such that $u_1(x, y) \neq 1$ and $u_1(x, y) = u_2(x, y)$.*

Here $\mathrm{PSL}(n, \mathbb{F}_q)$ denotes the projective special linear group of degree n over $\mathbb{F}_q$. For $q = 2^m$ we denote by $\mathrm{Sz}(q)$ the Suzuki group (the twisted form of ${}^2B_2$, see [HB, XI.3]).

Let us show that Theorem 1.2 implies Theorem 1.1.

Assume that Theorem 1.2 holds, and suppose that there exists a non-solvable finite group in which the identity $u_n \equiv 1$ holds. Denote by G a minimal counterexample, that is, a finite non-solvable group of the smallest order with identity $u_n \equiv 1$. Such a G must be simple. Indeed, if H is a proper normal subgroup of G, then both H and G/H are solvable (because any identity remains true in the subgroups and the quotients). But the list of groups in Theorem 1.2 contains Thompson's list of finite simple groups all of whose subgroups are solvable [Th], hence G is one of the groups (1)–(3). Since sequence (1.1) is correct, the assumption $u_1(x, y) = u_2(x, y)$ implies that $u_2(x, y) = u_3(x, y) = \dots$. From $u_1 \neq 1$ it follows that the identity $u_n \equiv 1$ does not hold in G, contradiction.

Theorem 1.2 admits a generalisation which can easily be deduced from the classification of finite simple groups.

Corollary 1.3. *Let G be a finite non-abelian simple group. Then there are x, $y \in G$ such that $u_1(x, y) \neq 1$ and $u_1(x, y) = u_2(x, y)$.*

For small groups from the above list it is an easy computer exercise to verify Theorem 1.2. There are for example altogether 44928 suitable pairs x, y in the group $\mathrm{PSL}(3, \mathbb{F}_3)$; here is one of them:

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

The general idea of our proof can be roughly described as follows. For a group G in the list of Theorem 1.2, using its standard matrix representation over $\mathbb{F}_q$, we regard the entries of the matrices corresponding to x and y in this representation as variables, and thus interpret solutions of the equation $u_1(x, y) = u_2(x, y)$ as $\mathbb{F}_q$-rational points of an algebraic variety. Lang–Weil type estimates [LW] 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. Small values of q are checked case by case. 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 the x, y from appropriate Zariski-closed subsets only. In the next two subsections we discuss more details.

1.2. **The case $G = \mathrm{PSL}(2, \mathbb{F}_q)$.** We shall explain here a more general setup which will also shed some light on the somewhat peculiar choice of the word u_1 in (1.1).

Let w be a word in x, x^{-1}, y, y^{-1}. Let G be a group and x, $y \in G$. Define

$$u_1^w(x, y) := w, \qquad \text{and inductively} \qquad u_{n+1}^w(x, y) := [\, x\, u_n^w(x, y)\, x^{-1},\ y\, u_n^w(x, y)\, y^{-1}\,].$$

Let $R := \mathbb{Z}[t, z_1, z_2, z_3, z_4]$ be the polynomial ring over $\mathbb{Z}$ in five variables. Consider further the two following 2×2-matrices over R.

$$x(t) = \begin{pmatrix} t & -1 \\ 1 & 0 \end{pmatrix}, \qquad y(z_1, \dots, z_4) = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix}.$$

Let $\mathfrak{a}$ be the ideal of R generated by the determinant of y and by the 4 polynomials arising from the matrix equation $u_1^w(x, y) = u_2^w(x, y)$, and let $\mathcal{V}^w \subset \mathbb{A}^5$ be the corresponding closed set of 5-dimensional affine space. Let further $\mathfrak{a}_0$ be the ideal of R generated by the determinant of y and by the matrix entries arising from the equation $u_1^w(x, y) = 1$, and let $\mathcal{V}_0^w \subset \mathbb{A}^5$ be the corresponding closed set. Our approach aims at showing that $\mathcal{V}^w \setminus \mathcal{V}_0^w$ has points over finite fields. We have therefore searched for words w satisfying $\dim(\mathcal{V}^w) - \dim(\mathcal{V}_0^w) \geq 1$ and also $\dim(\mathcal{V}^w) \geq 1$. We have only found the following words with this property:

$$x^{-2}y^{-1}x,\ y^{-1}xy,\ yx^{-1}y^{-1},\ yxy^{-1},\ x^{-1}yxy^{-1}x,\ x^{-1}yx^{-1}y^{-1}x \tag{1.2}$$

The extra freedom one might get by introducing variables for the entries of x does not lead to more suitable results. Indeed, elements of GL(2) act on the corresponding varieties by conjugation, and every matrix of determinant 1 except ± 1 is conjugate (over any field) to a matrix with entries like in $x(t)$.

For the last 5 words in (1.2) the corresponding closed sets $\mathcal{V}^w$ have no absolutely irreducible components outside $\mathcal{V}_0^w$, and in fact the analogue of Theorem 1.2 does not hold for them. For the first word $w = x^{-2}y^{-1}x$ the closed set $\mathcal{V}^w$ has 2 irreducible components. One of them is $\mathcal{V}_0^w$, the second which we call $\mathcal{S}$ has dimension 2 and is absolutely irreducible. The map $\varphi\colon \mathcal{S} \to \mathbb{A}^1 \setminus \{0\}$, $\varphi(x, y) = z_1$, is a fibration with curves of genus 8 as fibres. We now consider the fibre $\varphi^{-1}(1)$ and thus arrive to the matrices of the form

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

To give the precise form of this curve which is used in computations, we write the equation $u_1(x, y) = u_2(x, y)$ in an equivalent form

$$x^{-1}yx^{-1}y^{-1}x^2 = yx^{-2}y^{-1}xy^{-1}. \tag{1.3}$$

On substituting $x(t)$ instead of x and $y(b, c)$ instead of y, we obtain a matrix equation giving rise to the following

Definition 1.4. We denote by $I \subset \mathbb{Z}[b, c, t]$ be the ideal generated by the four polynomials arising after equating the matrix entries in (1.3), and let C be the corresponding algebraic set.

The following theorem will be proved in Section 2:

Theorem 1.5. *For any prime p the reduction of C modulo p is an absolutely irreducible curve.*

We now use the classical Hasse–Weil bound (in a slightly modified form adapted for singular curves, cf. [FJ, Th. 3.14], [AP], [LY]).

Lemma 1.6. *Let D be an absolutely irreducible projective algebraic curve defined over a finite field $\mathbb{F}_q$, and let $N_q = \#D(\mathbb{F}_q)$ denote the number of its rational points. Then $|N_q - (q + 1)| \leq 2p_a\sqrt{q}$, where p_a stands for the arithmetic genus of D (in particular, if D is a plane curve of degree d, $p_a = (d-1)(d-2)/2$).*

In fact, we need an affine version of the lower estimate of Lemma 1.6 (cf. [FJ, Th. 4.9, Cor. 4.10]) based on the fact that an affine curve C has at most $\deg(\overline{C})$ rational points less than the projective closure $\overline{C}$.

Corollary 1.7. *Let $C \subset \mathbb{A}^n$ be an absolutely irreducible affine curve defined over the finite field $\mathbb{F}_q$ and $\overline{C} \subset \mathbb{P}^n$ the projective closure. Then the number of $\mathbb{F}_q$-rational points of C is at least $q + 1 - 2p_a\sqrt{q} - d$ where d is the degree and p_a the arithmetic genus of $\overline{C}$.*

To apply Lemma 1.6 (or Corollary 1.7) we have to compute the arithmetic genus of the curve C (or the degree of some plane projection of C) and to prove that the curve is absolutely irreducible (which is the most technically difficult part of the proof, see Section 2 for more details). Computations give $d = 10$ and $p_a = 12$. This implies that for $q > 593$ there exist enough $\mathbb{F}_q$-rational points on C to prove Theorem 1.2 in the case of the groups PSL(2).

Remark 1.8. Consider the initial word $w = [x, y]$. The ideal $\mathfrak{a}$ corresponding to the variety $\mathcal{V}^w$ contains the polynomial $(-tz + v - w)(v + w)$. Let $\mathcal{V}_1^w$ be the closed set defined by the ideal generated by $\mathfrak{a}$ and $v + w$. This variety has 5 components: one is two-dimensional and equals $\mathcal{V}_0^w$, and 4 others are of dimension 0; each of them decomposes into 4 absolutely irreducibles components over a splitting field of the polynomial $5z^4 + 20z^3 + 36z^2 + 32z + 16$. Let $\mathcal{V}_2^w$ be the closed set defined by the ideal generated by $\mathfrak{a}$ and $-tz + v - w$. This variety also has 5 components, all of dimension 1; one of them is contained in $\mathcal{V}_0^w$ and each other decomposes into 3 absolutely irreducibles components over the splitting field of the polynomial $t^2 + t - 1$. Since none of the components, except for the one corresponding to trivial solutions of $u_1 = u_2$, is absolutely irreducible, our method fails for the initial word $w = [x, y]$. In fact, the analogue of Theorem 1.2 does not hold for this word.

1.3. **The case $G = \mathrm{Sz}(q)$.** To prove Theorem 1.2, the Suzuki groups $G = \mathrm{Sz}(q)$ ($q = 2^n$, n odd) provide the most difficult case. This is due to the fact that although $\mathrm{Sz}(q)$ is contained in $\mathrm{GL}(4, \mathbb{F}_q)$, it is not a Zariski-closed set. In fact the group $\mathrm{Sz}(q)$ is defined with the help of a field automorphism of $\mathbb{F}_q$ (the square root of the Frobenius), and hence the standard matrix representation for $\mathrm{Sz}(q)$ contains entries depending on q. We shall describe now how our problem can still be treated by methods of algebraic geometry.

Let $R := \mathbb{F}_2[a, b, c, d, a_0, b_0, c_0, d_0]$ be the polynomial ring over $\mathbb{F}_2$ in eight variables. Let $\pi\colon R \to R$ be its endomorphism defined by $\pi(a) = a_0$, $\pi(a_0) := a^2, \dots, \pi(d) := d_0$, $\pi(d_0) := d^2$. Let $\mathbb{F}$ be the algebraic closure of $\mathbb{F}_2$ and consider $a, \dots, d_0$ as the coordinates of eight dimensional affine space $\mathbb{A}^8$ over $\mathbb{F}$. The endomorhism π defines an algebraic bijection $\alpha : \mathbb{A}^8 \to \mathbb{A}^8$. The square of α is the Frobenius automorphism on $\mathbb{A}^8$ (note that a similar operator appears in [DL, Section 11]). Let $p \in \mathbb{A}^8$ be a fixed point of α^n, then its coordinates are in $\mathbb{F}_{2^n}$ if n is odd and in $\mathbb{F}_{2^{n/2}}$ if n is even.

Consider further the two following matrices in $\mathrm{GL}(4, R)$:

$$x = \begin{pmatrix} a^2a_0 + ab + b_0 & b & a & 1 \\ aa_0 + b & a_0 & 1 & 0 \\ a & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \qquad y = \begin{pmatrix} c^2c_0 + cd + d_0 & d & c & 1 \\ cc_0 + d & c_0 & 1 & 0 \\ c & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{1.4}$$

The matrices x, y also define maps from $\mathbb{A}^8$ to $\mathrm{GL}(4, \mathbb{F})$. It can easily be checked that the matrices corresponding to a fixed point of α^n (n odd and $n \geq 3$) lie in $\mathrm{Sz}(2^n)$.

Definition 1.9. Let $\mathfrak{a}$ be the ideal of R generated by the 16 polynomials arising from the matrix equation (1.3), where x and y are taken from (1.4), let $\mathfrak{a}' = \mathfrak{a} : a^3c_0^2$, and let V (resp. V') denote the closed set in $\mathbb{A}^8$ corresponding to $\mathfrak{a}$ (resp. $\mathfrak{a}'$). Let $U = V' \smallsetminus S$ where S is defined by the equation $cc_0 = 0$.

The varieties V' and U are needed to understand the geometric structure of V. In fact, V' is the unique top dimensional component of V, and U is a smooth open subset of V'. The following theorem will be proved in Section 3.

Theorem 1.10. (1) $\dim(V) = \dim(V') = 2$, (2) $\pi(\mathfrak{a}) = \mathfrak{a}$, $\pi(\mathfrak{a}') = \mathfrak{a}'$.

We thus see that α defines an algebraic map $\alpha\colon V \to V$. Our task now becomes to show that α^n (n odd and $n \geq 3$) has a non-zero fixed point on the surface V. Our basic tool is the Lefschetz trace formula resulting from Deligne's conjecture proved by Fujiwara [Fu]. To apply this formula, we replace V by U.

Theorem 1.11. *U is a smooth, affine, absolutely irreducible surface invariant under α. We have $b^1(U) \leq 675$ and $b^2(U) \leq 2^{22}$.*

Here $b^i(U) = \dim H^i_{\text{ét}}(U, \overline{\mathbb{Q}}_\ell)$ stand for the ℓ-adic Betti numbers of U. We use results of Adolphson–Sperber [AS] and Ghorpade–Lachaud [GL] to get the above estimates.

Since U is non-singular, the ordinary and compact Betti numbers of U are related by Poincaré duality, and we have $b^i_{\text{c}}(U) := \dim H^i_{\text{c}}(U, \overline{\mathbb{Q}}_\ell) = b^{4-i}(U)$. Since U is affine, $b^i(U) = 0$ for $i > 2$. Since U is absolutely irreducible, $b^0(U) = 1$ and the Frobenius acts on the one-dimensional vector space $H^0_{\text{ét}}(U, \overline{\mathbb{Q}}_\ell)$ as multiplication by 4. The operator α induces linear self-maps of all these cohomology groups. The above properties of U imply that the Lefschetz trace formula holds in the form

$$\#\operatorname{Fix}(U, n) = \sum_{i=0}^{4} (-1)^i \operatorname{tr}\bigl(\alpha^n \mid H^i_{\text{c}}(U, \overline{\mathbb{Q}}_\ell)\bigr),$$

where $\operatorname{Fix}(U, n)$ is the set of fixed points of α^n acting on U ($n > 1$ is an odd integer).

Note that α acts on $H^0_{\text{ét}}(U, \overline{\mathbb{Q}}_\ell)$ as multiplication by 2. (Indeed, if it were multiplication by (-2), for a sufficiently big power of α the right-hand side of the trace formula would be negative.) Hence α^n acts as multiplication by 2^n. Thus $\operatorname{tr}(\alpha^n | H^4_{\text{c}}(U, \overline{\mathbb{Q}}_\ell)) = 2^n$.

We infer from Deligne's estimates for the eigenvalues of the endomorphism induced by α on étale cohomology the following inequality:

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

An easy estimate then shows that $\#\mathrm{Fix}(U, n) \ne 0$ for $n > 48$. The cases $n < 48$ are checked with the help of MAGMA, and this finishes the proof of Theorem 1.2 (and hence Theorem 1.1). More details can be found in Section 3.

Remark 1.12. As a by-product of these computations, we found the first terms of the zeta-function of the operator α acting on the set U. This is a rational function defined by

$$Z_U(\alpha, T) := \exp\left(-\sum_{n=1}^{\infty} \frac{\#\operatorname{Fix}(U, n)}{n} T^n\right).$$

We have found that $Z_U(\alpha, T)$ equals

$$\frac{(1-2T)(1-T)(1-T^2)^8(1+T^2)^3(2T^4+2T^2+1)(4T^8+2T^4+1)(2T^2+2T+1)(8T^6+4T^5+T+1)}{(1-2T^2)^3}$$

up to terms of order T^{33}. Note that the absolute values of the zeros and poles of this rational function are all equal to 1, $1/2$, $1/\sqrt{2}$, or $1/\sqrt[4]{2}$, as general theory predicts. This formula suggests heuristic values $b^4_c(U) = 1$, $b^3_c(U) = 6$, $b^2_c(U) = 43$.

1.4. **Analogues, problems, and generalisations.** First, let us mention the following analogue of Levi–van der Waerden's problems for nilpotent groups (cf. [H, §3.6]):

Problem 1.13. Fix $n \in \mathbb{N}$ and assume that a finite group G satisfies the identity $u_n(x, y) \equiv 1$. What can be said about the solvability length of G?

If $n = 1$, then $G = \{1\}$. If $n = 2$, then G is nilpotent class at most 3.

Further on, Theorem 1.1 admits some natural analogues in Lie-algebraic and group-schematic settings [GKNP]. In particular, the following analogue of the classical Engel theorem on nilpotent Lie algebras is true.

Theorem 1.14. [GKNP] *Let L be a finite dimensional Lie algebra defined over an infinite field k of characteristic different from 2, 3, 5. Define*

$$v_1 = [x, y], \quad v_{n+1} = [[v_n, x], [v_n, y]] \quad (n > 1). \tag{1.5}$$

Then L is solvable if and only if for some n one of the identities $v_n(x, y) \equiv 0$ holds in L.

(Here [,] are Lie brackets.)

A much more challenging question is related to the infinite-dimensional case. Namely, the remarkable Kostrikin–Zelmanov theorem on locally nilpotent Lie algebras [Ko], [Ze2], [Ze3] and Zelmanov's theorem [Ze1] lead to the following

Problem 1.15. Suppose that L is a Lie algebra over a field k, the v_n's are defined by formulas (1.5), and there is n such that the identity $v_n(x,y) \equiv 0$ holds in L. Is it true that L is locally solvable? If k is of characteristic 0, is it true that L is solvable?

Of course, it would be of significant interest to consider similar questions for arbitrary groups.

We call G an *Engel* group if there is an integer n such that the Engel identity $e_n(x,y) \equiv 1$ holds in G.

We call G an *unbounded Engel* group if for every $x, y \in G$ there is an integer $n = n(x,y)$ such that $e_n(x,y) = 1$.

We introduce the following

Definition 1.16. We call G a *quasi-Engel* group if there is an integer n such that the identity $u_n(x,y) \equiv 1$ holds in G.

Definition 1.17. We call G an *unbounded quasi-Engel* group if for every $x, y \in G$ there is an integer $n = n(x,y)$ such that $u_n(x,y) = 1$.

Problem 1.18. Is every Engel group locally nilpotent?

Problem 1.19. Is every quasi-Engel group locally solvable?

(A property is said to hold locally if it holds for all finitely generated subgroups.)

Problem 1.18 remains open for a long time, cf. [Plo3]. The answer in general is most likely negative, however, some positive results are known [BM], [Gr], [Plo1], [Plo2], [Wi], [WZ], etc. In the solvable case the situation is even less clear. We dare to state the following

Conjecture 1.20. *Every residually finite, quasi-Engel group is locally solvable.*

(A group is said to be residually finite if the intersection of all its normal subgroups of finite index is trivial.)

For pro-finite groups the situation looks more promising.

Theorem 1.21. [WZ, Th. 5] *Every pro-finite, unbounded Engel group is locally nilpotent.*

Conjecture 1.22. *Every pro-finite, unbounded quasi-Engel group is locally solvable.*

It is quite natural to consider restricted versions of Problems 1.18 and 1.19 as is considered for the Burnside problem. Let E_n be the Engel variety defined by the identity $e_n \equiv 1$. Let $F = F_{k,n}$ be the free group with k generators in the variety E_n. One can prove that the intersection of all co-nilpotent normal subgroups H_α in F is also co-nilpotent. Hence there exists a group $F^0_{n,k}$ in E_n such that every nilpotent group $G \in E_n$ with k generators is a homomorphic image of $F^0_{n,k}$. This implies that all locally nilpotent groups from E_n form a variety. In other words, the restricted Engel problem has a positive solution. The situation with the restricted quasi-Engel problem is unclear.

Problem 1.23. Let $F = F_{k,n}$ be the free group with k generators in the variety of all quasi-Engel groups with fixed n. Is it true that the intersection of all co-solvable normal subgroups in $F = F_{k,n}$ is also co-solvable?

Our main theorem can be reformulated in pro-finite terms.

Theorem 1.24. *Let $F = F(x,y)$ denote the free group in two variables, and let $\widehat{F}$ be its pro-finite completion. Let $v_1, v_2, \dots, v_m, \dots$ be any convergent subsequence of* (1.1) *with limit f from $\widehat{F}$. Then the identity $f \equiv 1$ defines the pro-finite variety of pro-solvable groups.*

(See Section 4.2 for more details.)

It would be of great interest to consider the restricted quasi-Engel problem for pro-finite groups.

Remark 1.25. There is no sense in generalising Conjecture 1.22 too far: from the Golod–Shafarevich counterexamples one can deduce an example of an unbounded quasi-Engel group which is not locally nilpotent (and hence not locally solvable). We thank B. Plotkin for this observation.

Consider an interesting particular case of linear groups.

Corollary 1.26. *Suppose that $G \subset \mathrm{GL}(n, K)$ where K is a field. Then G is solvable if and only if it is quasi-Engel.*

Proof. The "only if" part is obvious. The "if" part is an immediate consequence of Theorem 1.1 and Platonov's theorem [Pla] stating that every linear group over a field satisfying a non-trivial identity has a solvable subgroup of finite index. (Of course, if K is of characteristic zero, the assertion follows from the Tits alternative [Ti].) □

Here is one more application of Theorem 1.2: it generates short presentations of finite simple groups. Let $\mathbf{B}$ be the group generated by x, y with the single relation $u_1(x,y) = u_2(x,y)$, that is $\mathbf{B} = \langle x, y \,|\, u_1 = u_2 \rangle$. The solvable quotients of $\mathbf{B}$ are all cyclic, but $\mathbf{B}$ has at least all minimal simple groups from Thompson's list as quotients. We for example found that $\mathrm{PSL}(2, \mathbb{F}_5) = \big\langle x, y \,|\, u_1 = u_2,\, x^3 = y^2 = 1 \big\rangle$ and

$$\mathbf{Sz}(8) = \big\langle x, y \,|\, u_1 = u_2,\, x^7 = y^5 = (xy^2)^5 = (x^{-1}y^{-1}xy^2)^2 = 1 \big\rangle .$$

Acknowledgements. Bandman, Kunyavskiĭ, and Plotkin were partially supported by the Ministry of Absorption (Israel), the Israeli Science Foundation founded by the Israeli Academy of Sciences — Center of Excellence Program, and the Minerva Foundation through the Emmy Noether Research Institute of Mathematics. Kunyavskiĭ and Plotkin were also supported by the RTN network HPRN-CT-2002-00287 and INTAS 00-566. Greuel and Pfister were partially supported by the DFG project "Globale Methoden in der komplexen Geometrie" as well as by the Stiftung Rheinland–Pfalz für Innovation. Greuel was also supported by the German–Israeli Foundation for Scientific Research and Development, G-616-15.6/1999. Bandman and Kunyavskiĭ thank the Max-Planck-Institut für Mathematik (Bonn) which they visited when preparing this paper for publication.

We are grateful to N. Gordeev, D. Grayson, L. Illusie, A. Lubotzky, A. Mann, S. Margolis, R. Pink, J. Piontkowski, L. Rowen, D. Segal, Y. Segev, J.-P. Serre, M. Stoll, Y. Varshavsky, and N. Vavilov for useful comments and advice. We thank D. Nikolova and R. Shklyar for help in computer experiments and H. Schönemann for extending the functionality of the Singular kernel. Our special thanks go to B. Plotkin for numerous enlightening, encouraging, and inspiring discussions.

Notation. Because of extensive use of the Singular package our notation sometimes differs from the standard one: say, in the output of computer sessions, powers like a^{12} are denoted as `a12` . We refer the reader to [GP3], [GP4], [GPS] for definitions of Singular commands and their usage, and to [Bu], [GP1]–[GP3] for details on Gröbner bases.

All other notations are more or less standard.

Rings and fields: All rings are assumed commutative with 1; $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{F}_q$ denote the ring of integers, the field of rational numbers, the field of q elements, respectively. $\overline{k}$ denotes a (fixed) algebraic closure of a field k.

Ideals and varieties: If I is an ideal in R and $i\colon R \to S$ is a ring homomorphism, IS stands for the image of I under i. The ideal generated by $f_1, \dots f_k$ is denoted $\langle f_1, \dots, f_k\rangle$.

For $f \in R$ we denote $I : f^\infty = \bigcup_{n=1}^{\infty} I : f^n$. If R is noetherian, the chain of ideals $I : f \subseteq I : f^2 \subseteq \dots$ stabilises, and we have $I : f^\infty = I : f^n$ for some n.

$\mathbb{A}^n$ and $\mathbb{P}^n$ denote affine and projective spaces. $\overline{C}$ denotes the projective closure of an affine set $C \subset \mathbb{A}^n$, and I_h stands for the homogenisation of an ideal I. $\mathbf{V}(J)$ denotes the affine variety defined of the ideal J. If $\mathbf{V}(J) \subset \mathbb{A}^n$, we denote $\mathbf{D}(J) = \mathbb{A}^n \setminus \mathbf{V}(J)$. We shorten $\mathbf{V}(\langle f_1, \dots, f_k\rangle)$ to $\mathbf{V}(f_1, \dots, f_k)$, and $\mathbf{D}(\langle f_1, \dots, f_k\rangle)$ to $\mathbf{D}(f_1, \dots, f_k)$. We denote by $V(k)$ the set of rational points of a k-variety V.

$\chi(V)$ denotes the Euler characteristic of a variety V.

If D is a projective curve (maybe singular), $p_a(D)$ is the arithmetic genus of D, and $g(D)$ denotes the genus of the normalisation of D.

All other notations will be explained when needed.

2. The details of the PSL(2) case

2.1. **Proof of the main result.** Our goal is to prove Theorem 1.5 and to compute the arithmetic genus of C. This will lead us to the following

Proposition 2.1. *If $q = p^k$ for a prime p and $q \neq 2, 3$, then there are x, y in $\mathrm{PSL}(2, \mathbb{F}_q)$ with $y \neq x^{-1}$ and $u_1(x, y) = u_2(x, y)$.*

Note that for $w = x^{-2}y^{-1}x$, the equation $u_1(x, y) = u_2(x, y)$ has a non-trivial solution if and only if it has a solution with $y \neq x^{-1}$.

The proof will use some explicit computations with the following matrices. Let R be a commutative ring with identity. Recall that we defined

$$x(t) = \begin{pmatrix} t & -1 \\ 1 & 0 \end{pmatrix}, \qquad y(b, c) = \begin{pmatrix} 1 & b \\ c & 1 + bc \end{pmatrix} \in \mathrm{SL}(2, R)$$

for $t, b, c \in R$.

Remark 2.2.

(1) We have

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

for t, b, c.

(2) For any $t, b, c \in R$ we have $y(b, c) \neq x(t)^{-1}$, even for the images of $x(t)$ and $y(b, c)$ in $\mathrm{PSL}(2, R)$.

The equation $u_1 = u_2$ is equivalent to $x^{-1}yx^{-1}y^{-1}x^2 = yx^{-2}y^{-1}xy^{-1}$; we put $x = x(t)$, $y = y(b, c)$, and write

$$x^{-1}yx^{-1}y^{-1}x^2 - yx^{-2}y^{-1}xy^{-1} = \begin{pmatrix} n_1(t, b, c) & n_2(t, b, c) \\ n_3(t, b, c) & n_4(t, b, c) \end{pmatrix}.$$

Let $I = \langle n_1, n_2, n_3, n_4\rangle \subseteq \mathbb{Z}[b, c, t]$ be the ideal generated by the entries of the matrix.

Using Singular we can obtain I as follows:[1]

[1]A file with all Singular computations can be found at `http://www.mathematik.uni-kl.de/~pfister/SolubleGroups`.

```
LIB"linalg.lib";  option(redSB);
ring R = 0,(c,b,t),(c,lp);
matrix X[2][2] = t, -1,
                 1,  0;
matrix Y[2][2] = 1, b,
                 c, 1+bc;
matrix iX = inverse(X);     matrix iY = inverse(Y);
matrix M=iX*Y*iX*iY*X*X-Y*iX*iX*iY*X*iY;  ideal I=flatten(M);   I;
I[1]=c2b3t2+c2b2t3-c2b2t2+c2b2t+c2b2-c2bt3+2c2bt2+c2bt-c2t2+c2t+c2-cb3t
     +cb2t2+cb2t+cbt3-cbt2+cbt+2cb-ct3+ct2+2ct+c-b2t+bt+1
I[2]=c2b2t+c2bt2+c2t-cb3t2-cb2t3-cb2t-cb2-2cbt2+cbt+ct2-ct-c+b3t-bt-b-1
I[3]=c3b3t2+c3b2t3+c3b2t+2c3bt2+c3t-c2b3t-c2b2t3+2c2b2t2+c2b2t-c2bt4
     +2c2bt3+c2bt2+c2bt-c2t3+2c2t2+c2t+2cb2t2-2cb2t-cb2+cbt2+cbt+cb-ct4
     +ct3+3ct2-c-b2t+bt2-bt-b+1
I[4]=-c2b3t2-c2b2t3+c2b2t2-c2b2t+c2bt3-2c2bt2+c2t2-c2t+cb3t-cb2t2-2cb2t
     -cbt3-cbt+ct3-ct2-2ct+b2t+b2-bt-b-t+1
```

Denote by C the $\mathbb{F}_q$-variety defined by the ideal $I\mathbb{F}_q[b,c,t]$.

To prove Proposition 2.1, it is enough to prove

Proposition 2.3. *Let q be as in Proposition* 2.1*, then the set $C(\mathbb{F}_q)$ of rational points of C is not empty.*

The proof is based on the Hasse–Weil estimate (see Corollary 1.7).

Note that the Hilbert function of $\overline{C}$, $H(t) = dt - p_a + 1$, can be computed from the homogeneous ideal I_h of $\overline{C}$, hence we can compute d and p_a without any knowledge about the singularities of $\overline{C}$. The ideal I_h can be computed by homogenising the elements of a Gröbner basis of I with respect to a degree ordering (cf. [GP3]).

In the following let $q = p^k$ be an arbitrary, fixed prime power and L the algebraic closure of $\mathbb{F}_q$. To apply Corollary 1.7, we have to prove

Proposition 2.4. *$IL[b,c,t]$ is a prime ideal.*

We start with the following

Lemma 2.5. *The following polynomials form a Gröbner basis of $IL[c,b,t]$ with respect to the lexicographical ordering $c > b > t$,*

```
J[1]=(t2)*b4+(-t4+2t3)*b3+(-t5+3t4-2t3+2t+1)*b2+(t5-4t4+3t3+2t2)*b
      +(t4-4t3+2t2+4t+1)
J[2]=(t3-2t2-t)*c+(t2)*b3+(-t4+2t3)*b2+(-t5+3t4-2t3+2t+1)*b+(t5-4t4+3t3+2t2)
J[3]=(t)*cb+(-t2+2t+1)
J[4]=cb2+(-t2+2t+1)*c+(-t)*b3+(t3-2t2)*b2+(t4-3t3+2t2-t)*b+(-t4+4t3-3t2-2t)
J[5]=(t)*c2-cb+(t)*c+(-t2)*b3+(t4-2t3+t)*b2+(t5-3t4+t3+2t2-2t-1)*b
      +(-t5+3t4-4t2+t)
```

Proof. The Gröbner basis can be computed in Singular as follows (in characteristic 0):

```
ideal J=std(I);
```

Instead of relying on the Singular computation, we can verify "by hand" that J is a Gröbner basis for each q. Indeed, given some intermediate data obtained with the help of a computer, the truth of the lemma can be verified without computer. We first show that I and J generate the same ideal.

```
matrix M=lift(I,J); M;

M[1,1]=b2t4+2bt3-t4-t3+3t2+2t
M[1,2]=bt4-t4+2t3+t2+t
M[1,3]=t
M[1,4]=-bt3+t3-2t2-t
M[1,5]=-cb2t3-cbt4-ct3+ct+b2t2-bt4-bt+t6-t5-t4+t3-2t2-t
M[2,1]=-cbt5+cbt4+2cbt3-cbt2-cbt-ct4+ct2+ct-bt3+bt-t5+t4+2t3-2t2-2t
M[2,2]=cbt4-cbt2-cbt+ct2+ct+t4-t3-t2
M[2,3]=-cbt-t
M[2,4]=-cbt3+2cbt-ct-t3+t2+2t-1
M[2,5]=-cbt4+2cbt2-cbt+ct5-2ct4+ct3+ct2-3ct-bt+t5-3t4+t3+3t2-3t+1
M[3,1]=bt4-bt3-2bt2+bt+b+t3-t-1
M[3,2]=-bt3+bt+b-t-1
M[3,3]=b
M[3,4]=bt2-2b+1
M[3,5]=bt5-2bt4+bt3+2bt2-3bt+b-t5+2t4-2t2+3t
M[4,1]=cbt4-cbt3-2cbt2+cbt+cb+ct3-ct-c+b2t4+bt5-bt4+2bt2+bt+t4-2t3+3t+2
M[4,2]=-cbt3+cbt+cb-ct-c+bt2+bt-t4+t3+t2+t+1
M[4,3]=cb+bt+t+1
M[4,4]=cbt2-2cb+c-2bt+t3-t2-t-2
M[4,5]=-cb2t3+cbt5-3cbt4+cbt3+2cbt2-4cbt+cb-ct3-ct2+2ct+c+b2t2-bt2+t5-t4-3t+1
```

This implies that over $\mathbb{Z}$ and, hence, over each $\mathbb{F}_q$ the k-th generator of J satisfies

$$J[k] = \sum_{\ell=1}^{4} M[\ell,k]\cdot I[\ell]\,, \quad k=1,\dots,5\,.$$

```
matrix N= lift(J,I); N;

N[1,1]=-cb+c-1                              N[1,2]=b
N[1,3]=-c2b-1                               N[1,4]=cb-c+b+t
N[2,1]=cb2-cb+b-1                           N[2,2]=-b2+1
N[2,3]=c2b2+c+b-1                           N[2,4]=-cb2+cb-b2-bt+t-1
N[3,1]=cb2t+cbt+2cb+bt+t+2                  N[3,2]=cb+ct-b2t-2bt-2b-t-1
N[3,3]=c2b2t+2c2bt+2c2b+2c2t-cb2+ct-c+bt-b+t+2
N[3,4]=-cb2t-cbt-2cb+c+b2-bt-2b-3t+1        N[4,1]=c
N[4,2]= -1                                  N[4,3]=0
N[4,4]=0                                    N[5,1]=-1
N[5,2]=1                                    N[5,3]=c-1
N[5,4]=t-1
```

In the same way this implies that

$$I[k] = \sum_{\ell=1}^{5} N[\ell,k]\cdot J[\ell]\,, \quad k=1,\dots,4.$$

We proved that the polynomials $J[1],\dots,J[5]$ generate the ideal I.

To show that $J[1],\dots,J[5]$ is a Gröbner basis, we use Buchberger's criterion (cf. [GP3], Theorem 1.7.3). To see this for any q, we can use the same trick as above. Let $s = \mathrm{spoly}(J[i],J[j])$, $i<j$, be the s-polynomial of $J[i]$ and $J[j]$. We have to show that the normal form of s with respect to $J[1],\dots,J[5]$ is 0. We apply `lift(s,J);` in SINGULAR and use the result to check by hand that s is a linear combination of $J[1],\dots,J[5]$ in all characteristics. As this is similar to above, we dispense with the output. □

Lemma 2.6. *Let*

$$\begin{aligned}
f_1 =&t^2b^4 - t^3(t-2)b^3 + (-t^5+3t^4-2t^3+2t+1)b^2 + t^2(t^2-2t-1)(t-2)b + (t^2-2t-1)^2\\
f_2 =&t(t^2-2t-1)c + t^2b^3 + (-t^4+2t^3)b^2 + (-t^5+3t^4-2t^3+2t+1)b + (t^5-4t^4+3t^3+2t^2)\,,\\
h =&t(t^2-2t-1)\,.
\end{aligned}$$

Then the following holds for any prime power q.

(1) $\{f_1, f_2\}$ *is a Gröbner basis of* $IL(t)[b,c]$ *with respect to the lexicographical ordering* $c > b$*;*
(2) $I : h = I$*;*
(3) $IL(t)[b,c] \cap L[t,b,c] = \langle f_1, f_2\rangle : h^2 = I$.

Proof. Because J is a Gröbner basis of I with respect to the lexicographical ordering $c > b > t$, J is a Gröbner basis of $IL(t)[b,c]$ with respect to the lexicographical ordering $c > b$ (cf. [GP3, Chapter 4.3]). But $J[1] = f_1$ and $J[2] = f_2$ and, considered in $IL(t)[b,c]$, the leading monomials of f_1 and f_2 generate already the leading ideal of $IL(t)[b,c]$. This shows (1).

(3) is a consequence of (2) because $IL(t)[b,c] \cap L[t,b,c] = \langle f_1, f_2\rangle : h^\infty$, see [GP3, Prop. 4.3.1], and $h^2 I \subset \langle f_1, f_2\rangle$ that we shall see now.

```
M=lift(ideal(J[1],J[2]),h^2*I); M;

M[1,1]=(-t5+2t4+t3)*cb2+(-t6+3t5-2t4+3t2+t)*cb+(t6-4t5+2t4+4t3+t2)*c
       +(-t3)*b2+(t3)*b+(t2)
M[1,2]=(-t4+2t3+t2)*cb+(-t5+2t4+t3)*c+(t5-2t4)*b2+(t6-3t5+2t4-t3-3t2-t)*b
       +(-t6+4t5-3t4-2t3-t2)
M[1,3]=(-t5+2t4+t3)*c2b2+(-t6+2t5+2t3+t2)*c2b+(-2t5+4t4+2t3)*c2+(t4-t3-t2)*cb2
       +(-t5+2t4)*cb+(-t6+2t5+t4-t2)*c+(-t3)*b2+(t4-t3-t2)*b+(t2)
M[1,4]=(t5-2t4-t3)*cb2+(t6-3t5+2t4-t3-t2)*cb+(-t6+4t5-3t4-2t3)*c+(t3+t2)*b2
       +(-t3-t2)*b+(-t3+t2)
M[2,1]=(t5-2t4-t3)*cb3+(t6-3t5+2t4-3t2-t)*cb2+(-t6+4t5-2t4-4t3-t2)*cb
       +(-t5+3t4-3t2-t)*c+(t3)*b3+(-t3)*b2+(-t2)*b
M[2,2]=(t4-2t3-t2)*cb2+(t5-2t4-t3)*cb+(t4-2t3-t2)*c+(-t5+2t4)*b3
       +(-t6+3t5-2t4+t3+3t2+t)*b2+(t6-4t5+3t4+2t3+t2)*b+(t5-4t4+2t3+4t2+t)
M[2,3]=(t5-2t4-t3)*c2b3+(t6-2t5-2t3-t2)*c2b2+(2t5-4t4-2t3)*c2b+(t4-2t3-t2)*c2
       +(-t4+t3+t2)*cb3+(t5-2t4)*cb2+(t6-2t5-t4+t2)*cb+(-t4+2t3+t2)*c
       +(t3)*b3+(-t4+t3+t2)*b2+(-t2)*b
M[2,4]=(-t5+2t4+t3)*cb3+(-t6+3t5-2t4+t3+t2)*cb2+(t6-4t5+3t4+2t3)*cb+(t5-3t4+t3
       +t2)*c+(-t3-t2)*b3+(t3+t2)*b2+(t3-t2)*b
```

This implies

$$h^2 \cdot n_i = M[1,i]\cdot f_1 + M[2,i]\cdot f_2, \quad i = 1,\dots,4.$$

To prove (2) we can use the following Singular commands to see that $I : h \subset I$.

```
poly h=t*(t2-2t-1);
reduce(quotient(I,h),std(I));
_[1]=0      _[2]=0      _[3]=0      _[4]=0
```

If we want to check this by hand, we can use the following method to compute the quotient (cf. [GP3, 2.8.5]):

If $U = \langle [g_1, 0], \dots, [g_n, 0], [h, 1]\rangle$ is a submodule of the free module $L[c,b,t]^2$ and $[0, h_1], \dots, [0, h_r]$ is the part of the Gröbner basis of U (with respect to the ordering $(c, >)$ giving priority to the components (cf. [GP3, 2.3])), having the first component zero, then $\langle g_1, \dots, g_n\rangle : h = \langle h_1, \dots, h_r\rangle$.

```
module N=[J[1],0],[J[2],0],[J[3],0],[J[4],0],[J[5],0],[h,1];
module N1=std(N);  N1;
N1[1]=[0,b4t2-b3t4+2b3t3-b2t5+3b2t4-2b2t3+2b2t+b2+bt5-4bt4+3bt3+2bt2+t4
       -4t3+2t2+4t+1]
N1[2]=[0,ct3-2ct2-ct+b3t2-b2t4+2b2t3-bt5+3bt4-2bt3+2bt+b+t5-4t4+3t3+2t2]
N1[3]=[0,cbt-t2+2t+1]
N1[4]=[0,cb2-ct2+2ct+c-b3t+b2t3-2b2t2+bt4-3bt3+2bt2-bt-t4+4t3-3t2-2t]
N1[5]=[0,c2t-cb+ct-b3t2+b2t4-2b2t3+b2t+bt5-3bt4+bt3+2bt2-2bt-b-t5+3t4-4t2+t]
N1[6]=[t2-2t-1,-cb+t-2]
N1[7]=[b3-b2-bt+2b,cb-4ct2+10ct-c+b5t-b4t-b3t5+3b3t4-4b3t2-4b3t+4b3-b2t6
      +5b2t5-6b2t4+b2t3-3b2t2+4b2t-6b2+bt6-6bt5+14bt4-13bt3+4bt2-7bt-b+t5
      -10t4+27t3-17t2-8t-4]
N1[8]=[cb,-c2+2cb-ct2+2ct-b2+bt2-2bt+t3-3t2+4]
N1[9]=[c2+c+b2-b-t+2,c3b+c2b+4c2-5cb+ct2-3ct+5c+b4t-b3t3+b3t2-b3t+2b2t3
       -4b2t2+b2t+4b2+bt5-bt4-6bt3+5bt2+5bt-5b-t5+2t4+3t3-4t2-2t-1]
```

We see that in the second component of $N1[1], \ldots, N1[5]$ we have exactly the Gröbner basis J.

We have to check that $N = N1$ and that $N1$ is a Gröbner basis. The last claim follows again by using Buchberger's criterion ([GP3], Theorem 1.7.3). To see that $N = N1$, we compute

```
M=lift(N1,N); M;

M[1,1]=-c+b2t-b2-bt2+2bt+4b-t2+2t                M[1,2]=-b-2t2+3t+3
M[1,3]=0                                         M[1,4]=b+t2-2t-1
M[1,5]=b+2t2-3t-4                                M[1,6]=0
M[2,1]=-b3t-b3-b2t2+b2t+bt2-2bt+8b+t-2           M[2,2]=-b2+b+8
M[2,3]=1                                         M[2,4]=-4
M[2,5]=b2-b-9                                    M[2,6]=0
M[3,1]=-b2t-4b2-2bt+2b+t-2                       M[3,2]=c-b2t-4b-t+1
M[3,3]=-2                                        M[3,4]=b2+bt-b+1
M[3,5]=-c2-c+b2t+3b+4                            M[3,6]=1
M[4,1]=b3-b2+1                                   M[4,2]=b2-b
M[4,3]=0                                         M[4,4]=-2
M[4,5]=-b2+b                                     M[4,6]=0
M[5,1]=0                                         M[5,2]=0
M[5,3]=1                                         M[5,4]=0
M[5,5]=-t-2                                      M[5,6]=0
M[6,1]=b4-b3t2-b3-b2t3+b2t2-b2t+b2+bt3-2bt2+t2-2t-1
M[6,2]=ct+b3-b2t2-b2-bt3+bt2-bt+b+t3-2t2         M[6,3]=-1
M[6,4]=-c+b2t+bt2-bt-t2+2t
M[6,5]=-b3+b2t2+b2+bt3-bt2-b-t3+t2+t             M[6,6]=t
M[7,1]=2bt+b                                     M[7,2]=2t+1
M[7,3]=0                                         M[7,4]=-t
M[7,5]=-2t-1                                     M[7,6]=0
M[8,1]=0                                         M[8,2]=0
M[8,3]=t                                         M[8,4]=b
M[8,5]=-1                                        M[8,6]=0
M[9,1]=0                                         M[9,2]=0
M[9,3]=0                                         M[9,4]=0
M[9,5]=t                                         M[9,6]=0
```

This implies

$$N[k] = \sum_{\ell=1}^{9} M[\ell, k] \cdot N1[\ell], \quad k = 1, \ldots, 6.$$

```
M=lift(N,N1);

M[1,1]=0     M[1,2]=0     M[1,3]=0     M[1,4]=0     M[1,5]=0     M[1,6]=0
M[1,7]=0     M[1,8]=0     M[1,9]=0     M[2,1]=0     M[2,2]=0     M[2,3]=0
M[2,4]=0     M[2,5]=0     M[2,6]=0     M[2,7]=0     M[2,8]=0     M[2,9]=0
M[3,1]=-b2t3+2b2t2+b2t+t5-4t4+2t3+4t2+t              M[3,2]=-bt3+2bt2+bt
M[3,3]=-t3+2t2+t                                     M[3,4]=0
M[3,5]=0                                             M[3,6]=t2-2t-1
M[3,7]=-b2t4+3b2t3-3b2t-b2+bt5-5bt4+7bt3+3bt2-12bt+b+2t6-11t5+16t4+2t3-10t2-3t
M[3,8]=bt2-2bt-b-2t2+5t
M[3,9]=b2t5-5b2t4+5b2t3+2b2t2+2b2t+b2+2bt6-11bt5+18bt4-5bt3-7bt2+3bt-b+t7
       -6t6+13t5-11t4-4t3+14t2-4t+2
M[4,1]=bt4-2bt3-bt2                                  M[4,2]=t4-2t3-t2
M[4,3]=0                                             M[4,4]=-t3+2t2+t
M[4,5]=0                                             M[4,6]=0
M[4,7]=-ct3+2ct2+ct+bt5-4bt4+2bt3+5bt2-bt-b+t5-4t4-2t3+14t2
M[4,8]=-t3+2t2+t
M[4,9]=-ct4+2ct3+2ct+c-b2t4+2b2t3+b2t2+2bt3-4bt2-2bt+t6-4t5+3t4+t3+5t+1
M[5,1]=0                                             M[5,2]=0 M[5,3]=0
M[5,4]=0                                             M[5,5]=-t3+2t2+t
M[5,6]=0
M[5,7]=b2t2-2b2t-b2-t4+4t3-2t2-4t-1                  M[5,8]=t2-2t-1
M[5,9]=-cbt2+2cbt+cb+b2t3-2b2t2-b2t-t5+4t4-3t3-4t2+5t
M[6,1]=b4t2-b3t4+2b3t3-b2t5+3b2t4-2b2t3+2b2t+b2+bt5-4bt4+3bt3+2bt2+t4
       -4t3+2t2+4t+1
M[6,2]=ct3-2ct2-ct+b3t2-b2t4+2b2t3-bt5+3bt4-2bt3+2bt+b+t5-4t4+3t3+2t2
M[6,3]=cbt-t2+2t+1
M[6,4]=cb2-ct2+2ct+c-b3t+b2t3-2b2t2+bt4-3bt3+2bt2-bt-t4+4t3-3t2-2t
M[6,5]=c2t-cb+ct-b3t2+b2t4-2b2t3+b2t+bt5-3bt4+bt3+2bt2-2bt-b-t5+3t4-4t2+t
M[6,6]=-cb+t-2
M[6,7]=cb-4ct2+10ct-c+b5t-b4t-b3t5+3b3t4-4b3t2-4b3t+4b3-b2t6+5b2t5-6b2t4+b2t3
       -3b2t2+4b2t-6b2+bt6-6bt5+14bt4-13bt3+4bt2-7bt-b+t5-10t4+27t3-17t2-8t-4
M[6,8]=-c2+2cb-ct2+2ct-b2+bt2-2bt+t3-3t2+4
M[6,9]=c3b+c2b+4c2-5cb+ct2-3ct+5c+b4t-b3t3+b3t2-b3t+2b2t3-4b2t2+b2t+4b2+bt5
       -bt4-6bt3+5bt2+5bt-5b-t5+2t4+3t3-4t2-2t-1
```

This implies

$$N1[k] = \sum_{\ell=1}^{6} M[\ell,k]\cdot N[\ell]\,,\quad k=1,\dots,9\,,$$

and we obtain finally $N = N1$. □

We now continue the proof of Proposition 2.4. We have $IL(t)[b,c]\cap L[b,c,t] = \langle f_1, f_2\rangle : h^2 = IL[b,c,t]$. Therefore, if $IL[b,c,t]$ were reducible, then $IL(t)[b,c]$ would be reducible too. We are going to prove that this is not the case.

In $L(t)[b,c]$ the polynomial f_2 is linear in c. Since f_1 does not depend on c, we have $L(t)[b,c]/I \cong L(t)[b]/\langle f_1\rangle$ and, hence, it suffices to prove that the polynomial f_1 is irreducible.

Set $x = bt$, and let $p(x,t) = t^2 f_1(\frac{x}{t}, t)$, then

$$\begin{aligned} p(x,t) = {} & x^4 - t^2(t-2)x^3 + (-t^5+3t^4-2t^3+2t+1)x^2 + t^3(t-2)(t^2-2t-1)x \\ & + t^2(t^2-2t-1)^2\,. \end{aligned}$$

To prove that $f_1 \in L[t,b]$ is irreducible, it suffices to prove that $p \in L[x,t] = L[t][x]$ is irreducible. We first show that p has no linear and no quadratic factor with respect to x.

First we prove that p has no linear factor, that is, that $p(x)=0$ has no solution in $L[t]$.

Assume that $x(t)\in L[t]$ is a zero of $p(x)=0$. Then $x(t)\mid t^2(t^2-2t-1)^2$. If the characteristic of L is not 2, it is not difficult to see that $x(t)$ cannot contain the square of an irreducible factor of $t^2(t^2-2t-1)^2$. If the characteristic of L is 2, it is not possible that $t^2\mid x(t)$ or $(t+1)^3\mid x(t)$. Moreover, it is easy to see that the leading coefficient of $x(t)$ is $(-1)^{\deg(x(t))-1}$.

The following list gives the candidates for a zero of $p(x)$ and the value of $p(x)$.

If $\mathrm{char}(L)>2$:

x	leading term of $p(x)$
-1	$-t^5$
t	$-t^6$
$t-1-\sqrt{2}$	$\sqrt{2}t^6$
$t-1+\sqrt{2}$	$\sqrt{2}t^6$
$-t(t-1+\sqrt{2})$	$-\sqrt{2}t^8$
$-t(t-1-\sqrt{2})$	$\sqrt{2}t^8$
$-t^2+2t+1$	$-t^8$
$t(t^2-2t-1)$	$-t^{11}$.

If $\mathrm{char}(L)=2$:

x	$p(x)$
1	$t^5+t^3+t^2$
t	$t^6+t^5+t^4$
$t+1$	t^5+t^3
$t(t+1)$	t^7+t^5
$(t+1)^2$	$t^8+t^7+t^6+t^4+t^3+t^2$
$t(t+1)^2$	$t^{11}+t^9+t^8+t^7+t^5+t^4$.

This implies that $p(x)$ has no linear factor with respect to x in $L[x,t]$.

Now assume that $p(x)=(x^2+ax+b)(x^2+gx+d)$, $a,b,g,d\in L[t]$.

This implies:

$$\begin{aligned}
&(1) & bd &= t^2(t^2-2t-1)^2\\
&(2) & ad+bg &= t^3(t-2)(t^2-2t-1)\\
&(3) & d+ag+b &= -t^5+3t^4-2t^3+2t+1\\
&(4) & a+g &= -t^2(t-2).
\end{aligned}$$

If $t^2|b$ then, because of (2), we obtain $t^2|a$. (4) implies $t^2|g$ and (2) implies $t^3|a$. (3) implies that $d\equiv 1+2t \mod (t^2)$ and (4) implies that $g\equiv 2t^2 \mod (t^3)$. If $\mathrm{char}(L)\neq 2$, we obtain $d=-(t^2-2t-1)$ and $b=-t^2(t^2-2t-1)$, because $(t^2-2t-1)^2\equiv 1+4t \mod (t^2)$. If $\mathrm{char}(L)=2$, then $t^3|a$ and $t^3|g$. (2) implies that $\frac{a}{t^3}\cdot d+\frac{g}{t^3}b=(t+1)^2$. This implies $(t+1)^2|b$ and $(t+1)^2|d$. Therefore, we have in any characteristic $b=-t^2(t^2-2t-1)$ and $d=-(t^2-2t-1)$. (3) implies that $ag=-t^3(t-2)^2$. This is a contradiction to the fact that $t^3|a$ and $t^2|g$.

We showed that $t^2\nmid b$. Similarly, we obtain that $t^2\nmid d$. This implies that $t|b$ and $t|d$. If $(t^2-2t-1)^2|b$, then (2) implies that $t^2-2t-1|a$. Let $d=d_1t$ for a suitable $d_1\in L$, then (3) implies that $t^2-2t-1\mid -t^5+3t^4-2t^3+2t+1-d_1t$, that is, $d_1=-1$. Then $b=-t(t^2-2t-1)^2$. Now (3) implies that $ag=-t^4+4t^2+4t+1=-(t^2-2t-1)(t+1)^2$.

But $t^2-2t-1|a$ and (4) implies that $\deg(a)=3$ and $\deg(g)=1$. This implies that $t+1|a$ and $t+1|g$, which is a contradiction to (4).

Similarly, we obtain that $(t^2-2t-1)^2\nmid d$. This implies that $b=b_3t(t^2-2t-1)$ and $d=\frac{1}{b_3}t(t^2-2t-1)$ for a suitable $b_3\in L$. (3) implies that $\deg(ag)=5$. Because of (4), we may assume that $\deg(a)=3$ and $\deg(g)=2$. (4) implies that $a=-t^3+$ terms of lower degree. (3) implies that $g=t^2+$ terms of lower degree. (4) implies that $a = -t^3 + t^2+$ terms of lower degree. (2) implies that $b_3 = -1$. (3) implies that $ag = -t^5 + 3t^4 - 4t^2 + 1$. Let $a = t^3 + t^2 + a_1 t + a_0$ for suitable $a_1, a_0 \in L$ then, because of (4), $g = t^2 - a_1 t - a_0$. (3) implies that $a_0^2 = -1$. Now $-t^5 + 3t^4 - 4t^2 + 1 = a \cdot g$ implies that $a_0 = 0$, which is a contradiction. This proves that p is irreducible, and hence the proposition is proved. □

We can now apply Corollary 1.7 to prove Proposition 2.3.

We compute the Hilbert polynomial $H(t)$ of the projective curve corresponding to I_h, the homogenisation of I. We obtain $H(t) = 10t - 11$. The corresponding Singular session is:

```
ring S=0,(b,c,t,w),dp;   ideal J=imap(R,J); ideal K=std(J); K;

K[1]=bct-t2+2t+1
K[2]=bt3-ct3+t4-b2t-c2t-2bt2+2ct2-3t3+bc+2t2-t
K[3]=b2c2-bt2+ct2-t3+b2+2bc+c2+2bt-2ct+2t2+2
K[4]=c2t3-ct4+c3t-2c2t2+3ct3-t4-bc2+bt2-2ct2+4t3-2bt+ct-3t2-b-2t
```

We now compute matrices to represent the generators of J in terms of the generators of K, and vice versa, in order to see that in any characteristic $KL[b,c,t,w] = JL[b,c,t,w]$. Moreover, using Buchberger's criterion, it is not difficult to check that K is a Gröbner basis of $IL[b,c,t,w]$ in any characteristic.

```
lift(J,K);

_[1,1]=0   _[1,2]=0   _[1,3]=t-1               _[1,4]=-1
_[2,1]=0   _[2,2]=-1  _[2,3]=-bt-t2+2b+3t-2    _[2,4]=b+c-1
_[3,1]=1   _[3,2]=0   _[3,3]=-t+3              _[3,4]=1
_[4,1]=0   _[4,2]=0   _[4,3]=-t3+bt+2t2+c      _[4,4]=-t
_[5,1]=0   _[5,2]=-1  _[5,3]=t-2               _[5,4]=c

lift(K,J);

_[1,1]=-bt3-bct+2bt2+b2-t2+2t+1  _[1,2]=-t3-ct+2t2+b
_[1,3]=1      _[1,4]=t2+c-2t     _[1,5]=t3+ct-2t2-b    _[2,1]=-b2t
_[2,2]=-bt    _[2,3]=0           _[2,4]=b              _[2,5]=bt-1
_[3,1]=0      _[3,2]=0           _[3,3]=0              _[3,4]=0
_[3,5]=0      _[4,1]=0           _[4,2]=0              _[4,3]=0
_[4,4]=0      _[4,5]=0
```

We homogenise K with respect to w and obtain again a Gröbner basis, cf. [GP3], with respect to the lexicographical ordering. Since the leading ideal is independent of the characteristic, the Hilbert polynomial is the same in any characteristic. We compute

```
K=homog(K,w); hilbPoly(K);

-11,10
```

Hence, the Hilbert polynomial is $10t - 11$. From this we obtain the degree $d = 10$ and the arithmetic genus $p_a = 12$ of the projective closure. Using Corollary 1.7, we obtain:

$$N_q \geq q + 1 - 24\sqrt{q} - 10.$$

This implies that $C(\mathbb{F}_q)$ is not empty if $q > 593$. For small q, we give a list of points (Tables 1 and 2) to prove that $C(\mathbb{F}_q)$ is not empty. Proposition 2.3 and, hence, 2.1 are proved.

Remark 2.7. Using the leading terms of J, we can even compute the Hilbert polynomial without computer. Hence (once the matrices are computed by the `lift` command and the Gröbner bases are given) we can check everything by hand, since only simple (although tedious) manipulations are necessary. Therefore, the PSL(2) case can be verified without computer.

p	point in $C(\mathbb{F}_p)$	p	point in $C(\mathbb{F}_p)$	p	point in $C(\mathbb{F}_p)$	p	point in $C(\mathbb{F}_p)$
5	$(1,2,2)$	113	$(0,37,52)$	263	$(0,47,154)$	421	$(2,331,151)$
7	$(0,1,4)$	127	$(0,10,112)$	269	$(2,205,73)$	431	$(0,100,189)$
11	$(1,9,1)$	131	$(1,14,22)$	271	$(0,64,97)$	433	$(0,67,228)$
13	$(1,1,8)$	137	$(0,5,32)$	277	$(4,21,7)$	439	$(0,4,22)$
17	$(0,7,7)$	139	$(1,19,109)$	281	$(0,98,150)$	443	$(2,213,143)$
19	$(3,2,10)$	149	$(1,87,63)$	283	$(1,188,250)$	449	$(2,215,286)$
23	$(0,11,19)$	151	$(1,99,108)$	293	$(1,26,270)$	457	$(0,63,378)$
29	$(2,12,8)$	157	$(1,22,62)$	307	$(1,100,10)$	461	$(5,5,267)$
31	$(1,18,26)$	163	$(1,67,8)$	311	$(2,56,162)$	463	$(0,62,204)$
37	$(1,25,22)$	167	$(0,3,14)$	313	$(0,45,194)$	467	$(1,70,461)$
41	$(1,4,19)$	173	$(1,101,119)$	317	$(2,34,146)$	479	$(0,202,293)$
43	$(1,15,3)$	179	$(1,11,71)$	331	$(1,197,323)$	487	$(0,9,92)$
47	$(0,2,8)$	181	$(1,3,75)$	337	$(0,138,312)$	491	$(1,31,439)$
53	$(2,16,12)$	191	$(0,7,58)$	347	$(1,252,267)$	499	$(1,275,40)$
59	$(3,33,39)$	193	$(0,45,142)$	349	$(2,314,255)$	503	$(0,12,158)$
61	$(2,21,49)$	197	$(1,18,145)$	353	$(0,142,187)$	509	$(7,424,256)$
67	$(1,11,63)$	199	$(0,67,180)$	359	$(0,80,20)$	521	$(0,219,250)$
71	$(0,18,60)$	211	$(1,51,92)$	367	$(0,28,80)$	523	$(3,8,369)$
73	$(1,44,49)$	223	$(5,6,157)$	373	$(1,82,336)$	541	$(1,220,80)$
79	$(0,17,71)$	227	$(1,118,74)$	379	$(2,9,197)$	547	$(2,264,122)$
83	$(1,54,39)$	229	$(3,220,92)$	383	$(0,149,138)$	557	$(2,42,261)$
89	$(0,19,26)$	233	$(0,19,149)$	389	$(1,27,379)$	563	$(1,317,485)$
97	$(0,10,15)$	239	$(1,179,126)$	397	$(3,271,169)$	569	$(0,269,369)$
101	$(2,1,47)$	241	$(0,67,220)$	401	$(0,48,349)$	571	$(1,443,422)$
103	$(0,23,39)$	251	$(3,15,112)$	409	$(0,50,98)$	577	$(2,169,514)$
107	$(1,61,26)$	257	$(3,97,135)$	419	$(1,121,65)$	587	$(1,45,229)$
109	$(1,69,102)$					593	$(1,240,5)$.

TABLE 1. $q = p = 5, \ldots, 593$

n	point in $C(\mathbb{F}_q)$
2	$(a, 0, 1)$
3	(a, a^2, a^2)
4	(a^3, a^{12}, a^5)
5	(a^3, a^{20}, a^{22})
6	(a^9, a^9, a^{54})
7	(a, a^{62}, a^{48})
8	(a, a^{70}, a^{200})
9	(a, a^{191}, a^{121}).

$q = 2^n, \; n = 2, \ldots, 9$

n	point in $C(\mathbb{F}_q)$
2	$(a, 0, a)$
3	(a, a^3, a^{10})
4	$(a, -1, a^{66})$
5	(a^2, a^{10}, a^2).

$q = 3^n, \; n = 2, \ldots, 5$

TABLE 2. a denotes a generator of the multiplicative group $\mathbb{F}_q \smallsetminus \{0\}$.

2.2. **The geometric structure of C.** In this subsection we study the singularities of the reduction of the curve C modulo primes. This will result in another proof of the main theorem in the PSL(2) case (Proposition 2.1). We use the notations of Subsection 2.1 and consider the curve defined by the ideal I. The difference to the proof in Subsection 2.1 is the proof of the absolute irreducibility of the polynomial

$f_1 = J[1]$, which uses here the analysis of the singularities. Furthermore, the Hasse–Weil Theorem is applied here to the normalisation of the plane curve defined by f_1, while in Subsection 2.1 it was applied to the curve defined by I and not to its projection defined by f_1.

Lemma 2.8. *With the notations of Lemma 2.6 we obtain, substituting $c = (t^2 - 2t - 1)/tb$,*

$$\begin{aligned} bJ[2]\big(t, b, (t^2 - 2t - 1)/tb\big) &= J[1] \\ J[3]\big(t, b, (t^2 - 2t - 1)/tb\big) &= 0 \\ tbJ[4]\big(t, b, (t^2 - 2t - 1)/tb\big) &= -J[1] \\ tb^2J[5]\big(t, b, (t^2 - 2t - 1)/tb\big) &= (1 - tb)J[1]\,. \end{aligned}$$

Proof. This is an easy computation. □

Corollary 2.9. *A point (t, b) of the plane curve defined by $J[1] = 0$ with $tb \neq 0$ defines a point $\big(t, b, (t^2 - 2t - 1)/tb\big)$ of the curve defined by the ideal I.*

Proof. Just note that $J[1], \ldots, J[5]$ is a Gröbner basis of I and use Lemma 2.8. □

Remark 2.10. In Subsection 2.1 we did not use this reduction to the case of a plane curve since this allowed a verification without computer. We used the Hasse–Weil theorem involving the arithmetic genus which avoids an analysis of the singularities. The arithmetic genus is 12 for the curve defined by I and 15 for its projection to the plane defined by f_1. The analysis of singularities allows us to use the geometric genus, which is 8. In principle, this does not make a big difference because we are using a computer for small fields $\mathbb{F}_q$, anyway. For genus 15, resp. 12, resp. 8, Hasse–Weil guarantees rational points if $q \geq 977$, resp. $q \geq 593$, resp. $q \geq 277$. Hence, the analysis of the singularities reduces the number of small fields which have to be treated by computer. On the other hand, when analysing the singularities, we have the disadvantage of treating the field $\mathbb{F}_{864007}$. That such a large prime will play a special role in the analysis of singularities was unexpected for us.

We reduced the problem to find a point $(t, b) \in \mathbb{F}_q$ on the plane curve $\mathbf{V}(f_1)$ with $tb \neq 0$. Note that $f_1(0, b) = b^2 + 1$ and $f_1(t, 0) = (t^2 - 2t - 1)^2$. Hence there are at most four points on the curve with $t = 0$ or $b = 0$. We shall show that there are at least five points on such a curve. We did the calculations in Singular and MAGMA to work with independent computer algebra systems.

From now on, we denote $P(t, b) = f_1(t, b)$.

We shall analyse the plane algebraic curve given by the polynomial $P(t, b)$ over various (finite) fields. We put $C, \overline{C}$ for this curve (over the complex numbers $\mathbb{C}$) and its projective closure, respectively. We use the coordinate system $(t : b : z)$ in the projective plane. The projective curve $\overline{C}$ is then given by the homogeneous polynomial

$$\begin{aligned} \overline{P}(t, b, z) = &- b^3t^4 - b^2t^5 + b^4t^2z + 2b^3t^3z + 3b^2t^4z + bt^5z - 2b^2t^3z^2 - 4bt^4z^2 + 3bt^3z^3 + t^4z^3 \\ &+ 2b^2tz^4 + 2bt^2z^4 - 4t^3z^4 + b^2z^5 + 2t^2z^5 + 4tz^6 + z^7\,. \end{aligned}$$

If p is a prime number, we put $C_p, \overline{C}_p$ for these curves over $\overline{\mathbb{F}}_p$. The curves C_p and $\overline{C}_p$ are then defined by the reductions of the polynomials $P(t, b)$ and $\overline{P}(t, b, z)$ modulo p respectively.

We use the standard formula:

$$g(C) = \frac{\big(\mathrm{degree}(\overline{C}) - 1)(\mathrm{degree}(\overline{C}) - 2\big)}{2} - \sum_{Q \in \overline{C}(\bar{k})} \delta_Q$$

where the local contributions of singular points δ_Q are defined as $\dim_{\bar{k}} \widetilde{\mathcal{O}}_Q/\mathcal{O}_Q$; here $\mathcal{O}_Q$ is the local ring of Q and $\widetilde{\mathcal{O}}_Q$ is the integral closure of $\mathcal{O}_Q$ in the function field of $\overline{C}$.

We also define

$$A(\mathbb{F}_q) = \#\{\, (t,b) \in \mathbb{F}_q^2 \;\big|\; P(t,b) = 0 \,\} = \#C_p(\mathbb{F}_q)$$

for a prime power q. The following tables contain the solution numbers $A(L)$ for various finite fields L.

$A(\mathbb{F}_2) = 2$	$A(\mathbb{F}_{2^7}) = 128$	$A(\mathbb{F}_{2^{13}}) = 7880$	$A(\mathbb{F}_{3^5}) = 190$
$A(\mathbb{F}_{2^2}) = 6$	$A(\mathbb{F}_{2^8}) = 218$	$A(\mathbb{F}_{2^{14}}) = 16722$	$A(\mathbb{F}_{3^6}) = 734$
$A(\mathbb{F}_{2^3}) = 11$	$A(\mathbb{F}_{2^9}) = 551$	$A(\mathbb{F}_3) = 0$	$A(\mathbb{F}_{3^7}) = 2380$
$A(\mathbb{F}_{2^4}) = 10$	$A(\mathbb{F}_{2^{10}}) = 1026$	$A(\mathbb{F}_{3^2}) = 14$	$A(\mathbb{F}_{3^8}) = 6806$
$A(\mathbb{F}_{2^5}) = 32$	$A(\mathbb{F}_{2^{11}}) = 2048$	$A(\mathbb{F}_{3^3}) = 36$	
$A(\mathbb{F}_{2^6}) = 39$	$A(\mathbb{F}_{2^{12}}) = 4279$	$A(\mathbb{F}_{3^4}) = 78$	

TABLE 3. Number of points on $C(\mathbb{F}_q)$, $q = 2^n$ or 3^n

The numbers contained in Table 3 and also those in Table 4 can be obtained in microseconds on a computer. We have, in fact, used MAGMA and verified this with SINGULAR.

$A(\mathbb{F}_5) = 11$	$A(\mathbb{F}_{61}) = 72$	$A(\mathbb{F}_{131}) = 121$	$A(\mathbb{F}_{199}) = 167$
$A(\mathbb{F}_7) = 5$	$A(\mathbb{F}_{67}) = 57$	$A(\mathbb{F}_{137}) = 121$	$A(\mathbb{F}_{211}) = 229$
$A(\mathbb{F}_{11}) = 8$	$A(\mathbb{F}_{71}) = 76$	$A(\mathbb{F}_{139}) = 134$	$A(\mathbb{F}_{223}) = 203$
$A(\mathbb{F}_{13}) = 16$	$A(\mathbb{F}_{73}) = 78$	$A(\mathbb{F}_{149}) = 140$	$A(\mathbb{F}_{227}) = 230$
$A(\mathbb{F}_{17}) = 19$	$A(\mathbb{F}_{79}) = 89$	$A(\mathbb{F}_{151}) = 164$	$A(\mathbb{F}_{229}) = 220$
$A(\mathbb{F}_{19}) = 15$	$A(\mathbb{F}_{83}) = 76$	$A(\mathbb{F}_{157}) = 161$	$A(\mathbb{F}_{233}) = 250$
$A(\mathbb{F}_{23}) = 9$	$A(\mathbb{F}_{89}) = 82$	$A(\mathbb{F}_{163}) = 170$	$A(\mathbb{F}_{239}) = 272$
$A(\mathbb{F}_{29}) = 45$	$A(\mathbb{F}_{97}) = 92$	$A(\mathbb{F}_{167}) = 136$	$A(\mathbb{F}_{241}) = 277$
$A(\mathbb{F}_{31}) = 33$	$A(\mathbb{F}_{101}) = 98$	$A(\mathbb{F}_{173}) = 167$	$A(\mathbb{F}_{251}) = 233$
$A(\mathbb{F}_{37}) = 36$	$A(\mathbb{F}_{103}) = 97$	$A(\mathbb{F}_{179}) = 128$	$A(\mathbb{F}_{257}) = 271$
$A(\mathbb{F}_{41}) = 61$	$A(\mathbb{F}_{107}) = 98$	$A(\mathbb{F}_{181}) = 210$	$A(\mathbb{F}_{263}) = 246$
$A(\mathbb{F}_{43}) = 32$	$A(\mathbb{F}_{109}) = 121$	$A(\mathbb{F}_{191}) = 233$	$A(\mathbb{F}_{269}) = 305$
$A(\mathbb{F}_{47}) = 42$	$A(\mathbb{F}_{113}) = 122$	$A(\mathbb{F}_{193}) = 223$	$A(\mathbb{F}_{271}) = 240$
$A(\mathbb{F}_{53}) = 53$	$A(\mathbb{F}_{127}) = 136$	$A(\mathbb{F}_{197}) = 201$	$A(\mathbb{F}_{277}) = 263$
$A(\mathbb{F}_{59}) = 36$			

TABLE 4. Number of points on $C(\mathbb{F}_p)$, p prime

We add the information:

$$(*) \qquad A(\mathbb{F}_{523}) = 474, \qquad A(\mathbb{F}_{864007}) = 864867$$

which can also be obtained by a simple computer calculation.

We shall show:

Proposition 2.11. *If $q = p^k$ for a prime p and $q \neq 2, 3$ then $A(\mathbb{F}_q) \geq 5$.*

Our theorems would be much easier to prove if a rational point on C could be found. Unfortunately, even an extensive computer search has not revealed such a point.

We proceed by our analysis of the curve C. Consider affine charts $C^1 = \mathbf{D}(z) \cap \overline{C}$, $C^2 = \mathbf{D}(b) \cap \overline{C}$, and $C^3 = \mathbf{D}(t) \cap \overline{C}$. The part at infinity $\mathbf{V}(z) \cap \overline{C}$ is denoted by C^∞. Putting a prime p as an index to C stands then for the analogous construction over $\overline{\mathbb{F}}_p$. We have:

Lemma 2.12. *The part at infinity $C^\infty(\mathbb{C})$ consists exactly of the points $(0:1:0)$, $(1:0:0)$, $(1:1:0)$. Also $C_p^\infty(\overline{\mathbb{F}}_p)$ consists exactly of the points $(0:1:0)$, $(1:0:0)$, $(1:1:0)$ for every prime p.*

Proof. We find

$$\overline{P}(t,b,0) = -b^2t^4(b+t)$$

and the statement follows. □

We shall later prove Proposition 2.11 by an application of a Hasse–Weil estimate for the number of points on $\overline{C}_p$. To do this, we have to understand the singularities of $\overline{C}_p$ and also prove the absolute irreducibility as the prime p varies. The following contains a description of the singularities of $\overline{C}$.

Lemma 2.13. *The projective curve $\overline{C}$ has the 4 singular points*

$$Q_1 = (\omega+1:0:1),\ Q_2 = (-\omega+1:0:1),\ Q_3 = (1:0:0),\ Q_4 = (0:1:0)$$

where $\omega = \sqrt{2}$. The points Q_1, Q_2, Q_3 are ordinary double points whereas Q_4 is a singularity of type D_6, that is Q_4 is a triple point with 3 branches, two of which are simply tangent. The projective curve $\overline{C}_{\mathbb{Q}}$ is absolutely irreducible and $g(\overline{C}_{\mathbb{Q}}) = 8$.

Proof. Most of this statement is computed by MAGMA and Singular, the absolute irreducibility follows from Bezout. We shall not carry this out here since we shall give the same argument over the finite fields $\mathbb{F}_p$ later. □

From general theory it is clear that Lemma 2.13 also holds for the curve $\overline{C}_p$ for almost all primes p. To get later explicit estimates, we have to find the exceptional set of primes. We put:

$$S = \{\,2,\ 23,\ 37,\ 523,\ 864007\,\}$$

and prove:

Proposition 2.14. *Let p be a prime with $p \notin S$. Then the projective curve $\overline{C}_p$ has the 4 singular points*

$$Q_1 = (\omega+1:0:1),\ Q_2 = (-\omega+1:0:1),\ Q_3 = (1:0:0),\ Q_4 = (0:1:0)$$

where ω is a root of x^2-2 in $\overline{\mathbb{F}}_p$. The points Q_1, Q_2, Q_3 are ordinary double points whereas Q_4 is a triple point with 3 branches, two of which meet in Q_4 of order 2 and the third intersects them transversally (D_6-configuration). The projective curve $\overline{C}_p$ is absolutely irreducible and $g(\overline{C}_p) = 8$.

Proof. We shall first find the singularities of $\overline{C}_p$. The description of the singularities is obtained by looking at the blow ups of $\overline{C}_p$ in the four singular points. These can be computed by Singular or MAGMA.

We shall now analyse the singularities on the first affine patch C_p^1. Let $\mathfrak{a}_1$ be the ideal in $\mathbb{Z}[t,b]$ generated by P and its derivatives with respect to t, b. A Gröbner basis computation over $\mathbb{Z}$ carried out in Singular or MAGMA shows that $sb \in \mathfrak{a}_1$ where

$$s = 35378249251012 = 4\cdot 23^2\cdot 37\cdot 523\cdot 864007.$$

We have $\langle \mathfrak{a}_1, b\rangle = \langle b, t^2-2t-1\rangle$. This shows that the affine patch C_p^1 contains only the (distinct) singular points Q_1, Q_2.

Let $\mathcal{O}$ be the ring of integers in $\mathbb{Q}[\sqrt{2}]$. Note that $\mathcal{O} = \mathbb{Z}[\sqrt{2}]$. The points Q_1, Q_2 have their coordinates in $\mathcal{O}/\mathfrak{p}$ where $\mathfrak{p}$ is a prime ideal of $\mathcal{O}$ containing p. The polynomial $P(v+\sqrt{2}+1, b)$ has $H_2(b,v) = -(\sqrt{2}+1)b^2+2(\sqrt{2}+2)bv+8v^2$ as its homogeneous part of lowest degree. A simple computation shows that the only prime ideals $\mathfrak{p}$ of $\mathcal{O}$ with the property that $H_2(b,v)$ is a square modulo $\mathfrak{p}$ are $\mathfrak{p}_1 = \sqrt{2}\mathcal{O}$ and $\mathfrak{p}_2 = (-3+4\sqrt{2})\mathcal{O}$. Note that $\mathfrak{p}_2$ contains 23. The point Q_2 is analysed similarly.

This shows that for $p \notin S$ the affine patch C_p^1 only contains the ordinary double points Q_1, Q_2 as singularities. Note that $\delta_{Q_1} = \delta_{Q_2} = 1$.

We shall now analyse the singularities on the second affine patch C_p^2. Put $P_2(t,z) = \overline{P}(t,1,z)$. Let $\mathfrak{a}_2$ be the ideal in $\mathbb{Z}[t,z]$ generated by P_2 and its derivatives with respect to t, z. A Gröbner basis computation over $\mathbb{Z}$ carried out in MAGMA shows that $s_2 b \in \mathfrak{a}_2$ where

$$s_2 = 66877597828 = 4 \cdot 37 \cdot 523 \cdot 864007.$$

We have $\langle \mathfrak{a}_2, z\rangle = \langle z, t^2\rangle$. This shows that this affine patch contains only the singular point Q_4. The polynomial $P_2(t,z)$ has t^2z as its homogeneous part of lowest degree, hence Q_4 is a triple point. Let C be the affine curve over $\overline{\mathbb{F}}_p$ given by P_2 and C_1 be the curve given by the polynomial $T_1(t,z) = P_2(t,zt)/t^3$. The polynomial T_1 has $t+z$ as its homogeneous part of lowest degree. This shows that the blown up curve C_1 has only a simple point lying over $(0,0) \in C$. Let C_2 be the curve given by the polynomial $T_2(t,z) = P_2(tz,z)/z^3$. The polynomial T_2 has t^2+z^2 as its degree 2 homogeneous part. This shows that the blown up curve C_2 has an ordinary double point lying over $(0,0) \in C$.

This shows that 3 branches meet in Q_4. Two of them intersect of order 2 and the third intersects these transversally. By M. Noether's formula (δ_{Q_4} equals the sum of $m_Q(m_Q-1)/2$ where Q runs over all points in all blow-ups lying over Q_4 and m_Q is the multiplicity of Q) we find $\delta_{Q_4} = 4$.

We shall now analyse the singularities on the third affine patch C_p^3. Put $P_3(b,z) = \overline{P}(1,b,z)$. Let $\mathfrak{a}_3$ be the ideal in $\mathbb{Z}[b,z]$ generated by P_3 and its derivatives with respect to b, z. A Gröbner basis computation over $\mathbb{Z}$ carried out in MAGMA shows that $sb \in \mathfrak{a}_3$. We have $\langle \mathfrak{a}_3, b\rangle = \langle b, z(z^2+2z-1)\rangle$. This shows that Q_3 is the only singular point on this patch which was not found on the previous affine patches. The polynomial $P_3(b,z)$ has $-(b+z)z$ as its homogeneous part of lowest degree.

This shows that Q_3 is an ordinary double point and $\delta_{Q_3} = 1$.

So far we have described the singularities of $\overline{C}_p$.h The degree of C_p being also 7, we find $g(\overline{C}_p) = 15-1-1-1-4 = 8$.

It remains to prove the absolute irreducibility of $\overline{C}_p$. Suppose $\overline{C}_p$ had 2 components C_1, C_2. From the description of the singularities we infer the following possibilities for the intersection numbers:

$$I(C_1,C_2;Q_1) = 0,\,1, \quad I(C_1,C_2;Q_2) = 0,1, \quad I(C_1,C_2;Q_3) = 0,\,1, \quad I(C_1,C_2;Q_4) = 0,\,2,\,3\,.$$

Note that $Q_1,\dots,Q_4$ do not lie on a common line. The degree of $\overline{C}_p$ being 7, Bezout's theorem shows that $\overline{C}_p$ is absolutely irreducible. □

Although we shall not need all of it, we shall also describe the situation for the exceptional primes p in S. We start with $p=2$.

Proposition 2.15. *The projective curve $\overline{C}_2$ has the 4 singular points*

$$Q_1 = (1:0:1), \quad Q_2 = (0:1:1), \quad Q_3 = (1:0:0), \quad Q_4 = (0:1:0)\,.$$

The points Q_1, Q_2, Q_3 are double points whereas Q_4 is a triple point. The point Q_3 is ordinary, at Q_1 two branches with a common tangent touch of order 2, Q_2 is an ordinary cusp, at Q_4 two branches with distinct tangents meet, one of them behaves like a third order cusp, the other is smooth in Q_4 (a D_9-configuration). The projective curve $\overline{C}_2$ is absolutely irreducible and $g(\overline{C}_2) = 6$.

Proof. We shall first find the singularities of $\overline{C}_2$. The description of the singularities is obtained by looking at the blow ups of $\overline{C}_2$ in the four singular points. These can be computed by MAGMA or SINGULAR.

We shall now analyse the singularities on the first affine patch C_2^1 ($z=1$). The Jacobian ideal of $P(t,b)$ is generated by $b^2(b^2+1)$ and t^2+b^2+1. This shows that Q_1, Q_2 are the only singularities on this affine patch.

Put $L_1(t,b) = P(t+1,b)$. The polynomial $L_1(t,b)$ has b^2 as homogeneous component of lowest degree. Let C be the affine curve over $\overline{\mathbb{F}}_2$ given by L_1 and C_1 be the curve given by the polynomial $T_1(t,b) = L_1(tb,b)/b^2$. A look at T_1 shows that there is no point of C_1 lying over $(0,0) \in C$. Let C_2 be the curve given by the polynomial $T_2(t,b) = L_1(t,tb)/t^2$. The polynomial has $b(b+t)$ as its homogeneous component of lowest degree. This shows that there is an ordinary double point over $(0,0) \in C$ on C_2. Altogether we find that two branches with a common tangent touch of order 2 in Q_1. This implies $\delta_{Q_1} = 2$.

Put $L_2(t,b) = P(t,b+1)$. The polynomial $L_2(t,b)$ has $(t+b)^2$ as homogeneous component of lowest degree. Let C be the affine curve over $\overline{\mathbb{F}}_2$ given by L_2. Both blow-ups of $(0,0) \in C$ contain (the same) smooth point over $(0,0) \in C$. This shows that Q_2 is a cusp (one branch passing through Q_2) and $\delta_{Q_1} = 1$.

We shall now analyse the singularities on the second affine patch C_2^2 $(b=1)$. The points Q_2 and Q_4 are the only singularities on this affine patch. To analyse Q_4, put $P_2(t,z) = \overline{P}(t,1,z)$. The polynomial P_2 has t^2z as its homogeneous component of lowest degree. Hence Q_4 is a triple point with two distinct tangents. Let C be the affine curve over $\overline{\mathbb{F}}_2$ given by P_2. In the first blow-up $(z = zt)$ we find a simple point over $(0,0) \in C$, In the second blow-up we find a point Q_5 of multiplicity 2 with a double tangent over $(0,0) \in C$. The blow-ups of Q_5 give one double point with a double tangent Q_6. The blow-ups of Q_6 give one simple point Q_7. This shows that at Q_4 two branches with distinct tangents meet, one of them behaves like a third order cusp, the other is smooth in Q_4. By M. Noether's formula we find $\delta_{Q_4} = 3+1+1 = 5$.

We shall now analyse the singularities on the third affine patch C_2^3 $(t=1)$. The points Q_1 and Q_3 are the only singularities on this affine patch. To analyse Q_3 put $P_3(b,z) = \overline{P}(1,b,z)$. The polynomial P_3 has bz as its homogeneous component of lowest degree. This shows that Q_3 is an ordinary double point and $\delta_{Q_3} = 1$.

The analysis of the singularities being completed, we have found $g(\overline{C}_2) = 15-2-1-1-5 = 6$.

Suppose $\overline{C}_2$ had 2 components C_1, C_2. From the description of the singularities we infer the following possibilities for the intersection numbers:

$$I(C_1,C_2;Q_1) = 0,\ 2, \quad I(C_1,C_2;Q_2) = 0, \quad I(C_1,C_2;Q_3) = 0,\ 1, \quad I(C_1,C_2;Q_4) = 0,\ 2\,.$$

These numbers cannot add up to 6 or more. The degree of $\overline{C}_{\mathbb{F}_2}$ being 7, Bezout's theorem shows that $\overline{C}_{\mathbb{F}_2}$ is absolutely irreducible. □

Proposition 2.16. *The projective curve $\overline{C}_{23}$ has the 4 singular points*

$$Q_1 = (19:0:1), \quad Q_2 = (6:0:1), \quad Q_3 = (1:0:0), \quad Q_4 = (0:1:0)\,.$$

The points Q_2, Q_3 are ordinary double points, Q_1 is an ordinary cusp, whereas Q_4 is a triple point with 3 branches, two of which meet in Q_4 of order 2 and the third intersects them transversally (D_6-configuration). Q_1 is a cusp singularity. The projective curve $\overline{C}_{23}$ is absolutely irreducible and $g(\overline{C}_{23}) = 8$.

Proof. The singular points and their types were computed by MAGMA. To complete the Bezout argument notice that Q_4 does not lie on a line with at the three double points. □

Proposition 2.17. *The projective curve $\overline{C}_{37}$ has the 8 singular points*

$$\begin{aligned} &Q_1 = (\omega+1:0:1), && Q_2 = (-\omega+1:0:1), && Q_3 = (1:0:0), && Q_4 = (0:1:0), \\ &Q_5 = (27:17:1), && Q_6 = (10:10:1), && Q_7 = (10:24:1), && Q_8 = (27:34:1)\,, \end{aligned}$$

where ω is a root of x^2-2 in $\overline{\mathbb{F}}_{37}$. The points Q_1, Q_2, Q_3, Q_5, Q_6, Q_7, Q_8 are ordinary double points whereas Q_4 is a triple point with 3 branches, two of which meet in Q_4 of order 2 and the third intersects them transversally (D_6-configuration). The projective curve $\overline{C}_{37}$ is absolutely irreducible and $g(\overline{C}_{37}) = 4$.

Proof. The singular points and their types were computed by MAGMA. To complete the Bezout argument notice that Q_4 does not lie on a line with at least three of the double points, and also that the points Q_1,…,Q_8 do not lie on a quadric. □

Proposition 2.18. *The projective curve $\overline{C}_{523}$ has the 5 singular points*

$$Q_1 = (\omega + 1 : 0 : 1), \quad Q_2 = (-\omega + 1 : 0 : 1), \quad Q_3 = (1 : 0 : 0), \quad Q_4 = (0 : 1 : 0), \quad Q_5 = (479 : 463 : 1)\,,$$

where ω is a root of $x^2 - 2$ in $\overline{\mathbb{F}}_{523}$. The points Q_1, Q_2, Q_3, Q_5 are ordinary double points whereas Q_4 is a triple point with 3 branches, two of which meet in Q_4 of order 2 and the third intersects them transversally (D_6-configuration). The projective curve $\overline{C}_{523}$ is absolutely irreducible and $g(\overline{C}_{523}) = 7$.

Proof. The singular points and their types were computed by MAGMA. To complete the Bezout argument notice that Q_4 does not lie on a line with at least three of the double points. $\square$

Proposition 2.19. *The projective curve $\overline{C}_{864007}$ has the 5 singular points*

$$\begin{array}{lll} Q_1 = (767405 : 0 : 1), & Q_2 = (96604 : 0 : 1), & Q_3 = (1 : 0 : 0), \\ Q_4 = (0 : 1 : 0), & Q_5 = (395579 : 564628 : 1)\,. & \end{array}$$

The points Q_1, Q_2, Q_3, Q_5 are ordinary double points whereas Q_4 is a triple point with 3 branches, two of which meet in Q_4 of order 2 and the third intersects them transversally (D_6-configuration). The projective curve $\overline{C}_{864007}$ is absolutely irreducible and $g(\overline{C}_{864007}) = 7$.

Proof. The singular points and their types were computed by MAGMA. To complete the Bezout argument notice that Q_4 does not lie on a line with at least three of the double points. $\square$

We are now ready for the

Proof of Proposition 2.11: We first assume that the prime p satisfies $p \notin S$ and also $p \neq 3$. We shall then show that the statement of Proposition 2.11 is already true for $q = p$. We have to show that $\#C_p(\mathbb{F}_p) \geq 5$, which is by Lemma 2.12 equivalent to $\#\overline{C}_p(\mathbb{F}_p) > 7$. We write $\mathcal{D}_p$ for a nonsingular model of $\overline{C}_p$ and $\pi_p \colon \mathcal{D}_p \to \overline{C}_p$ for the birational projection. The map π_p is defined over $\mathbb{F}_p$. Let $M \subset \overline{C}_p(\mathbb{F}_p)$ be the set of singular points. The map π_p defines a bijection

$$\pi_p : \mathcal{D}_p(\mathbb{F}_p) \smallsetminus \pi_p^{-1}(M) \to \overline{C}_p(\mathbb{F}_p) \smallsetminus M\,.$$

Since the singularities of $\overline{C}_p$ are three double and a triple point, we find:

$$\#C_p(\mathbb{F}_p) \geq \#\mathcal{D}_p(\mathbb{F}_p) - 5\,.$$

Hence, it is sufficient to show that $\#\mathcal{D}_p(\mathbb{F}_p) > 12$. By the Hasse–Weil estimate we know that $p - 16\sqrt{p} \leq \#\mathcal{D}_p(\mathbb{F}_p)$. If $p \geq 280$, we infer $\#\mathcal{D}_p(\mathbb{F}_p) > 12$. If $p \leq 280$, we use Table 4.

For the primes $p \in S$, $p \neq 2, 3$ we also have $\#C_p(\mathbb{F}_p) \geq 5$ by Table 4 or the addition (*). For $q = 2^k$, $k \geq 2$, we use Proposition 2.15 and an argument similar to the above to show $A(\mathbb{F}_{2^k}) \geq 5$ for $k \geq 8$. The remaining values can be found in Table 3. For $q = 3^k$, $k \geq 2$, we use Proposition 2.14 and an argument similar to the above to show $A(\mathbb{F}_{3^k}) \geq 5$ for $k \geq 6$. The remaining values can be found in Table 3.

3. The details of the Suzuki case

3.1. **The variety V and the Suzuki groups.** We shall explain here in more detail the relationship of the variety V constructed in Subsection 1.3 to the Suzuki groups. We use the following representation for $\mathrm{Sz}(q)$. Let $n = 2m+1$ and $q = 2^n$ and consider the automorphism $\theta \colon \mathbb{F}_q \longrightarrow \mathbb{F}_q$, $\theta(a) = a^{2^{m+1}}$. We have $\theta^2(a) = a^2$, that is, π is the square root of the Frobenius.

Let

$$U(a,b) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ a\theta(a) + b & \theta(a) & 1 & 0 \\ a^2\theta(a) + ab + \theta(b) & b & a & 1 \end{pmatrix}, \tag{3.1}$$

$$M(c) = \begin{pmatrix} c^{1+2^m} & 0 & 0 & 0 \\ 0 & c^{2^m} & 0 & 0 \\ 0 & 0 & c^{-2^m} & 0 \\ 0 & 0 & 0 & c^{-1-2^m} \end{pmatrix}, \qquad T = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Then $\mathrm{Sz}(q) = \langle U(a,b), M(c), T \mid a,b,c \in \mathbb{F}_q, c \neq 0\rangle \subset \mathrm{SL}(4,\mathbb{F}_q)$.

To show that $u_1(x,y) = u_2(x,y)$ has a solution with $y \neq x^{-1}$, we consider the matrices $X = TU(a,b)$ and $Y = TU(c,d)$ in $\mathrm{Sz}(q)$. It is easy to see that $Y = X^{-1}$ in $Sz(q)$ if and only if a,b,c,d are all 0.

To eliminate the dependence of X and Y on q, we replace $\theta(a),\dots,\theta(d)$ with $a_0,\dots,d_0$ which we regard as indeterminates, along with $a,\dots,d$. We thus arrive to the matrices $x,y \in \mathbf{GL}(4,R)$ defined in (1.4), where $R = \mathbb{F}_2[a,\dots,d,a_0,\dots,d_0]$ is the polynomial ring in 8 variables.

Using the definition of the ideal $\mathfrak{a} \subset R$ and the variety V, (see Subsection 1.3), we can easily produce 16 generators of $\mathfrak{a}$ and prove that $\pi(\mathfrak{a}) = \mathfrak{a}$ and $\dim(V) = 2$ (a Singular computation). Hence V is preserved by the operator α, and we have

Proposition 3.1. *The matrices corresponding to a fixed point of α^n (n odd and $n \geq 3$) lie in* $\mathrm{Sz}(2^n)$.

Proof. Let $p = (a,\dots,d_0) \in V$ be a fixed point of α^{2m+1}. We have $a = a_0^{2^{m+1}}$, $a_0 = a^{2^{m+1}}$ (and hence $a = a^{2^n}$, $a_0 = \theta(a)$), and the same formulas hold for b,c,d,b_0,c_0,d_0. To finish the proof, it remains to combine this with formulas (1.4) and (3.1). ☐

To sum up, we obtained the following reduction.

Theorem 3.2. *Suppose that for every odd $n > 1$ the operator α^n has a non-zero $\mathbb{F}_q$-rational fixed point on the variety V. Then the equation $u_1 = u_2$ has a non-trivial solution in* $\mathrm{Sz}(q)$ *for every $q = 2^n$.*

3.2. **The geometric structure of V.** In this subsection we study the 2-dimensional component V' of V. We prove that it is absolutely irreducible and find a smooth open affine subset $U \subset V'$ invariant under α.

In order to facilitate computer output, we slightly change notation: we denote $a_0 = v$, $b_0 = w$, $c_0 = x$, $d_0 = y$; we will replace $\mathfrak{a}$ by I, and $\mathfrak{a}'$ by J. Singular gives 16 polynomials generating the ideal I:

```
ring A=2,(a,b,c,d,v,w,x,y),dp;

matrix S1[4][4] =1            0, 0, 0,
                 a,           1, 0, 0,
                 av+b,        v, 1, 0,
                 a2v+a*b+w,   b, a, 1;

matrix S2[4][4] =1,           0, 0, 0,
                 c,           1, 0, 0,
                 cx+d,        x, 1, 0,
                 c2x+c*d+y,   d, c, 1;

matrix T[4][4] = 0, 0, 0, 1,
                 0, 0, 1, 0,
                 0, 1, 0, 0,
                 1, 0, 0, 0;

matrix X=T*S1; matrix Y=T*S2;
matrix iX = inverse(X); matrix iY = inverse(Y);
```

```
matrix M=iX*Y*iX*iY*X*X-Y*iX*iX*iY*X*iY;
ideal I=flatten(M); I;
```

(we dispense with the output).

To show that α^n has a rational fixed point on V, we want to apply the Lefschetz trace formula, which requires that V is absolutely irreducible. This is not the case. Therefore, we exhibit a subvariety $V' \subset V$ for which we can show that it is absolutely irreducible. Then we apply the Lefschetz trace formula to the non-singular locus of V' which happens to be affine.

Set $J = I : a^3x^2$, then $J \supset I$ and $V' := \mathbf{V}(J) \subset \mathbf{V}(I) = V$. Computing $I : a^3x^2$ is not an easy task. However, once J is given, it is much simpler to check $J \supset I$, which is all we need. The ideal J is computed as follows:

```
ideal J=quotient(I,a3x2); J;

J[1]=d2+adv+cdv+a2v2+c2v2+abx+bcx+wx+c2x2+vy+xy+c2;
J[2]=a2b+acd+a2cv+aw+a3x+a2cx+ac2x+ay+av+cx;
J[3]=bcw+acvw+w2+a2wx+acwx+b2+bd+d2+abv+bcv+c2v2+bcx+adx+a4
     +a3c+vx+x2+ac+1;
J[4]=adv2+cdv2+d2x+abvx+bcvx+advx+cdvx+vwx+abx2+bcx2+wx2
     +c2x3+v2y+vxy+x2y+ab+cd+acv+c2v+w+a2x+acx+c2x+y;
J[5]=abd+abcv+bc2v+a2dv+dw+avw+cvw+bc2x+c2dx+ac2vx+awx
     +a2cx2+ac2x2+c3x2+by+cxy+dv+av2+cv2+bx+cx2+ac2+a+c;
J[6]=bcd+cd2+a2bv+abcv+a2dv+c2dv+bw+avw+cvw+a2dx+c2dx+c3vx
     +a3x2+a2cx2+ac2x2+by+dy+cvy+axy+bv+dv+cv2+dx+cvx+ax2+
     a3+a+c;
J[7]=a3v2+a2cv2+c2dx+a3vx+ac2vx+a2cx2+ac2x2+c3x2+cxy+cx2;
J[8]=d2v+acv3+c2v3+cdvx+a2vx2+acvx2+a2bc+ac2d+ac3v+acw+a3cx
     +vx2+acy+a2v+acx+v;
J[9]=advx+cdvx+a2v2x+c2v2x+abx2+bcx2+a2vx2+c2vx2+wx2+vxy+c3d
     +a3cv+a2c2v+a3cx+a2c2x+c4x+c2y+cd+a2v+c2v+c2x+y;
J[10]=a2vw+acvw+c2vw+w2+ac2dx+c3dx+a3cvx+ac3vx+acwx+c2wx
     +a3cx2+c4x2+aby+acxy+c2xy+a2v2+acv2+abx+adx+cdx+a2vx
     +acvx+c2vx+a2x2+c2x2+a4+a2c2+v2+1;
```

It turns out to be sufficient, and easier, to work temporarily with generic a and c, that is, to work over $\mathbb{F}_2(a,c)[w,y,b,d,x,y]$, where $\mathbb{F}_2(a,c)$ denotes the field of fractions of $\mathbb{F}_2[a,c]$. Fortunately, SINGULAR allows the computation of Gröbner bases over such rings. We compute a Gröbner basis of the ideal $J3 = J\mathbb{F}_2(a,c)[w,y,b,d,x,v]$ with respect to the lexicographical ordering.

```
ring s=(2,a,c),(w,y,b,d,x,v),lp;
ideal J3=std(imap(r,J));J3;
J3[1]=(a8+a6c2+a4c4+a2c6)*v6+(a8+a7c3+a6c2+a5c3+a4c4+a3c7+a2c6
      +ac7)*v4+(a7c3+a6c2+a5c5+a5c3+a3c7+a3c5+a2c6+a2c4+ac9+c6)
      *v2+(ac9+ac5+c8+c4);
J3[2]=(a4c4+a3c7+a3c5+a3c3+a2c8+a2c4+ac7+c4)*x+(a8+a7c+a4c4
      +a3c5)*v5+(a8+a7c+a6c2+a5c3+a4c4+a4c2+a2c6+a2c4)*v3
      +(a4c4+a4c2+a3c7+a3c3+a2c8+a2c6)*v;
J3[3]=(c2+1)*d2+(xc3+xc)*d+(v3xa2+v3xc2+v2a4+v2a3c+v2ac3+v2c4
      +vxa4+vxa3c+vxa2c2+vxc4+x2a2c2+x2a2+x2ac3+x2c2+c4+c2);
J3[4]=(ac5+ac)*b+(v4a2c2+v4a2+v3xac+v2x2c4+v2x2c2+v2a5c+v2a4
      +v2a2c4+v2ac3+vx3ac+vx3+vxa5c+vxa4+vxa3c3+vxa2c4+vxa2c2
      +vxac5+vxac3+vxac+vxc2+vx+x2a3c3+x2a2c2+x2ac5+x2c4+a2c4
```

```
      +a2c2+ac3+ac+c4+c2)*d+(v2xa2c3+v2xc5+vx2a3+vx2ac2+va5c2
      +va4c+va3+va2c+vac6+vac4+vac2+vc5+xa5c2+xa4c+xa2c3+xa2c
      +xac6+xac2+xc3);
J3[5]=(c)*y+(va2c+va)*bd+(v2ac+v2c2+x4+c4+1)*b+(v4a3c+v4ac
      +v3xc2+v2x2ac3+v2x2ac+v2a4+v2a3c3+v2ac3+v2c2+vx3c2+vxa4
      +vxa3c3+vxa3c+vxa2c4+vxa2+vxac3+vxac+x2a3c+x2a2c4+x2a2
      +x2ac+x2+ac3+ac+c2+1)*d+(v3x2ac2+v3x2c3+v3a3c4+v3a3c2
      +v3a2c3+v3a2c+v3ac4+v3c3+v2xa3c2+v2xa3+v2xa2c5+v2xc5
      +v2xc+vx4a3+vx4a2c+vx4a+vx4c+vx2a7+vx2a5+vx2a3c2+vx2a2c3
      +vx2a+vx2c3+vx2c+va7c2+va7+va4c+va3c6+va3c4+va3c2+va2c5
      +va2c+vac6+vac4+vc3+vc+x5a+x5c+x3a6c+x3a5+x3a4c+x3a3
      +x3a2c+x3ac2+x3c+xa7c2+xa6c+xa5c2+xa4c3+xa3c6+xa3c4
      +xa3c2+xa3+xa2c5+xa2c3+xa2c+xac6+xac4+xa);
J3[6]=w+(vx+1)*y+(a)*b+(vxc+c)*d+(v3a2+v3ac+v2xa2+v2xc2+vx2a2
      +vx2ac+vx2c2+vx2+va2c2+va2+vac3+v+xa2c2+xa2+xac3+xc2);
```

`dim(J3);` returns 0, hence V' is a surface.

Let $f = (a^3 + a^2c^3 + a^2c + ac^4 + ac^2 + c)(ac+1)(a+c)(c+1)ac$ be the least common multiple of the leading coefficients of this Gröbner basis. Then, using Singular[2], we obtain

$$J3 \cap \mathbb{F}_2[a,c,w,y,b,d,x,v] = \langle J3[1],\ldots,J3[6]\rangle : f^\infty = \langle J3[1],\ldots,J3[6]\rangle : f^6 = J.$$

Since $J : f = J$, no factor of f divides all elements of J. That is why the irreducibility of $J3$ as an ideal of $\mathbb{F}_2(a,c)[w,y,b,d,x,v]$ implies the irreducibility of J.

Furthermore, we compute the vector space dimension over $\mathbb{F}_2(a,c)$ as

$$\dim_{\mathbb{F}_2(a,c)} \mathbb{F}_2(a,c)[w,y,b,d,x,v]/J3 = 12\,.$$

Next we show that $J3 \cap \mathbb{F}_2(a,c)[b] = \langle h\rangle$ with the following polynomial h, which we compute directly by elimination (using Singular).

```
poly h=(a18c2+a16+a14c6+a12c4+a10c10+a8c8+a6c14+a4c12)*b12
+(a20c2+a19c5+a18+a17c7+a17c5+a17c3+a16c6+a15c7+a15c5+a15c3
+a14c4+a13c5+a12c10+a11c13+a10c8+a9c15+a9c13+a9c11+a8c14
+a7c15+a7c13+a7c11+a6c12+a5c13)*b10+(a21c5+a20c4+a19c5+a19c3
+a18c2+a17c9+a17c3+a16c6+a16+a15c7+a14c6+a14c4+a14c2+a13c13
+a12c12+a11c13+a11c11+a10c10+a10c6+a9c17+a9c11+a8c14+a8c8
+a7c15+a6c14+a6c12+a6c10+a2c14)*b8+(a24c2+a22c4+a22+a18c6
+a18c4+a17c11+a17c3+a16c8+a15c13+a15c9+a15c7+a15c5+a14c10
+a14c8+a13c15+a13c11+a13c9+a13c7+a12c14+a12c12+a12c8+a11c17
+a11c13+a11c11+a11c5+a10c16+a10c12+a10c10+a10c4+a9c15+a9c13
+a9c11+a9c9+a8c12+a8c6+a7c13+a7c11+a6c14+a6c8+a5c15+a5c13
+a4c10+a3c15+a3c13)*b6+(a26c2+a25c5+a24c4+a24+a23c5+a23c3
+a22c2+a21c9+a21c5+a21c3+a20c8+a20c6+a20+a19c5+a18c10+a18c8
+a18c6+a18c2+a17c13+a17c5+a17c3+a16c6+a16c4+a16+a15c13
+a15c11+a15c9+a15c7+a15c5+a14c12+a14c8+a14c6+a14c4+a13c17
+a13c11+a13c5+a12c14+a12c12+a12c10+a12c8+a12c6+a12c4+a11c11
+a11c9+a11c7+a10c12+a10c10+a9c17+a9c7+a8c12+a8c10+a8c8+a8c4
+a7c11+a6c14+a6c12+a6c6+a5c17+a5c15+a5c13+a5c11+a4c12+a2c10
+c12)*b4+(a27c5+a26c4+a25c7+a25c5+a25c3+a24c6+a24c2+a23c7
+a23c5+a23c3+a22c6+a21c7+a21c5+a21c3+a20c8+a19c13+a19c9+a19c7
```

[2] The first equality is a general fact (cf. [GP3]). To see that $\langle J3[1],\ldots,J3[6]\rangle : f^\infty = J$, it is sufficient to know that $J \supset \langle J3[1],\ldots,J3[6]\rangle$, $J = J : f$ and that $\langle J3[1],\ldots,J3[6]\rangle : f^\infty$ is a prime ideal, which we shall see later. This is, computationally, much easier to check than a direct computation.

```
+a19c5+a18c12+a18c10+a18c8+a18c6+a18c4+a18+a17c15+a17c13+a17c9
+a17c5+a16c14+a16c12+a16c8+a16c4+a15c15+a15c3+a14c12+a14c10
+a14c6+a14c4+a13c11+a13c5+a12c14+a12c8+a11c13+a11c9+a11c5
+a10c14+a10c12+a10c10+a9c13+a9c11+a9c9+a8c12+a8c10+a7c13+a6c14
+a5c15+a4c14+a4c12+a4c8+a3c15+a3c13+a2c14+a2c10)*b2
+(a26c6+a24c4+a22c6+a20+a18c14+a16c12+a16c4+a16+a14c14+a14c10
+a14c2+a8c12+a8c8+a8c4+a6c14+a6c10+a4c12+a2c14+a2c10+c8);
```

h is a polynomial of degree 12 with respect to b and therefore $\dim_{\mathbb{F}_2(a,c)} \mathbb{F}_2(a,c)[b]/(J3 \cap \mathbb{F}_2(a,c)[b]) = 12$. Since $\dim_{\mathbb{F}_2(a,c)} \mathbb{F}_2(a,c)[w,y,b,d,x,v]/J3$ is also 12, we know that a lexicographical Gröbner basis with respect to $b < v < x < d < y < w$ of $J3$ must have leading polynomials as follows: b^{12}, v, x, d, y, w.[3]

It follows that the projection

$$[a,b,c,d,v,w,x,y] \to (a,b,c)$$

over the field $\mathbb{F}_2(a,c)$ is birational on $V(J3)$. The image of $\mathbf{V}(J3)$ in $\mathbb{F}_2(a,c)[b]$ is defined by the polynomial h.

This implies that $J3\overline{\mathbb{F}}_2(a,c)[w,y,d,x,v,b]$ is a prime ideal if h is absolutely irreducible. In particular, we obtain that J is absolutely irreducible if h is absolutely irreducible.

To prove that h is absolutely irreducible, we proceed as follows:

First we show that the radical of the ideal of the coefficients of h in $\overline{\mathbb{F}}_2[a,c]$ with respect to b is $\langle a,c\rangle \cap \langle a+1, c+1\rangle$. We do this using the factorising Gröbner basis algorithm.

```
ideal JF=coeffs(h,b); facstd(JF);
[1]:
   _[1]=c
   _[2]=a
[2]:
   _[1]=c+1
   _[2]=a+1
```

This implies that h cannot have a nontrivial factor in $\mathbb{F}_2[a,c]$. Then we consider $\tilde{h}(b,c) = h(1,b,c)$.

```
subst(h,a,1);

(c+1)^14*b12+(c+1)^14*b10+(c+1)^11*(c6+c5+c4+c+1)*b8+(c+1)^11
*(c6+c4+c2+c+1)*b6+(c+1)^8*(c9+c7+c5+c4+c3+c2+1)*b4+(c+1)^10
*b2+(c+1)^10*c2;
```

It is sufficient to show that $f(x,c) = \tilde{h}\left(\frac{x}{c+1}, c\right)/(c+1)^2$ is absolutely irreducible. To simplify the situation, we make the transformation $c \mapsto c+1$.

Let $a_4 = c^6 + c^5 + c^4 + c^2 + 1$ and $a_2 = c^9 + c^8 + c^7 + c^6 + c^4 + c^2 + 1$.

Lemma 3.3. *The polynomial*

$$f = x^{12} + c^2x^{10} + ca_4x^8 + c^3(c^6+c+1)x^6 + c^2a_2x^4 + c^6x^2 + c^8(c+1)^2$$

is irreducible in $\overline{\mathbb{F}}_2[x,c]$.

Proof. We check that $f(x,c^2)$ is the square of some polynomial g, that is, g is defined by $g^2(x,c) = f(x,c^2)$. It suffices to prove that g is irreducible: if $f = f_1f_2$ is a non-trivial decomposition, then $g^2 = f_1(x,c^2)f_2(x,c^2)$. If g is irreducible, we obtain $g = f_1(x,c^2) = f_2(x,c^2)$. This implies $f = f_1^2$, which is obviously not true.

[3]We do not need to compute directly $J3 \cap \mathbb{F}_2(a,c)[b] = \langle h\rangle$ which is difficult. Once h is given, it suffices to know that h is irreducible of degree 12, $\dim_{\mathbb{F}_2(a,c)} \mathbb{F}_2(a,c)[w,y,b,d,x,v]/J3 = 12$ and $h \in J3$, which is much easier to check.

The polynomial g is $g = x^6 + c^2x^5 + ca_4x^4 + c^3(c^6 + c + 1)x^3 + c^2a_2x^2 + c^6x + c^8(c + 1)^2$.

First step: g has no linear factor in x.

A linear factor of g has to be of the form $x - x_0c^i(c + 1)^j$ for some $x_0 \in \overline{\mathbb{F}}_2$ and $i \leq 8$, $j \leq 2$. Now it is easy to see, using divisibility by c, that $g\big(x_0c^i(c + 1)^j, c\big) \neq 0$ for $i = 0, 1, 2, 4, \ldots, 8$. In the case $i = 3$, $g\big(x_0c^3(c + 1)^j, c)\big) \neq 0$ because $\big(x_0c^3(c + 1)^j\big)^6$ has degree $18 + 6j$ with respect to c, which is strictly larger than the degree of the other summands.

Second step: g has no quadratic factor in x.

Assume that $g = (x^4 + \alpha x^3 + \beta x^2 + \gamma x + \delta)(x^2 + \varepsilon x + \mu)$ for $\alpha, \ldots, \mu \in \overline{\mathbb{F}}_2[c]$. Then we obtain

$$\begin{array}{lrcl}
(1) & \mu\delta & = & c^8(c+1)^2 \\
(2) & \mu\gamma + \varepsilon\delta & = & c^6 \\
(3) & \mu\beta + \varepsilon\gamma + \delta & = & c^2a_2 \\
(4) & \mu\alpha + \varepsilon\beta + \gamma & = & c^3(c^6 + c + 1) \\
(5) & \mu + \varepsilon\alpha + \beta & = & ca_4 \\
(6) & \varepsilon + \alpha & = & c^2.
\end{array}$$

Now $g(x, 0) = x^6$ implies that $c|\alpha, \beta, \gamma, \delta, \varepsilon, \mu$. Therefore, they all have degree ≤ 10. Equation (2) implies that $(c + 1)|\mu$ and $(c + 1)|\delta$ are not possible. (3) and (4) imply that $c^2|\delta$ and $c^2|\gamma$ and, therefore $\deg(\mu) \leq 8$. (4), (5) and (6) imply that $\deg(\varepsilon) \leq 4$ and $\deg(\alpha) \leq 4$.

If $\deg(\mu) = 8$, then $\deg(\mu\gamma) \geq 10$ and (2) implies that $\deg(\varepsilon\delta) \geq 10$. This implies that $\deg(\varepsilon) \geq 8$, which is not possible, as we already saw.

If $\deg(\varepsilon) = \deg(\alpha) = 4$, then (5) implies $\deg(\beta) = 8$. This implies $\deg(\varepsilon\beta) = 12$ and, therefore, by (4), $\deg(\mu\alpha) = 12$. This contradicts $\deg(\mu) \leq 7$ and $\deg(\alpha) = 4$. Thus, we have $\deg(\varepsilon) \leq 3$ and $\deg(\alpha) \leq 3$. This implies, using (5), $\deg(\beta) \leq 7$.

If $\deg(\mu) = 6$, then $\deg(\delta) = 4$ implies $\deg(\mu\gamma) \geq 8$ and $\deg(\delta\varepsilon) \leq 7$, contradicting (2).

We obtain $\deg(\mu) \leq 5$ and, using (5), $\deg(\beta) = 7$. If $\deg(\mu) \leq 3$, then $\deg(\mu\beta) \leq 10$ and (3) implies $\deg(\varepsilon\gamma) = 11$. We shall see that this is not possible.

If $\deg(\varepsilon) = 3$, then $\deg(\varepsilon\beta) = 10$ and (4) implies that $\deg(\gamma) = 10$. This contradicts $\deg(\varepsilon\gamma) = 11$.

If $\deg(\varepsilon) = 2$ then $\deg(\varepsilon\beta) = 9$ and $\deg(\gamma) = 9$. This contradicts (2).

If $\deg(\varepsilon) = 1$, then $\deg(\varepsilon\beta) = 8$ and $\deg(\gamma) = 10$. This contradicts (4).

Finally, we obtain $4 \leq \deg(\mu) \leq 5$. This implies $c^2|\mu$ and $c^3|\delta$ and, consequently, $c^3|\mu\beta$. But we know already that $c^2|\gamma$ and, therefore, $c^3|\varepsilon\gamma$ and obtain a contradiction to (3).

Third step: g has no cubic factor in x.

Let $g = (x^3 + \alpha x^2 + \beta x + \gamma)(x^3 + \delta x^2 + \varepsilon x + \mu)$, then we obtain

$$\begin{array}{lrcl}
(1) & \gamma\mu & = & c^8(c+1)^2 \\
(2) & \gamma\varepsilon + \mu\beta & = & c^6 \\
(3) & \mu\alpha + \varepsilon\beta + \gamma\delta & = & c^2a_2 \\
(4) & \mu + \alpha\varepsilon + \beta\delta + \gamma & = & c^3(c^6 + c + 1) \\
(5) & \varepsilon + \alpha\delta + \beta & = & ca_4 \\
(6) & \alpha + \delta & = & c^2.
\end{array}$$

Now, $g(x, 0) = x^6$ implies $c|\alpha, \beta, \gamma, \delta, \varepsilon, \mu$.

As in the previous case, $(c+1)|\gamma, \mu$ is not possible. If $c^4 \nmid \gamma$, then $c^5|\mu$ and, by (2), $c^6|\varepsilon\gamma$, which implies $c^3|\varepsilon$. This contradicts (3) and (4), because (3) implies $c^3 \nmid \gamma\delta$ and, therefore, $c^2 \nmid \gamma$.

We obtain that $c^4|\gamma$ and, by symmetry, $c^4|\mu$. We may assume that $\gamma = \gamma_0 c^4 (c+1)^2$ and $\mu = \mu_0 c^4$ for suitable $\gamma_0, \mu_0 \in \overline{\mathbb{F}}_2$. This implies $\deg(\delta) \leq 5$, $\deg(\varepsilon) \leq 5$ and $\deg(\alpha) \leq 7$, $\deg(\beta) \leq 7$ by using (3), since a_2 is of degree 9.

If $\deg(\alpha) \geq 4$, then $\deg(\delta) \geq 4$ by (6). This implies $\deg(\alpha\delta) \geq 8$, which contradicts (5). We obtain that $\deg(\alpha) \leq 3$, $\deg(\delta) \leq 3$. This implies, using (5), that $\deg(\beta) = 7$. Now (4) implies that $\deg(\delta) = 2$ and we obtain, using (3), that $\deg(\varepsilon) = 4$. This is a contradiction to (2) and finishes the third step. $\square$

Altogether, we proved now that $V' = \mathbf{V}(J)$ is absolutely irreducible. Next we compute the singular locus of $\mathbf{V}(J)$, using Singular (with a special procedure).

Lemma 3.4. *The singular locus of $\mathbf{V}(J)$ is the union of the following six smooth curves defined by the ideals $S1, \ldots, S6$.*

```
S1[1]=y;          S1[2]=x;          S1[3]=v2+vw+w2+1;  S1[4]=d+1;        S1[5]=c+
1;
S1[6]=b+w+1;      S1[7]=a+1;

S2[1]=y+1;        S2[2]=x+1;        S2[3]=v+w+1;        S2[4]=d;          S2[5]=c;
S2[6]=b2+w2+w+1;  S2[7]=a;

S3[1]=y+1;        S3[2]=x+1;        S3[3]=v+1;          S3[4]=d           S3[5]=c;

S3[6]=b2+w+1;     S3[7]=a2+ab+w;

S4[1]=y;          S4[2]=x;          S4[3]=v;            S4[4]=d+          S4[5]=c+
1;
S4[6]=b2+b+w+1;   S4[7]=a+b+1;

S5[1]=x2+y;       S5[2]=wy+x;       S5[3]=wx+1;         S5[4]=v;          S5[5]=d2
+xy+x;
S5[6]=c;          S5[7]=by+b+dw+d;  S5[8]=bx+b+dw;      S5[9]=bw+b+dw2;
S5[10]=bd+x+1;    S5[11]=b2+w;      S5[12]=a+dw;

S6[1]=x;         S6[2]=w3y+w2+1;    S6[3]=v+w2y;        S6[4]=d+wy+1;     S6[5]=c+
w2y+w;
S6[6]=b+w;       S6[7]=a;
```

Corollary 3.5. *The singular locus of V' is contained in the set $S = V' \cap \mathbf{V}(xc)$. The variety $U = V' \smallsetminus S$ is a smooth irreducible affine surface invariant under the morphism α. For any odd n, α^n has no fixed points in S.*

Proof. The first two assertions are checked directly, looking at the equations S1–S6 and the equation for the action of α. To prove the third, assume that $p = (a, b, \ldots, y)$ is a fixed point of α^n lying on S. Let $x = 0$. Then, since p is α^n invariant, we have $c = 0$. Since $p \in V'$, equation J3[1] gives $a^8v^6 + a^8v^4 = 0$. (The variety defined by the ideal J3 contains V' as a component, so p must satisfy all the equations of J3.) Hence we have either $a = 0$, or $v = 0$, or $v = 1$.

In any of the two first cases we have $a = v = c = x = 0$, and equation J3[3] gives $d = 0$. Since p is an invariant point, we get $y = 0$. Furthermore, equation J[4] gives $w = 0$. Hence $b = 0$, contradiction.

If $v = 1$, then $a = 1$ which, taking into account $a = c = 0$, contradicts J[7]. $\square$

3.3. **Trace formula.** Throughout this subsection k denotes a (fixed) algebraic closure of $\mathbb{F}_2$. All varieties under consideration, even those defined over $\mathbb{F}_2$, are viewed as k-varieties.

Let V' be the variety defined by equations $J[1], \dots, J[10]$ (see Subsection 3.2). We have seen that this is an irreducible affine surface. Computations in Subsection 3.2 show that the singular locus of V' is contained in the set $S = V' \cap \mathbf{V}(xc)$. By Corollary 3.5 the variety $U = V' \setminus S$ is a smooth irreducible affine surface invariant under the morphism α acting in $\mathbb{A}^8$ as

$$\alpha(a,b,c,d,v,w,x,y) = (v,w,x,y,a^2,b^2,c^2,d^2) \tag{3.2}$$

(see Subsection 3.1).

Our goal is to prove that for n odd and large enough, the set U has an α^n-invariant point. In this subsection we prove an estimate of Lang–Weil type:

Theorem 3.6. *With the above notation, let $\#\operatorname{Fix}(U,n)$ be the number of fixed points of α^n (counted with their multiplicities). Then for any odd $n > 1$ the following inequality holds:*

$$|\#\operatorname{Fix}(U,n) - 2^n| \le b^1(U)2^{3n/4} + b^2(U)2^{n/2}, \tag{3.3}$$

where $b^i(U) = \dim H^i_{\text{ét}}(U, \overline{\mathbb{Q}}_\ell)$ are ℓ-adic Betti numbers ($\ell \neq 2$).

The strategy of proof is as follows. The operator α and all its powers act on the étale ℓ-adic cohomology groups $H^i_c(U, \overline{\mathbb{Q}}_\ell)$ of U (with compact support). We are going to apply Deligne's conjecture (proved by T. Zink for surfaces [Zi], by Pink [Pi] in arbitrary dimension (modulo resolution of singularities), and by Fujiwara [Fu] in the general case) saying that the Lefschetz(–Weil–Grothendieck–Verdier) trace formula is valid for any operator on U composed with sufficiently large power of the Frobenius (in our case this means sufficiently large odd power of α). We shall show that in our case the trace formula is already valid after twisting with the first power of the Frobenius. This fact is a consequence of the above mentioned results on Deligne's conjecture together with the following crucial observation: roughly speaking, if we consider the closure $\overline{U}$ of U in $\mathbb{P}^8$, α (as well as any of its odd powers) has no fixed points at the boundary (that is, on $\overline{U} \setminus U$). As soon as the trace formula is established, the proof can be finished by applying Deligne's estimates of the eigenvalues of the Frobenius.

Let us make all this more precise.

Denote by Γ (the transpose of) the graph of α acting on $\mathbb{A}^8$ by formulas (3.2), that is, $\Gamma = \{(\alpha(M), M) : M \in \mathbb{A}^8\}$, and let $\Gamma_U = \Gamma \cap (U \times U)$.

Consider the natural embedding $\mathbb{A}^8 \subset \mathbb{P}^8$, and denote by $\overline{\Gamma}$ (resp. $\overline{\Gamma}_U$) $\subset \mathbb{P}^8 \times \mathbb{P}^8$ the closure of Γ (resp. Γ_U) with respect to this embedding. Let $H_0 = (\mathbb{P}^8 \times \mathbb{P}^8) \setminus (\mathbb{A}^8 \times \mathbb{A}^8)$, $H_1 = (V' \times V') \setminus (U \times U)$, $H = H_0 \cup H_1$. Let Δ denote the diagonal of $\mathbb{A}^8 \times \mathbb{A}^8$, $\overline{\Delta}$ the diagonal of $\mathbb{P}^8 \times \mathbb{P}^8$, $\Delta_U = \Delta \cap \Gamma_U$, and $\overline{\Delta}_U = \overline{\Delta} \cap \overline{\Gamma}_U$. If n is a positive integer, denote the corresponding objects related to α^n by $\Gamma^{(n)}$, $\overline{\Gamma}^{(n)}$, $\Gamma_U^{(n)}$, $\overline{\Gamma}_U^{(n)}$, $\Delta_U^{(n)}$, $\overline{\Delta}_U^{(n)}$.

Lemma 3.7. *If n is odd, $\overline{\Delta}_U^{(n)} = \Delta_U^{(n)}$.*

Proof. We have

$$\overline{\Delta}_U^{(n)} \setminus \Delta_U^{(n)} = \overline{\Gamma}_U^{(n)} \cap \overline{\Delta} \cap H.$$

We wish to prove that this set is empty. Since

$$\overline{\Gamma}_U^{(n)} \cap \overline{\Delta} \cap H \subseteq \overline{\Gamma}^{(n)} \cap (\overline{U \times U}) \cap \overline{\Delta} \cap H,$$

it is enough to prove that

$$\overline{\Gamma}^{(n)} \cap \overline{\Delta} \cap H = \emptyset.$$

First note that

$$\overline{\Gamma}^{(n)} \cap \overline{\Delta} \cap H_1 = \Gamma^{(n)} \cap \Delta \cap H_1 = \emptyset$$

(the first equality is obvious since H_1 is contained in $\mathbb{A}^8 \times \mathbb{A}^8$, and the second one immediately follows from Corollary 3.5). Hence we only have to prove that $\overline{\Gamma}^{(n)} \cap \overline{\Delta} \cap H_0 = \emptyset$.

Let $(a, b, c, d, v, w, x, y), (a', b', \dots, y')$ be the coordinates in $\mathbb{A}^8 \times \mathbb{A}^8$, and let $(a : b : \dots : t), (a' : b' : \dots : t')$ be the homogeneous coordinates in $\mathbb{P}^8 \times \mathbb{P}^8$. Suppose that

$$M = ((a : b : \dots : t), (a' : b' : \dots : t')) \in \overline{\Gamma}^{(n)} \cap \overline{\Delta} \cap H_0.$$

If $n = 2m + 1$, denote $s = 2^m$. With this notation, since $M \in \overline{\Gamma}^{(n)}$, formulas (3.2) imply that

$$a' = v^s t^s,\ b' = w^s t^s,\ c' = x^s t^s,\ d' = y^s t^s,\ v' = a^{2s},\ w' = b^{2s},\ x' = c^{2s},\ y' = d^{2s},\ t' = t^{2s}\,.$$

On the other hand, since $M \in H_0$, we have $t = t' = 0$, and hence $a' = b' = c' = d' = 0$. Furthermore, since $M \in \overline{\Delta}$, we have $a' = \lambda a$, $b' = \lambda b$, $c' = \lambda c$, $d' = \lambda d$ for some $\lambda \in k$, and hence $a = b = c = d = 0$. This implies $v' = w' = x' = y' = 0$, contradiction. $\square$

The next goal is to show that the Lefschetz trace formula holds for all odd nth powers of α ($n > 1$). We shall do it using the above mentioned results on Deligne's conjecture. First we briefly recall the general approach ([SGA5], [Zi], [Pi], [Fu]); we mainly use the notation of [Pi] and refer the reader to that paper for more details.

(i) *Global term.* We can (and shall) view our operator α as a particular case of the correspondence a:

$$U \xleftarrow{a_1} \Gamma_U \xrightarrow{a_2} U$$

(here a_1 and a_2 stand for the first and second projections, respectively). We regard an odd power α^{2m+1} as a "twisted" correspondence $b = \mathrm{Fr}^m \circ a$ with $b_1 = \mathrm{Fr}^m \circ a_1$, $b_2 = a_2$.

Let Λ denote a finite field extension of $\mathbb{Q}_\ell$, L a constructible Λ-sheaf (in our situation it suffices to consider the constant sheaf $L = \overline{\mathbb{Q}}_\ell$). Then a cohomological correspondence u on L with support in b is a morphism $u\colon b_1^* L \to b_2^! L$, where * stands for the inverse image functor, and $^!$ for the extraordinary inverse image functor (cf. [Pi, Section 1] and references therein); in our situation $b_2 = \mathrm{id}$ and hence $b_2^! L = L$. Since b_1 is a proper morphism, u induces an endomorphism $u_!\colon H_c^\bullet(U, L) \to H_c^\bullet(U, L)$ which possesses a well-defined trace $\mathrm{tr}(u_!) \in \Lambda$; this is the global term in the desired trace formula. In down-to-earth terms, in our situation we have

$$\mathrm{tr}(u_!) = \sum_{i=0}^{4} (-1)^i \mathrm{tr}(\alpha^n | H_c^i(U, \overline{\mathbb{Q}}_\ell). \tag{3.4}$$

(ii) *Compactification.* Furthermore, since b_1 is proper, our correspondence b can be extended to a compactification $\bar{b}$

$$\begin{array}{ccccc} U & \xleftarrow{b_1} & \Gamma_U & \xrightarrow{b_2} & U \\ {\scriptstyle j}\downarrow & & \downarrow & & \downarrow \\ \bar{U} & \xleftarrow{\bar{b}_1} & \overline{\Gamma}_U & \xrightarrow{\bar{b}_2} & \bar{U} \end{array}$$

where the vertical arrows are open embeddings and the bottom line is proper. This gives rise to a cohomological correspondence $\bar{u}_!$ on the sheaf $j_! L$ with support in $\bar{b}$; here $_!$ stands for the direct image functor with compact support (extension by 0), cf. [Pi, 2.3].

The global term does not change after compactification:

$$\mathrm{tr}(\bar{u}_!) = \mathrm{tr}(u_!) \tag{3.5}$$

(see [Pi, Lemma 2.3.1]).

For a compactified correspondence the Lefschetz–Verdier trace formula is known (cf. [Pi, 2.2.1]):

$$\mathrm{tr}(\bar{u}_!) = \sum_D LT_D(\bar{u}) \tag{3.6}$$

where D runs over all the connected components of $\mathrm{Fix}(\bar{b})$, and the local terms $LT_D(\bar{u})$ are defined as in [Pi, 2.1]. In our case $\mathrm{Fix}(\bar{b})$ consists of isolated points (since this is true for the Frobenius), and all these points are contained in U (because of Lemma 3.7 there are no fixed points at the boundary, neither on the singular locus, nor at infinity).

(iii) *Local terms.* Suppose that b_2 is quasifinite and y is a point not at infinity. Let $x = b_2(y)$, then

$$d(y) = [k(y)/k(x)]_i \cdot \operatorname{length} O_{\Gamma_U,y}/b_2^*(m_{U,x}O_{U,x}),$$

where $[k(y)/k(x)]_i$ denotes the inseparable degree of the residue field extension. Clearly, in our case $b_2 = \mathrm{id}$ implies $d(y) = 1$.

By [Fu, Th. 5.2.1], for an isolated fixed point y at finite distance we have

$$LT_y(u) = \mathrm{tr}_y(u) \tag{3.7}$$

provided $2^m > d(y)$. In our setting,

$$\mathrm{tr}_y(u) \text{ equals the multiplicity of } y \tag{3.8}$$

(cf. [Zi, p. 338], [Pi, 8.3.1]).

(iv) Summing up, (i)–(iii) (or, more precisely, formulas (3.4), (3.5), (3.6), (3.7), (3.8), together with Lemma 3.7) imply

Proposition 3.8. *If $n > 1$ is an odd integer, then*

$$\#\,\mathrm{Fix}(U,n) = \sum_{i=0}^{4}(-1)^i \mathrm{tr}\big(\alpha^n \mid H^i_{\mathrm{c}}(U,\overline{\mathbb{Q}}_\ell)\big)\,. \tag{3.9}$$

We are now ready to prove Theorem 3.6. Since U is non-singular, the ordinary and compact Betti numbers of U are related by Poincaré duality [Ka2, p. 6], and we have $b^i_{\mathrm{c}} = b^{4-i}$. Since U is affine, $b^i(U) = 0$ for $i > 2$ [Ka2, loc. cit.]. Since U is geometrically integral, $b^0(U) = 1$ and Fr acts on the one-dimensional vector space $H^0(U,\overline{\mathbb{Q}}_\ell)$ as multiplication by 4 [Ka2, loc. cit.]. Hence α acts on the same space as multiplication by 2. (Indeed, if it were multiplication by (-2), for a sufficiently big power of α the right-hand side of (3.9) would be negative.) Hence α^n acts as multiplication by 2^n. Thus $\mathrm{tr}(\alpha^n|H^4_{\mathrm{c}}(U,\overline{\mathbb{Q}}_\ell)) = 2^n$.

On the other hand, according to Deligne [De, Th. 1] for every eigenvalue α_{ij} of Fr acting on $H^i_{\mathrm{c}}(U,\overline{\mathbb{Q}}_\ell)$ we have $|\alpha_{ij}| \leq 2^{i/2}$. This yields similar inequalities for the eigenvalues β_{ij} of α: $|\beta_{ij}| \leq 2^{i/4}$ and the eigenvalues $\beta_{ij,n}$ of α^n: $|\beta_{ij,n}| \leq 2^{ni/4}$. We thus obtain

$$\big|\mathrm{tr}\big(\alpha^n|H^3_{\mathrm{c}}(U,\overline{\mathbb{Q}}_\ell)\big)\big| \leq b^1(U)2^{3n/4}, \quad \big|\mathrm{tr}\big(\alpha^n|H^2_{\mathrm{c}}(U,\overline{\mathbb{Q}}_\ell)\big)\big| \leq b^2(U)2^{n/2}\,.$$

This proves the theorem. □

Remark 3.9. Probably one can get another proof of Proposition 3.8 (and hence Theorem 3.6) using an approach of [DL]. In that paper the Lefschetz trace formula is established for any endomorphism of finite order. A remark in Section 11 of the above cited paper (see also [SGA4, Sommes trig., 8.2, p. 231]) says that the results of the paper can be extended to the case of an endomorphism α with the property $\alpha^2 = \mathrm{Fr}$.

3.4. **Estimates of Betti numbers.** As in the previous subsection, we assume that the ground field is $k = \bar{\mathbb{F}}_2$.

Recall that we consider the variety V' defined by equations $J[1–10]$ (see Subsection 3.2) whose singular locus is contained in the set $S = V' \cap \mathbf{V}(xc)$. As before, we denote $U = V' \setminus S$; it is a smooth irreducible affine variety invariant under the morphism α. Our aim is to estimate $b^1(U)$ and $b^2(U)$.

First we deal with $b^1(U)$. We want to use the Lefschetz Theorem on hyperplane sections. For technical reasons we want to use hyperplanes of special type, namely those defined by equations $\alpha a + \beta c + \gamma = 0$.

These hyperplane sections are not general, and in order to apply the Lefschetz Theorem, we have to provide a quasifinite map of the surface V' onto $\mathbb{A}^2$ with coordinates a, c.

The next step is to estimate the Euler characteristic of U. To do this, we represent U as the union of an open subset U' and a finite number of curves. We estimate the Euler characteristics of these curves and of U' separately, using the fact that U' is a double cover of a simpler variety. Having in hand bounds for $b^1(U)$ and $\chi(U)$, we estimate $b^2(U)$.

Proposition 3.10. *A regular map* $\pi\colon \mathbb{A}^8 \to \mathbb{A}^2$ *defined as* $\pi(a,b,c,d,v,w,x,y) = (a,c)$ *is quasi-finite on* U.

Proof. Consider the variety $\widetilde{W}$ defined in $\mathbb{A}^8$ by equations $J3[1-6]$.

We have $\widetilde{W} \supset V'$ and $\widetilde{W} \smallsetminus V' \subset \mathbf{V}(f) \subseteq \mathbb{A}^8$, where

$$f(a,c) = c(ac+1)a(a+c)(c+1)(a^3 + a^2c^3 + a^2c + ac^4 + ac^2 + c)$$

(see Subsection 3.2). This means that the coordinates of any point of V' (and, in particular, of U), satisfy the equations $J3[1-10]$. If $f(a,c) \neq 0$, the equation $J3[1]$ provides at most six different possible values for v. The equation $J3[2]$ implies that for each of these six values only one value of x is possible. The equation $J3[3]$ gives at most two values for d, and all the proceeding equations provide one value for b, y and w. Hence, for any point $(a,c) \in \mathbb{A}^2$ with $f(a,c) \neq 0$, the preimage $\pi^{-1}(a,c)$ is finite in V' and hence in U.

Let now $A = \mathbf{V}(f) \subset \mathbb{A}^2$. Then $\pi^{-1}(A) \cap U = \cup A_i$, $i = 1, \dots, 6$ which may be described as follows.

(1) $A_1 = U \cap \mathbf{V}(c-1)$.

According to calculations, $A_1 = A_1^1 \cup A_1^2$, where $A_1^1 = U \cap \mathbf{V}(c-1) \bigcap \mathbf{D}\big(a(a+1)(a^2+a+1)\big)$ and $A_1^2 = U \cap \mathbf{V}\big(c-1, a(a+1)(a^2+a+1)\big)$.

The set A_1^1 is defined by the ideal L.

```
L[1]=c+1;
L[2]=(a5+a4+a3+a2)*v4+(a5+a)*v2+(a4+a2+1);
L[3]=x+(a3+a2)*v3+(a3+a2+a)*v;
L[4]=(a+1)*d2+(a4+a2)*dv3+(a4+a)*dv+(a8+a6+a5+a4+a3+a2+1)*v2+(a8+a5+a+1);
L[5]=(a4+a2+1)*b+(a5+a4+a2+a)*dv2+(a4+a)*d+(a6+a4)*v3+(a6+a2+a+1)*v;
L[6]=(a2+a+1)*y+(a2+a+1)*d+(a7+a6+a5+a2)*v3+(a7+a6+a4+a3+a2+a)*v;
L[7]=(a4+a2+1)*w+(a6+a5+a3+a2)*dv2+(a5+a2)*d+(a7+a4+a2+a)*v3+(a7+a6+a5+a)*v;
```

These equations show that for a fixed value of a, if $a(a+1)(a^2+a+1) \neq 0$, there are at most four points in $U \cap \pi^{-1}(a,1)$. The set A_1^2 is defined by the ideal L1.

```
L1[1]=a2+a+1        L1[2]=c+1        L1[3]=v+a+1     L1[4]=x+a
L1[5]=d2+da+1       L1[6]=b+da+d     L1[7]=y+d+a     L1[8]=w+d+1
```

It follows that $\pi^{-1}(1,1) = \emptyset$; $\pi^{-1}(0,1) = \emptyset$; $\pi^{-1}(a_0,1)$, where a_0 is a root of a^2+a+1, consists of two points.

(2) $A_2 = U \cap \mathbf{V}(c) = \emptyset$.
(3) $A_3 = U \cap \mathbf{V}(a) = \emptyset$.
(4) $A_4 = U \cap \mathbf{V}(a+c) = \emptyset$.
(5) $A_5 = U \cap \mathbf{V}(ac+1) = A_5^1 \cup A_5^2$.
(5.1) $A_5^1 = U \cap \mathbf{V}(ac+1) \cap \mathbf{D}\big((a^2+a+1)(a+1)a\big)$. This set is defined by the ideal D.

```
D[1]=(a)*c+1;
D[2]=(a3+a2)*v2+(a3+a2+a)*v+1;
D[3]=x+(a4)*v3+(a4)*v;
D[4]=(a6+a2)*d2+(a9+a5)*dv3+(a9+a5)*dv+(a10+a6+a4+a2+1)*v2+(a10+a2+1);
D[5]=(a3+a)*b+(a5+a3)*dv2+(a)*d+(a4+a2)*v3+(a4+a2+1)*v;
D[6]=(a)*y+d+(a7+a5)*v3+(a7+a5+a3)*v;
D[7]=(a2+1)*w+(a5+a3)*dv2+(a)*d+(a4+a2)*v3+v;
```

which show that for any point $a \neq 0, 1$, or a_0 (a root of a^2+a+1), the set $\pi^{-1}(a, \frac{1}{a})$ contains at most four points.

(5.2) $A_5^2 = U \cap \mathbf{V}\big((ac+1), (a+1)(a^2+a+1)a\big)$.

This set consists of four points defined by the ideal D1.

```
D1[1]=a2+a+1        D1[2]=c+a+1      D1[3]=v+a            D1[4]=x+1
D1[5]=d2+da+d+1     D1[6]=b+d+a      D1[7]=y+da+d+a+1     D1[8]=w+da+a
```

(6) $A_6 = U \cap \mathbf{V}(h_1)$, where $h_1(a,c) = a^3 + a^2c^3 + a^2c + ac^4 + ac^2 + c$.
$A_6 = A_6^1 \cup A_6^2 \cup A_6^3$, where
$A_6^1 = U \cap \mathbf{V}(h_1) \cap \mathbf{D}\big(v^2+ac^3+c^2+a^2, a(a+1)(a^2+a+1)\big)$;
$A_6^2 = U \cap \mathbf{V}(h_1, v^2+ac^3+c^2+a^2) \cap \mathbf{D}\big(a(a+1)(a^2+a+1)\big)$;
$A_6^3 = U \cap \mathbf{V}\big(h_1, v^2+ac^3+c^2+a^2, a(a+1)(a^2+a+1)\big)$.

The set A_6^1 is defined by the ideal K1.

```
K1[1]=(a)*c4+(a2)*c3+(a)*c2+(a2+1)*c+(a3);
K1[2]=(a3+a)*v4+(a6+a4+a2)*v2c3+(a5+a)*v2c2+v2c+(a7+a3+a)*v2+(a10+a8+1)*c3
     +(a9+a5)*c2+(a4+a2+1)*c+(a11+a7+a5);
K1[3]=(a2+1)*xv2+(a3+a)*xc3+(a2+1)*xc2+(a4+a2)*x+(a)*v3c3+(a)*v3c+v3+(a5)*vc3
     +(a4+a2)*vc2+(a)*vc+(a6+a4+a2)*v;
K1[4]=(a6+a4+a2+1)*x2+(a7+a5)*xvc3+(a8+a4)*xvc2+(a7+a)*xvc+(a8+a2)*xv
     +(a3)*v2c3+(a8+a6)*v2c2+(a7+a5+a3)*v2c+(a6)*v2+(a11+a3+a)*c3+(a10+a8)*c2
     +(a5+a3+a)*c+(a12+a10+a6+a4+1);
K1[5]=(a10+a6+a4+1)*d2+(a10+a6+a4+1)*dxc+(a13+a11+a9+a7+a3+a)*xvc3
     +(a14+a12+a4+a2)*xvc2+(a13+a9+a7+a5)*xvc+(a6+a4)*xv+(a13+a11+a9+a7+a3)
     *v2c3+(a14+a12+a8+a6+a2+1)*v2c2+(a13+a11+a5+a3+a)*v2c+(a8+a4+a2)*v2
     +(a13+a9+a5)*c3+(a12+a6)*c2+(a5+a3+a)*c+(a14+a12+a8+a4+a2);
K1[6]=(a12+a10+a8+a6+a4+a2)*b+(a3+a)*dxvc3+(a8+a6)*dxvc2+(a5+a)*dxvc
     +(a10+a4+a2+1)*dxv+(a5+a3)*dv2c3+(a10+a8)*dv2c2+(a7+a3)*dv2c
     +(a12+a6+a4+a2)*dv2+(a7+a5+a3)*dc3+(a12+a8+a6+a2)*dc2+(a11+a9+a7)*dc
     +(a12+a10+a8)*d+(a8+a2)*xc3+(a7+a)*xc2+(a10+a8+a4+a2)*xc+(a11+a9+a5+a3)
     *x+(a13+a11+a9+a7+a5+a3)*v3c2+(a14+a12+a10+a8+a6+a4)*v3c+(a13+a11+a9+a7
     +a5+a3)*v3+(a6+a4+a2)*vc3+(a13+a11+a7+a5)*vc2+(a14+a10+a8+a6+a2)*vc
     +(a11+a9+a7)*v;
K1[7]=y+bdx+bdv3+bdv+bc3+bc+d3c+(a)*d3+dx2c3+dx2c+(a)*dxvc12+(a2)*dxvc11
     +(a2+1)*dxvc9+(a)*dxvc6+(a)*dxvc4+(a2)*dxvc3+(a)*dxvc2+(a2)*dxvc+(a)*dxv
     +dv4c+(a)*dv2c12+(a2)*dv2c11+(a2+1)*dv2c9+dv2c7+(a)*dv2c6+dv2c5+(a2+1)
     *dv2c3+dv2c+(a)*dc10+(a)*dc8+(a2)*dc7+(a2)*dc5+(a2)*dc3+(a)*dc2+(a)
     *d+x3c2+x3+x2vc2+x2v+(a)*xc21+(a2)*xc20+(a2+1)*xc18+(a)*xc17+(a2+1)
     *xc16+(a2)*xc14+xc12+(a)*xc11+(a2+1)*xc10+(a)*xc9+(a)*xc7+(a2)*xc6+(a)
     *xc5+xc4+(a)*xc3+(a)*xc+(a)*v3c19+(a2)*v3c18+(a)*v3c17+v3c16+(a2)*v3c14
     +(a)*v3c11+(a2+1)*v3c10+(a2)*v3c8+(a)*v3c7+(a2)*v3c6+v3c4+v3c2+(a)*v3c
     +(a2)*v3+(a)*vc21+(a2)*vc20+(a)*vc19+vc18+(a2)*vc16+(a)*vc15+(a2)*vc14
```

```
     +(a)*vc13+vc10+(a)*vc9+vc6+(a)*vc5+(a2)*vc4+(a2+1)*vc2+(a2+1)*v;
K1[8]=w+bdx+bdv3+bdvc2+(a)*b+(a)*d3+dx2c3+dx2c+(a)*dxvc12+(a2)*dxvc1
     1+(a)*dxvc10+dxvc9+(a)*dxvc8+(a2)*dxvc5+(a2+1)*dxvc3+(a)*dxvc2+(a2)*dxvc
     +(a)*dxv+(a)*dv2c12+(a2)*dv2c11+(a)*dv2c10+dv2c9+(a)*dv2c8+(a)*dv2c6
     +(a2)*dv2c5+dv2c3+dv2c+(a)*dc10+(a2)*dc7+dc5+(a2+1)*dc3+(a2)*dc
     +(a)*d+x3c2+(a)*xc21+(a2)*xc20+(a)*xc19+xc18+(a2)*xc16+(a)*xc15+(a2)
     *xc14+(a)*xc13+(a)*xc11+xc10+(a2)*xc6+(a2)*xc2+x+v5+(a)*v3c19+(a2)*v3c18
     +(a2+1)*v3c16+v3c14+(a)*v3c13+(a2+1)*v3c12+(a)*v3c11+v3c10+(a)*v3c9+v3c8
     +(a)*v3c7+v3c6+(a)*v3c5+(a)*v3c3+(a2)*v3c2+(a)*v3c+(a2+1)*v3+(a)*vc21
     +(a2)*vc20+(a2+1)*vc18+vc16+vc14+(a2+1)*vc10+(a2)*vc8+vc6+(a)*vc5
     +(a)*vc3+vc2+(a)*vc+v;
```

This shows that each point (a, c) satisfying $f(a, c) = 0$, $v^2+ac^3+c^2+a^2 \neq 0$ and $a(a+1)(a^2+a+1) \neq 0$ has at most four preimages in U_1. The set A_6^2 is defined by the ideal K2.

```
K2[1]=(a)*c4+(a2)*c3+(a)*c2+(a2+1)*c+(a3);
K2[2]=v2+(a)*c3+c2+(a2);
K2[3]=(a4+1)*x2+(a5)*xvc3+(a6+a4)*xvc2+(a5+a3+a)*xvc+(a6+a4+a2)*xv
     +(a3+a)*c3+(a6+a2)*c2+(a3+a)*c+(a2+1);
K2[4]=(a9+a7+a3+a)*d2+(a9+a7+a3+a)*dxc+(a12+a8+a2)*xvc3+(a13+a3)*xvc2
     +(a12+a10+a6)*xvc+(a5)*xv+(a12+a10+a8+a6+a4+a2)*c3+(a13+a11+a7+a5)*c2
     +(a6+1)*c+(a11+a5);
K2[5]=(a12+a10+a8+a6+a4+a2)*b+(a3+a)*dxvc3+(a8+a6)*dxvc2+(a5+a)*dxvc
     +(a10+a4+a2+1)*dxv+(a11+a7+a3)*dc3+(a12+a10+a8+a6+a4+a2)*dc2
     +(a11+a7+a3)*dc+(a12+a8+a4)*d+(a8+a2)*xc3+(a7+a)*xc2+(a10
     +a8+a4+a2)*xc+(a11+a9+a5+a3)*x+(a6+a4+a2)*vc3+(a9+a3)*vc2
     +(a10+a8+a6)*vc+(a11+a9+a7)*v;
K2[6]=(a4+a2+1)*y+(a4+a2+1)*dc+(a)*xc3+(a4+a2)*xc2+(a5)*xc+(a6+a4+a2+1)*x
     +(a5+a3+a)*vc+(a6+1)*v;
K2[7]=(a11+a9+a7+a5+a3+a)*w+(a3+a)*dxvc3+(a8+a6)*dxvc2+(a5+a)*dxvc+(a10
     +a4+a2+1)*dxv+(a11+a7+a3)*dc3+(a12+a10+a8+a6+a4+a2)*dc2+(a11+a7+a3)*dc
     +(a12+a8+a4)*d+(a10+a8+a6+a4+a2+1)*xc+(a11+a9+a7+a5+a3+a)*x+(a6+a4+a2)
     *vc3+(a9+a3)*vc2+(a10+a8+a6)*vc+(a13+a5+a3)*v;
```

It follows that in this case the preimage of each point contains at most four points as well.

The set A_6^3 consists of 54 points defined by the ideals W1, W2, W3, W4, H1, H2 and H3.

```
W1[1]=c3+c+1
W1[2]=a+1
W1[3]=v+1
W1[4]=x+c2+c
W1[5]=d+c2+1
W1[6]=b+c2+c
W1[7]=y+c+1
W1[8]=w+c2

W2[1]=c3+c+1
W2[2]=a+1
W2[3]=v+1
W2[4]=x+c2+c
W2[5]=d+c
W2[6]=b+c2+c+1
W2[7]=y+c2+c
W2[8]=w+c2+1

W3[1]=c3+c+1
W3[2]=a+1
W3[3]=v+c2
W3[4]=x+c
W3[5]=d2+dc2+1
W3[6]=b+d+c

W4[1]=c3+c+1
W4[2]=a+1
W4[3]=v+c2
W4[4]=x+1
W4[5]=d2+dc+c
W4[6]=b+dc
```

```
W3[7]=y+dc+1                       W4[7]=y+dc+c
W3[8]=w+d+c2+1                     W4[8]=w+dc

H1[1]=a2+a+1                       H2[1]=a2+a+1
H1[2]=c3+c2a+c2+ca+a+1             H2[2]=c3+c2a+c2+ca+a+1
H1[3]=v+c2a+c2+1                   H2[3]=v+c2a+c2+1
H1[4]=x+c2a+c2+a+1                 H2[4]=x+ca+1
H1[5]=d2+dc2a+dca+da+c2a+c2        H2[5]=d2+dc2a+dc+c
H1[6]=b+dc2+dc+da+d+c2a            H2[6]=b+dc2a+dc2+dca+da+c+a+1
H1[7]=y+dc+a                       H2[7]=y+dc+c2+ca+a+1
H1[8]=w+dc2a+dca+d+ca+c+1          H2[8]=w+dc2+dca+dc+da+d+c2a+c2+ca+c+a

H3[1]=a2+a+1
H3[2]=c3+c2a+c2+ca+a+1
H3[3]=v+c2a+ca+c+1
H3[4]=x+c2a+c2
H3[5]=d2+dc2a+dc+da+c2a+c2+ca+c
H3[6]=b+dc2a+dc2+dca+c2a
H3[7]=y+dc+c2a+ca+c+a
H3[8]=w+dc2+dca+dc+ca+c+a
```

Thus, any point in $\mathbb{A}^2$ has a finite (maybe, empty) preimage. Hence π is quasi-finite. $\square$

Further on we shall consider the following sets:

$V' \subset \mathbb{A}^8$, defined by the ideal J;

$\widetilde{W} \subset \mathbb{A}^8$, defined by the ideal J3;

$U = V' \smallsetminus \mathbf{V}(xc) \subset \mathbb{A}^8$;

$U' = V' \smallsetminus \mathbf{V}(f) \subset \mathbb{A}^8$;

$W \subset \mathbb{A}^4$ with coordinates (a, c, v, x), defined by the ideal $\langle$J3(1), J3(2)$\rangle$.

$L = W \cap \mathbf{V}(f) \subset \mathbb{A}^4$;

$Z = W \smallsetminus L \subset \mathbb{A}^4$;

$Y = \mathbf{V}(J3[1]) \cap \mathbf{D}(f) \subset A^3$ with coordinates (a, c, v).

These affine sets are included in the following diagram:

$$\begin{array}{ccc}
\widetilde{W} \supset V' \supset U & \supset & U' \\
{\scriptstyle \pi_1}\downarrow & & \downarrow{\scriptstyle \pi_1} \\
W & \supset & Z \\
 & & \downarrow{\scriptstyle \pi_2} \\
 & & Y\,.
\end{array}$$

The inclusion $U \supset U'$ follows from computations: we have $V' \cap \mathbf{V}(x) \subset \mathbf{V}(f) \cap V'$. The map $\pi_1 : U' \to Z$ is a double unramified cover. This follows from the structure of equations $J3[1], \ldots, J3[6]$: all the branch points are contained in the set $\mathbf{V}(f)$. The map π_2 is an isomorphism since x appears linearly in the equation $J3[2]$ and its coefficient does not vanish in U'.

Proposition 3.11. $b^1(U) \leq 675$.

Proof. This estimate follows from the Weak Lefschetz Theorem proved by N. Katz ([Ka1, Cor. 3.4.1]). Indeed, we have:

- an algebraically closed field of characteristic $2 \le \ell$;
- U, a separated k-scheme of finite type which is a local complete intersection, purely of dimension $2 > 0$
- $U \to \mathbb{A}^2$, a quasi-finite morphism (see Proposition 3.10).

Then, for a constant $\mathbb{Q}_\ell$-sheaf $\mathcal{F}$ on U, there exists a dense open set $\mathcal{U} \subset \mathbb{A}^3$ such that for any $(\alpha, \beta, \gamma) \in \mathcal{U}$ the restriction map

$$H^1(U, \mathcal{F}) \to H^1\big(U \cap \{\alpha a + \beta c + \gamma = 0\}, i^*\mathcal{F}\big)$$

is injective (i denotes the embedding of the hyperplane section into U).

Denote:

$$\begin{aligned} S_1 &= U \cap \mathbf{V}(\alpha a + \beta c + \gamma); \\ \widetilde{S} &= S_1 \cap U' = U' \cap \mathbf{V}(\alpha a + \beta c + \gamma) \subset S_1; \\ S &= Y \cap \mathbf{V}(\alpha a + \beta c + \gamma) \subset Y\,. \end{aligned}$$

Since U' is a double unramified cover of Y, $\widetilde{S}$ is a double unramified cover of S. The curve S is defined in $\mathbb{A}^3$ with coordinates (a, c, v) by $\mathbf{V}(J3[1], \alpha a + \beta c + \gamma) \cap \mathbf{D}(f)$ with

```
J3[1]=(a8+a6c2+a4c4+a2c6)*v6+(a8+a7c3+a6c2+a5c3+a4c4+a3c7+a2c6
      +ac7)*v4+(a7c3+a6c2+a5c5+a5c3+a3c7+a3c5+a2c6+a2c4+ac9+c6)
      *v2+(ac9+ac5+c8+c4)=0;
```

Let $\overline{S}$ be the projectivisation of S in $\mathbb{P}^3$. For a general triple (α, β, γ) it is an irreducible complete intersection of degree $d = 14$. By [GL, Cor. 7.4], we have

$$b^1(\overline{S}) \le (d-1)(d-2) \le 156\,.$$

Let B be the union of the plane at infinity with the closure of the set $\mathbf{V}\big((\alpha a + \beta c + \gamma) f(a, c)\big)$. Since $\deg f = 11$, we have $\deg B = 13$. Thus $\overline{S} \cap B$ contains at most $14 \cdot 13 = 182$ points. Hence $b^1(S) \le 156 + 182 = 338$. Since $\widetilde{S}$ is a double unramified cover of S, $b^1(\widetilde{S}) = 2b^1(S) - 1 \le 675$. Since $\widetilde{S} \subset S_1$, $b^1(S_1) \le b^1(\widetilde{S}) \le 675$. ☐

Proposition 3.12. *The Euler characteristic of L satisfies $\chi(L) \le 71430 < 2^{17}$.*

Proof. The set $L = W \cap \mathbf{V}(f)$ consists of several components. According to computations, the list of components is as follows:

$$\begin{array}{lcllcllcr} F_1 &=& \mathbf{V}(a,c); & \dim F_1 &=& 2, & \chi(F_1) &=& 1 \\ F_2 &=& \mathbf{V}(v,c); & \dim F_2 &=& 2, & \chi(F_2) &=& 1 \\ F_3 &=& \mathbf{V}(v-1,c); & \dim F_3 &=& 2, & \chi(F_3) &=& 1 \\ F_4 &=& \mathbf{V}(a-1,c-1); & \dim F_4 &=& 2, & \chi(F_4) &=& 1 \\ E &=& \mathbf{V}(ac-1,v); & \dim E &=& 2, & \chi(E) &=& 0 \\ G &=& \mathbf{V}(ac-1, av^2+c^2+av+cv+v^2+v), & \dim G &=& 2, & \chi(G) &=& -3 \\ C_1 &=& \mathbf{V}(x, a, c^2+cv+1), & \dim C_1 &=& 1, & \chi(C_1) &=& 0 \\ C_2 &=& \mathbf{V}(c-1, v, x), & \dim C_2 &=& 1, & \chi(C_2) &=& 1 \\ C_3 &=& \mathbf{V}(I_3), & \dim C_3 &=& 1, & & & \\ H_1 &=& \mathbf{V}(I_1), & \dim H_1 &=& 2, & & & \\ H_2 &=& \mathbf{V}(I_2), & \dim H_2 &=& 2, & & & \end{array}$$

where
$I_3 = \langle c-1, a^2v^2x + a^2v + v^2x + av + ax + v + x, a^4x^2 + a^2vx^3 + a^3v^2 + a^3x^2 + a^4 + a^2vx + vx^3 + avx +$

$ax^2+a^2+vx+1, av^2x^4+v^5x+v^4x^2+v^2x^4+av^4+avx^3+v^4+a^2vx+a^2x^2+vx^3+x^4+avx+vx+x^2\rangle$,
$I_1=\langle c^3+c^2v+c^2+av+cv+v^2, acv+ac+c^2+av+v^2+a+c+v, a^2v+a^2+ac+cv+v^2+c\rangle$,
and
$I_2=\langle ac^2v+c^3v+c^3+c^2v+av^2+cv^2+cv+a, c^4+acv+c^2v+ac+cv+v+1, a^3v^2+a^2v^3+acv^3+c^3v+a^2v^2+acv^2+cv^2\rangle$.

Let us explain how the Euler characteristics were computed. We have $\chi(F_i)=1, i=1,\dots,4$ because the F_i's are just affine spaces. E is isomorphic to $(\mathbb{A}^1\setminus\{0\})$ with coordinates a and x respectively, so $\chi(E)=1\cdot(1-1)=0$. The component G is the direct product of $\mathbb{A}^1$ with coordinate x and a curve T which is a ramified covering of $\mathbb{A}^1$ with coordinate a. For a fixed point (a,c,v) in T we have $c=\frac{1}{a}$, and v is defined by the quadratic equation

$$v^2(a^3+a)+v(a^3+a^2+a)+1=0\,.$$

It follows that if $a\neq 0, a\neq 1, a^2+a+1\neq 0$, there are precisely two points in T with this value of a. There are no points with $a=0$ and precisely one point for each value $a=1$ or $a^2+a+1=0$. Since the Euler characteristics of $\mathbb{A}^1$ without 4 points is -3, we have $\chi(G)=2(-3)+3=-3$.

The curve C_1 is isomorphic to $\mathbb{A}^1\setminus\{0\}$ with coordinate c: $v=(1+c^2)/c$, $\chi(\mathbb{A}^1\setminus\{0\})=0$.

C_2 is $\mathbb{A}^1$, $\chi(\mathbb{A}^1)=1$.

In order to estimate the Euler characteristics of C_3, H_1, H_2, we use the following theorem of Adolphson and Sperber:

Proposition 3.13. [AS, Th. 5.27], [Ka2] *If an affine variety V is defined in $\mathbb{A}^N$ by r polynomial equations all of degree $\leq d$, then*

$$|\chi(V)|\leq 2^r D_{N,r}\underbrace{(1,1+d,\dots,1+d)}_{r+1}, \tag{3.10}$$

where $D_{N,r}(x_0,\dots,x_r)=\sum_{|W|=N}X^W$ is the homogeneous form of degree N in $x_0,\dots,x_r$ all of whose coefficients equal 1.

According to formula (3.10),

$|\chi(C_3)|\leq 2^3D_{3,3}(1,8,8,8)\leq 44232<2^{16}$

$|\chi(H_1)|\leq 2^3D_{3,3}(1,4,4,4)\leq 5992<2^{13}$

$|\chi(H_2)|\leq 2^3D_{3,3}(1,6,6,6)\leq 19160<2^{15}$.

The pairwise intersection of these components is a union of 16 lines and 10 points. The triple intersections contain 3 lines and 3 points. No four of these components intersect. Thus, $|\chi(L)|\leq 5+3+44232+5992+21224+26+6=71488<2^{17}$. □

Proposition 3.14. $b^2(U)\leq 2^{22}$.

Proof. We consider two cases:

(I) $\chi(U)\leq 0$. Then $1-b^1(U)+b^2(U)\leq 0$ and $b^2(U)\leq b^1(U)<675$.

(II) $\chi(U)>0$. We first find $|(\chi(U')|$. Since U' is a double cover of Z, we have $|\chi(U')|=2|\chi(Z)|$. Since $Z=W\setminus L$, we have $\chi(Z)=\chi(W)-\chi(L)$. By formula (3.10), we get $|\chi(W)|\leq 2^2D_{4,2}(1,15,15)\leq 1069324$. In view of Proposition 3.12, we have $|\chi(L)|\leq 71488$. Hence $|\chi(Z)|\leq|\chi(W)|+|\chi(L)|\leq 1140812$, and therefore $|\chi(U')|\leq 2281624<2^{22}$. On the other hand, $\chi(U)=\chi(U')+\chi(U\setminus U')$. In order to find $\chi(U)$, we have to evaluate $\chi(U\setminus U')$. Let $N=U\bigcap\mathbf{V}(f)$. Since N is the intersection of the smooth affine surface U with the hypersurface $\mathbf{V}(f)$, all of its irreducible components N_i are curves (that is, $\dim N_i=1$). This follows from [Sh, Th.5, p.74], and is confirmed by calculations. Since by Proposition 3.10 the projection $\pi\colon U\to\mathbb{A}^2$ is quasi-finite, none of N_i is mapped into a point. Hence $\pi(N_i)\subset\mathbb{A}^2$ is a curve. This curve does not meet the lines $\mathbf{V}(c)$ and $\mathbf{V}(a)$ because $\mathbf{V}(a)\cap U=\emptyset$. This means that the ring $O(\pi(N_i))$ contains the non-vanishing function ac. If $ac = \text{const}$ on $\pi(N_i)$, then $\pi(N_i)$ has two punctures at infinity. If $ac \neq \text{const}$, then the normalisation of $\pi(N_i)$ has at least two punctures, as does any curve having a non-constant and non-vanishing regular function. Thus $\chi(\pi(N_i)) \leq 0$.

Let k denote the degree of the map $N_i \to \pi(N_i)$. By Hurwitz's formula, $\chi(N_i) = k\chi(\pi(N_i)) - s$, where $s \geq 0$ is the branching number. It follows that

$$\chi(N_i) \leq k\chi\big(\pi(N_i)\big) \leq 0\,.$$

It follows that

$$\chi(\bigcup N_i) = \sum \chi(N_i) - T,$$

where $T = \sum_{x \in \bigcup N_i} (k(x) - 1) \geq 0$, and $k(x)$ stands for the multiplicity of a point $x \in \bigcup N_i$. Hence

$$\chi\left(\bigcup N_i\right) = \chi(N) \leq 0;$$

$$\chi(U) = \chi(U') + \chi(N) \leq \chi(U') \leq 2281624,$$

and, therefore,

$$b^2(U) = \chi(U) + b^1(U) \leq 2282299 < 2^{22}.$$

□

Corollary 3.15. *Let $n > 48$, $q = 2^n$. Then V_n has an $\mathbb{F}_q$-point.*

Proof. On plugging the estimates of Propositions 3.11 and 3.14 into formula (3.3), we see that $\#\,\mathrm{Fix}(U, n) > 0$ as soon as $n > 48$. This proves the corollary. □

3.5. **Small fields.** The purpose of this section is to study the fixed points and also numbers of fixed points of the operator α^n on the variety V' given by the equations $J[1], \ldots, J[10]$. Let k denote the algebraic closure of $\mathbb{F}_2$ and N_n the number of fixed points of α^n on $V'(k)$. As explained before, if n is even ($n = 2k$) then N_n is just the number of points of V' in the field $\mathbb{F}_{2^k}$. We are interested here in the numbers N_p for odd primes p.

We first give a table of the numbers N_n for $1 \leq n \leq 23$.

n	N_n	n	N_n	n	N_n	n	N_n
1	0	9	516	17	130084	25	33545220
2	8	10	1088	18	263504	26	67068464
3	12	11	2332	19	523260	27	134231772
4	16	12	3904	20	1050016	28	268394688
5	20	13	8372	21	2102420	29	536948340
6	56	14	16416	22	4198752	30	1073676416
7	140	15	32012	23	8378348	31	2221330252
8	240	16	65360	24	16788720	32	4295197328

TABLE 5. Fixed point numbers

We shall explain now the strategy for computing the numbers in Table 5. From the defining polynomials for V' it can be read off that the map $\psi\colon V' \to \mathbb{A}^2$ which maps a point to its coordinates (a, a_0) has finite fibers. Some special features of MAGMA make it then possible to count the number of elements in the fiber of every point of $\mathbb{A}^2$ over a given finite field.

From the above table we have tried to understand heuristically the zeta-function of α acting on V'. The zeta-function $Z(\alpha, T)$ of the operator α is defined by

$$Z(\alpha, T) := \exp\left(-\sum_{n=1}^{\infty} \frac{N_n}{n} T^n\right).$$

It is known to be a rational function and we find that it agrees with the polynomial

$$(1-2T)(1-T)(1-T^2)^5(1+T^2)^2(2T^4+2T^2+1)(4T^8+2T^4+1)(2T^2+2T+1)(8T^6+4T^5+T+1)$$

up to terms of degree T^{32}. Note that the absolute values of the zeros of this polynomial are equal to 1, $1/2$, $1/\sqrt{2}$, or $1/\sqrt[4]{2}$, as general theory predicts.

The operator α also acts on the singular set $S = V' \cap \mathbf{V}(cx)$. We have computed that the correspondingly defined zeta-function $Z_S(\alpha, T)$ agrees with $(1-2T^2)^3/[(1-T^2)^3(1+T^2)]$ up to terms of order T^{45}.

These two formulas imply the statement in Remark 1.12 above.

The data in Table 5 are not enough to close the gap between Corollary 3.15 and Theorem 1.2. We have used our method of computation to exhibit, for each $3 \le p \le 47$, a fixed point of α^p acting on V'. In the next table we give such points by their coordinates a, b, c, d, v, w, x, y in the field $\mathbb{F}_{2^p}$. Here t is a primitive element of $\mathbb{F}_{2^p}$ and MP is its minimal polynomial over $\mathbb{F}_2$. This finishes the proof of Theorem 1.1. Although it is quite difficult to find fixed points of α^p, it is easily checked, given the coordinates of point, whether it is a fixed point or not.

$p=3,$	$\mathrm{MP}=t^3+t+1$						
$a=1,$	$b=t^4,$	$c=t,$	$d=t^6,$	$v=1,$	$w=t^2,$	$x=t^4,$	$y=t^3.$
$p=5,$	$\mathrm{MP}=t^5+t^2+1$						
$a=t^5,$	$b=0,$	$c=t^{14},$	$d=1,$	$v=t^9,$	$w=0,$	$x=t^{19},$	$y=1.$
$p=7,$	$\mathrm{MP}=t^7+t+1$						
$a=t^5,$	$b=t^3,$	$c=t^{56},$	$d=t^{91},$	$v=t^{80},$	$w=t^{48},$	$x=t^7,$	$y=t^{59}.$
$p=11,$	$\mathrm{MP}=t^{11}+t^2+1$						
$a=t^3,$	$b=t^{228},$	$c=t^{151},$	$d=t^8,$	$v=t^{192},$	$w=t^{263},$	$x=t^{1476},$	$y=t^{512}.$
$p=13,$	$\mathrm{MP}=t^{13}+t^4+t^3+t+1$						
$a=t^9,$	$b=t^{2129},$	$c=t^{6077},$	$d=t^{7814},$	$v=t^{1152},$	$w=t^{2209},$	$x=t^{7902},$	$y=t^{890}.$
$p=17,$	$\mathrm{MP}=t^{17}+t^3+1$						
$a=t^5,$	$b=t^{39028},$	$c=t^{30333},$	$d=t^{16060},$	$v=t^{2560},$	$w=t^{59544},$	$x=t^{64118},$	$y=t^{96318}.$
$p=19,$	$\mathrm{MP}=t^{19}+t^5+t^2+t+1$						
$a=t,$	$b=t^{45681},$	$c=t^{503015},$	$d=t^{8107},$	$v=t^{1024},$	$w=t^{115801},$	$x=t^{237526},$	$y=t^{437263}.$

$p = 23, \quad \mathrm{MP} = t^{23} + t^5 + 1$
$a = t,$
$b = t^{22} + t^{21} + t^{19} + t^{16} + t^{15} + t^{12} + t^{10} + t^7 + t^4 + t,$
$c = t^{21} + t^{20} + t^{17} + t^{16} + t^{13} + t^{12} + t^{11} + t^{10} + t^6 + t^5 + t^3 + t,$
$d = t^{22} + t^{21} + t^{20} + t^{19} + t^{16} + t^{14} + t^{13} + t^{12} + t^{10} + t^9 + t^6 + t^5 + t^2,$
$v = t^{22} + t^{21} + t^{20} + t^{19} + t^{17} + t^{15} + t^{13} + t^{12} + t^{10} + t^7 + t^5 + t^4,$
$w = t^{22} + t^{18} + t^{15} + t^{13} + t^{12} + t^8 + t^7 + t^5 + t,$
$x = t^{21} + t^{19} + t^{18} + t^{15} + t^{13} + t^{11} + t^9 + t^8 + t^6 + t^4 + t^2,$
$y = t^{19} + t^{18} + t^{15} + t^9 + t^7 + t^6 + t^3 + t^2 + t.$

$p = 29, \quad \mathrm{MP} = t^{29} + t^2 + 1$
$a = t^2 + t,$
$b = t^{28} + t^{27} + t^{26} + t^{25} + t^{24} + t^{23} + t^{22} + t^{19} + t^{16} + t^{13} + t^{12} + t^9 + t^8 + t^7 + t^6 + t^5 + t^4 + 1,$
$c = t^{25} + t^{23} + t^{20} + t^{19} + t^{17} + t^{16} + t^{15} + t^{14} + t^{13} + t^9 + t + 1,$
$d = t^{26} + t^{24} + t^{23} + t^{22} + t^{20} + t^{17} + t^{16} + t^{13} + t^{12} + t^6,$
$v = t^{28} + t^{26} + t^{20} + t^{18} + t^{17} + t^{11} + t^{10} + t^8 + t^4 + t^3 + t,$
$w = t^{23} + t^{19} + t^{17} + t^{16} + t^{15} + t^{14} + t^{13} + t^{12} + t^{11} + t^8 + t^6 + t^3 + t^2,$
$x = t^{26} + t^{25} + t^{19} + t^{16} + t^{13} + t^{11} + t^9 + t^8 + t^7 + t^6 + t^4 + 1,$
$y = t^{27} + t^{24} + t^{23} + t^{22} + t^{21} + t^{15} + t^{13} + t^9 + t^8 + t^6 + t^5 + t^4 + 1,$

$p = 31, \quad \mathrm{MP} = t^{31} + t^3 + 1$
$a = t^3,$
$b = t^{30} + t^{27} + t^{25} + t^{23} + t^{21} + t^{20} + t^{18} + t^{17} + t^{16} + t^{15} + t^{14} + t^{12} + t^{11} + t^9 + t^8 + t^5 + t^3 + t,$
$c = t^{27} + t^{24} + t^{20} + t^{19} + t^{18} + t^{16} + t^{14} + t^{12} + t^{10} + t^9 + t^8 + t^7 + t^5 + t^4,$
$d = t^{23} + t^{21} + t^{20} + t^{19} + t^{13} + t^8 + t^7 + t^6,$
$v = t^{24} + t^{10} + t^9 + t^2 + t,$
$w = t^{29} + t^{24} + t^{23} + t^{21} + t^{20} + t^{17} + t^{16} + t^{15} + t^{11} + t^9 + t^4 + t^3 + t^2 + t,$
$x = t^{30} + t^{28} + t^{26} + t^{25} + t^{24} + t^{20} + t^{18} + t^{17} + t^{16} + t^{15} + t^{11} + t^9 + t^8 + t^7 + t^5 + t,$
$y = t^{30} + t^{29} + t^{28} + t^{19} + t^{17} + t^{16} + t^{13} + t^{11} + t^{10} + t^8 + t^6 + t^5 + t^4 + t^3 + t.$

$p = 37, \quad \mathrm{MP} = t^{37} + t^5 + t^4 + t^3 + t^2 + t + 1$
$a = t^2 + t,$
$b = t^{36} + t^{34} + t^{30} + t^{29} + t^{28} + t^{27} + t^{26} + t^{25} + t^{23} + t^{21} + t^{17} + t^{14} + t^{11} + t^{10} + t^7 + t^4 + t^3 + t^2 + t + 1,$
$c = t^{36} + t^{35} + t^{33} + t^{32} + t^{31} + t^{29} + t^{19} + t^{14} + t^{11} + t^{10} + t^9 + t^7 + t^2,$
$d = t^{36} + t^{34} + t^{32} + t^{31} + t^{27} + t^{26} + t^{24} + t^{23} + t^{22} + t^{20} + t^{19} + t^{18} + t^{16} + t^{12} + t^{10} + t^9 + t^8 + t^7 + t^6 + t^4 + t^3 + t + 1,$
$v = t^{34} + t^{33} + t^{31} + t^{30} + t^{29} + t^{28} + t^{27} + t^{26} + t^{19} + t^{18} + t^{17} + t^{16} + t^{14} + t^{13} + t^{10} + t^8 + t^7 + t^6 + t^5 + t^4 + t^3 + t^2 + t + 1,$
$w = t^{36} + t^{35} + t^{34} + t^{33} + t^{32} + t^{29} + t^{26} + t^{25} + t^{20} + t^{19} + t^{18} + t^{17} + t^{15} + t^{14} + t^{12} + t^{10} + t^8 + t^7 + t^6 + t^5 + t^4 + t^2 + t + 1,$
$x = t^{36} + t^{31} + t^{29} + t^{28} + t^{27} + t^{24} + t^{21} + t^{19} + t^{18} + t^{16} + t^{15} + t^{14} + t^{12} + t^{10} + t^9 + t^8 + t^3 + t^2 + t,$
$y = t^{35} + t^{33} + t^{32} + t^{27} + t^{26} + t^{25} + t^{24} + t^{23} + t^{19} + t^{18} + t^{15} + t^{11} + t^{10} + t^9 + t^7 + t^4 + t^3 + t + 1.$

$p = 41$, $\quad \mathrm{MP} = t^{41} + t^3 + 1$

$a = t,$

$b = t^{39} + t^{37} + t^{36} + t^{35} + t^{34} + t^{33} + t^{30} + t^{28} + t^{26} + t^{24} + t^{23} + t^{20} + t^{19} + t^{18} + t^{17} + t^{16} + t^{15} + t^{13} + t^{11} + t^{10} + t^9 + t^6 + t^3 + t^2 + t + 1,$

$c = t^{40} + t^{38} + t^{36} + t^{33} + t^{32} + t^{31} + t^{29} + t^{28} + t^{27} + t^{26} + t^{24} + t^{23} + t^{22} + t^{19} + t^{18} + t^{14} + t^{12} + t^{10}t^9 + t^6 + t^7 + t^3 + 1,$

$d = t^{40} + t^{39} + t^{37} + t^{35} + t^{34} + t^{30} + t^{28} + t^{27} + t^{23} + t^{21} + t^{19} + t^{18} + t^{15} + t^{14} + t^{13} + t^{11} + t^9 + t^7 + t^4 + t^2,$

$v = t^{38} + t^{37} + t^{36} + t^{35} + t^{34} + t^{32} + t^{31} + t^{29} + t^{26} + t^{23} + t^{21} + t^{20} + t^{18} + t^{17} + t^{15} + t^{14} + t^{13} + t^{12} + t^{11} + t^{10} + t^7 + t^6 + t^4 + t^3 + t^2 + t,$

$w = t^{40} + t^{37} + t^{35} + t^{33} + t^{32} + t^{31} + t^{30} + t^{29} + t^{28} + t^{27} + t^{26} + t^{25} + t^{24} + t^{20} + t^{19} + t^{18} + t^{16} + t^{14} + t^{13} + t^9 + t^8 + t^7 + t^6 + t^4 + t^2 + 1,$

$x = t^{38} + t^{36} + t^{34} + t^{33} + t^{32} + t^{29} + t^{21} + t^{20} + t^{18} + t^{13} + t^{12} + t^7 + t^6 + t^2 + t + 1,$

$y = t^{40} + t^{39} + t^{35} + t^{26} + t^{23} + t^{22} + t^{19} + t^{17} + t^{16} + t^{15} + t^{14} + t^{13} + t^7 + t^5 + t^4 + t^2.$

$p = 43$, $\quad \mathrm{MP} = t^{43} + t^6 + t^4 + t^3 + 1$

$a = t^3,$

$b = t^{42} + t^{39} + t^{37} + t^{32} + t^{31} + t^{30} + t^{29} + t^{26} + t^{23} + t^{19} + t^{18} + t^{17} + t^{10} + t^7 + t^6 + t^5 + t^4 + 1,$

$c = t^{40} + t^{39} + t^{38} + t^{36} + t^{35} + t^{34} + t^{33} + t^{32} + t^{31} + t^{30} + t^{26} + t^{23} + t^{22} + t^{21} + t^{19} + t^{16} + t^{14} + t^{13} + t^{12} + t^9 + t^4 + t^3 + t^2 + 1,$

$d = t^{39} + t^{38} + t^{37} + t^{36} + t^{35} + t^{34} + t^{29} + t^{28} + t^{27} + t^{22} + t^{21} + t^{20} + t^{19} + t^{18} + t^{17} + t^{16} + t^{13} + t^9 + t^7 + t^6 + t^4 + t^3 + t^2 + +1,$

$v = t^{40} + t^{37} + t^{35} + t^{33} + t^{31} + t^{30} + t^{26} + t^{17} + t^{16} + t^{15} + t^{11} + t^8 + t^4 + t + 1,$

$w = t^{41} + t^{39} + t^{38} + t^{37} + t^{36} + t^{31} + t^{29} + t^{28} + t^{24} + t^{21} + t^{18} + t^{17} + t^{14} + t^9 + t^8 + t^7 + t^5 + t^4 + t^3 + 1,$

$x = t^{41} + t^{39} + t^{34} + t^{33} + t^{32} + t^{29} + t^{25} + t^{24} + t^{20} + t^{18} + t^{17} + t^{16} + t^{15} + t^{12} + t^{11} + t^{10} + t^9 + t^7 + t^6 + t^5 + t^4 + t + 1,$

$y = t^{41} + t^{40} + t^{38} + t^{37} + t^{35} + t^{34} + t^{33} + t^{30} + t^{29} + t^{28} + t^{27} + t^{26} + t^{24} + t^{19} + t^{18} + t^{16} + t^{15} + t^{12} + t^{11} + t^9 + t^6 + t^4 + t^3 + t^2 + t,$

$p = 47$, $\quad \mathrm{MP} = t^{47} + t^5 + 1$

$a = t^2 + t + 1,$

$b = t^{46} + t^{44} + t^{43} + t^{41} + t^{40} + t^{39} + t^{37} + t^{36} + t^{35} + t^{34} + t^{27} + t^{25} + t^{24} + t^{23} + t^{22} + t^{20} + t^{19} + t^{17} + t^{16} + t^{14} + t^{13} + t^{12} + t^{11} + t^9 + t^8 + t^7 + t^6 + t^5 + t^2,$

$c = t^{40} + t^{33} + t^{31} + t^{30} + t^{29} + t^{26} + t^{25} + t^{24} + t^{23} + t^{21} + t^{19} + t^{16} + t^{15} + t^{14} + t^{11} + t^{10} + t^9 + t^8 + t^7 + t^6,$

$d = t^{44} + t^{42} + t^{41} + t^{39} + t^{36} + t^{35} + t^{31} + t^{27} + t^{26} + t^{24} + t^{20} + t^{19} + t^{17} + t^{16} + t^{15} + t^{13} + t^9 + t^7 + t^5 + t^4 + t^2 + t,$

$v = t^{44} + t^{41} + t^{40} + t^{38} + t^{37} + t^{34} + t^{33} + t^{32} + t^{31} + t^{28} + t^{27} + t^{26} + t^{25} + t^{23} + t^{22} + t^{18} + t^{17} + t^{16} + t^{12} + t^{10} + t^9 + t^7 + t^6 + t^5 + t^2 + t + 1,$

$w = t^{45} + t^{43} + t^{42} + t^{41} + t^{39} + t^{38} + t^{36} + t^{35} + t^{34} + t^{32} + t^{30} + t^{24} + t^{23} + t^{21} + t^{20} + t^{16} + t^{12} + t^9 + t^5 + t^4 + t^3 + t^2 + t,$

$x = t^{46} + t^{45} + t^{44} + t^{41} + t^{39} + t^{37} + t^{35} + t^{33} + t^{31} + t^{30} + t^{29} + t^{27} + t^{26} + t^{25} + t^{21} + t^{20} + t^{18} + t^{17} + t^{11} + t^9 + t^8 + t^7 + t^6 + t^5 + t^3 + t^2,$

$y = t^{46} + t^{42} + t^{41} + t^{40} + t^{39} + t^{38} + t^{37} + t^{35} + t^{34} + t^{33} + t^{32} + t^{31} + t^{30} + t^{28} + t^{26} + t^{25} + t^{22} + t^{15} + t^{13} + t^{11} + t^{10} + t^9 + t^6 + t^5 + t^4 + t^2.$

TABLE 6. Fixed points in $\mathbb{F}_{2^p}$

4. Appendix

4.1. A variant of Zorn's theorem.

In this appendix we prove

Proposition 4.1. *Let G be a finite group, and let $w = w(x, y)$ be a word in two variables such that: 1) if $w(x, y) \equiv 1$ in G then $G = \{1\}$; 2) the words x and $w(x, y)$ generate the free group $F_2 = \langle x, y\rangle$. Then G is nilpotent if and only if it satisfies one of the identities $[w(x, y), x, x, \dots, x] = 1$.*

Proof. Necessity. Let G be a nilpotent group of class n. Since the element $e_n(x, y) = [w(x, y), x, \dots, x]$ lies in the nth term of the invariant series, $e_n(x, y)$ is an identity.

Sufficiency. We want to prove that any G satisfying the identity $e_n(x, y) \equiv 1$ for some n is nilpotent. Assume the contrary.

Suppose that $n = 1$. Then according to assumption (1) of the proposition, the group G is trivial. Let $n > 1$. Let Γ denote a minimal counterexample, that is, a non-nilpotent group of the smallest order satisfying the identity $e_n(x, y) \equiv 1$. Obviously, all subgroups of Γ are nilpotent. Then Γ is a Schmidt group, that is, a non-nilpotent group all of whose proper subgroups are nilpotent (see [Sch], [Re] for the description of these groups). In particular, the commutator subgroup Γ' is the unique maximal Sylow subgroup in Γ. Since Γ' is nilpotent, it contains a non-trivial center $Z(\Gamma')$. Take a nontrivial $a \in Z(\Gamma')$. For any element $x \notin \Gamma'$ there exists $y \in G$ such that $w(x, y) = a$ (condition (2)). Consider the sequence $[a, x, x, \dots, x] = [w(x, y), x, x, \dots, x] = e_n(x, y)$. There exists n such that $[w(x, y), x, x, \dots, x] \equiv 1$. Let n denote the smallest number satisfying this equality, and let $b = [w(x, y), x, x, \dots, x] = e_{n-1}(x, y)$. Clearly, $b \in Z(\Gamma')$. Moreover, $[b, x] = e_n(x, y) = 1$ and hence b is a nontrivial element from $Z(\Gamma)$. Take $\bar\Gamma = \Gamma/Z(\Gamma)$. Then the order of $\bar\Gamma$ is less than the order of Γ, hence $\bar\Gamma$ is nilpotent. Therefore Γ is nilpotent. Since $e_n(x, y)$ is an identity in Γ, we get a contradiction.

The proposition is proved. □

4.2. Pro-finite setting.

4.2.1. *Pseudo-varieties of finite groups.* A *variety* of groups is a class C of groups defined by some set of identities T (that is, $G \in C$ if and only if for every $u \in T$ the identity u holds in G). Birkhof's theorem says that C is a variety if and only if C is closed under taking subgroups, homomorphic images, and direct products. To work with classes of finite groups (which cannot be closed under taking infinite direct products), one needs a more general notion.

Definition 4.2. A *pseudo-variety* of groups is a class of groups closed under taking subgroups, homomorphic images, and **finite** direct products.

By Birkhoff's theorem every variety of groups is a pseudo-variety. We are interested in pseudo-varieties of all finite groups, all finite solvable groups, and all finite nilpotent groups.

Let $F = F(X^0)$ be a free group with countable set of generators X^0. Consider a sequence of words $u = u_1, u_2, \dots, u_n, \dots$ in F. The sequence u determines a class of groups V_u by the rule: a group G belongs to V_u if and only if almost all elements u are identities in G. The class V_u is a pseudo-variety. It turns out that this construction is universal:

Theorem 4.3. [ES] *For every pseudo-variety of finite groups V there exists a sequence of elements $u\colon \mathbb{N} \to F$, $u = u_1, u_2, \dots, u_n, \dots$ such that $V = V_u$.*

We shall consider a special class of sequences.

Definition 4.4. Let X be a finite set. We say that a sequence of elements (not necessarily distinct) $u = u_1, u_2, \dots, u_n, \dots$ of the free group $F(X)$ is *correct* if given any group G, as soon as an identity $u_n \equiv 1$ holds in G, for all $m > n$ the identities $u_m \equiv 1$ hold in G, too.

As above, a correct sequence u defines a pseudo-variety of groups V by the rule: $G \in V$ if and only if some identity $u_n \equiv 1$, $u_n \in u$, holds in G.

Remark 4.5. If u is a correct sequence defining a pseudo-variety V and v is a subsequence of u, then v is also correct and defines the same pseudo-variety V.

Let $F = F(x, y)$ and

$$e_1 = [x, y], \quad e_{n+1} = [e_n, y\ ], \dots \tag{4.1}$$

This sequence is correct and defines the pseudo-variety of all finite Engel groups. According to Zorn's theorem [Zo], this pseudo-variety coincides with the pseudo-variety of all finite nilpotent groups.

Our main sequence of quasi-Engel words

$$u_1 = w = x^{-2}y^{-1}x, \quad u_{n+1} = [x\, u_n\, x^{-1}, y\, u_n\, y^{-1}], \dots \tag{4.2}$$

is also correct, and according to Theorem 1.1 it defines the pseudo-variety of all finite solvable groups.

4.2.2. *Residually finite groups.*

Definition 4.6. We say that a group G is *residually finite* if the intersection of all its normal subgroups of finite index H_α, $\alpha \in I$, is trivial.

Define a partial order on the set I by: $\alpha < \beta$ if and only if $H_\beta \subset H_\alpha$. The intersection of two normal subgroups of finite index is also of finite index, and therefore for every $\alpha, \beta \in I$ there is $\gamma \in I$ such that $\alpha < \gamma$, $\beta < \gamma$. Thus the set I is directed.

Denote $G_\alpha = G/H_\alpha$. If $\alpha < \beta$ then there is a natural homomorphism $\varphi_\alpha^\beta : G_\beta \to G_\alpha$. If gH_β is an element of G_β then its image in G_α is gH_α. Let $\bar{G}$ be the direct product of all G_α. Then there is an embedding $G \to \bar{G}$ which associates to each $g \in G$ the element $\bar{g} = (gH_\alpha)_{\alpha \in I}$. Hence G can be approximated by finite groups G_α, that is, if f, g are distinct elements of G then there is α such that $\bar{f}_\alpha$ and $\bar{g}_\alpha$ are distinct elements of G_α.

A group G is regarded as a topological group, with the topology defined by the system of neighbourhoods of 1 consisting of all normal subgroups of finite index H_α. The system of neighbourhoods of an element $g \in G$ is given by the cosets gH_α. The group $\bar{G}$ is also a topological group. To define the topology, consider the projections $\pi_\alpha \colon \bar{G} \to G_\alpha$. Let $\ker \pi_\alpha = U_\alpha$. Then $\bar{G}/U_\alpha$ is isomorphic to $G_\alpha = G/H_\alpha$. For every $g \in G$ the element $\bar{g}$ lies in U_α if and only if $g \in H_\alpha$. The system of neighbourhoods of 1 in $\bar{G}$ consists of all finite intersections of normal subgroups U_α. This defines the Tikhonov topology on $\bar{G}$. Since all groups G_α are finite, the group $\bar{G}$ is compact.

Let $g_1, \dots, g_n, \dots$ be a sequence of elements of G. As usual, we say that this sequence tends to 1 if for every neighbourhood H_α there exists a natural number $N = N(\alpha)$ such that for all $n > N$ the element g_n lies in H_α.

Definition 4.7. Let $F = F(X)$ be a free group. We say that a sequence $u = u_1, \dots, u_n, \dots$ of elements of F *identically converges to 1* in a group G if for any homomorphism $\mu \colon F \to G$ the sequence $\mu(u) = \mu(u_1), \dots, \mu(u_n), \dots$ tends to 1 in G. In this case we write $u \equiv 1$ in G.

Proposition 4.8. *Let X be a finite set. If a sequence $u_1, \dots, u_n, \dots$ identically converges to 1 in G then for every neighbourhood H_α there exists $N = N(\alpha)$ such that all $u_n, n > N$, are identities of the group G/H_α.*

Proof. Take a homomorphism $\mu \colon F \to G$, and let $\mu^0 \colon G \to G/H_\alpha$ be the natural projection. Then $\nu = \mu^0\mu$ is a homomorphism $F \to G/H_\alpha$, and every homomorphism $\nu \colon F \to G/H_\alpha$ can be represented in this way. Since both G/H_α and X are finite, the set of different ν's is also finite. Denote them $\{\nu_1, \dots, \nu_k\}$.

Define an equivalence relation on the set of all homomorphisms $\mu\colon F \to G$ by: $\mu_1 \equiv \mu_2$ if $\mu^0\mu_1 = \mu^0\mu_2$. For an arbitrary $u \in F$ we have $\mu(u) \in H_\alpha$ if and only if $\mu^0\mu(u) = 1$. Thus, if $\mu_1 \equiv \mu_2$ then for every $u \in F$ we have $\mu_1(u) \in H_\alpha$ if and only if $\mu_2(u) \in H_\alpha$. Indeed, let $\mu_1(u) \in H_\alpha$. Then $\mu^0\mu_1(u) = 1 = \mu^0\mu_2(u) = 1$, and $\mu_2(u) \in H_\alpha$.

For every ν_i, $i = 1, \ldots, k$ take μ_i such that $\mu^0\mu_i = \nu_i$. Consider the equivalence classes $[\mu_1], \ldots, [\mu_k]$. Each $\mu\colon F \to G$ belongs to one of these classes. Since the sequence $u_1, \ldots, u_n, \ldots$ identically converges to 1 in G, for every $\mu\colon F \to G$ there exists $N = N(\alpha, \mu)$ such that $\mu(u_n) \in H_\alpha$ for $n > N$. Let N_0 be the maximum of $N(\alpha, \mu_i)$, $i = 1, \ldots, k$. If $n > N_0$ then $\mu_i(u_n) \in H_\alpha$ for every μ_i. Since every μ is equivalent to some μ_i, we have $\mu(u_n) \in H_\alpha$ for every μ. This means that $\nu(u_n) = 1$ for every $\nu : F \to G/H_\alpha$. Thus the element u_n defines an identity of the group G/H_α. □

4.2.3. *Pro-finite groups.* We now focus on pro-finite groups, with a goal to establish a relationship with pseudo-varieties and give another reformulation of our main result. Generalities on pro-finite groups can be found in [RZ], [Al1], etc. We recall here some basic notions.

Let V be a pseudo-variety of finite groups. Given a group G, consider all its normal subgroups of finite index H_α such that $G/H_\alpha = G_\alpha \in V$. If the intersection of all these H_α is trivial, we say that G is a *residually V-group.* This is a topological group with V-topology (the subgroups H_α as above are taken as the neighbourhoods of 1).

Let $\bar{G}$ be the direct product of all G_α. Denote by $\widehat{G}$ a subgroup in $\bar{G}$ defined as follows: an element $f \in \bar{G}$ belongs to $\widehat{G}$ if and only if for every α and β such that $H_\beta \subset H_\alpha$ the equality $\varphi_\alpha^\beta(f_\beta) = f_\alpha$ holds. Denote $f_\alpha = g_\alpha H_\alpha$. Then
$$\varphi_\alpha^\beta(g_\beta H_\beta) = g_\alpha H_\alpha = g_\beta H_\alpha\,.$$
Recall that φ_α^β are natural homomorphisms.

The group $\widehat{G}$ turns out to be the completion of G in its V-topology [ESt].

Such a group $\widehat{G}$ is called a *pro-V-group.* If V is the pseudo-variety of all finite groups, $\widehat{G}$ is called *a pro-finite group.* Thus in the class of all pro-finite groups one can distinguish subclasses related to particular pseudo-varieties V.

A free group $F = F(X)$ is residually finite. Take all normal subgroups of finite index in F. They define the pro-finite topology in F. Denote by $\widehat{F}$ the completion of F in this topology. This group is a free pro-finite group (see, for example, [RZ]).

Indeed, if $\widehat{G}$ is the pro-finite completion of an arbitrary residually finite group G, then every map $\mu\colon X \to \widehat{G}$ induces a homomorphism $\mu\colon F \to \widehat{G}$ which turns out to be a continuous homomorphism of topological groups and therefore induces a continuous homomorphism $\hat{\mu}\colon \widehat{F} \to \widehat{G}$.

Another approach to free pro-finite groups is based on the idea of implicit operations (cf. [Al1], [Al2], [AV], [MSW], [We], etc.). This approach has a lot of advantages but we do not use it since it needs additional notions which are not necessary for our aims.

Definition 4.9. Let $f \in \widehat{F}$. The expression $f \equiv 1$ is called a *pro-finite identity* of a pro-finite group $\widehat{G}$ if for every continuous homomorphism $\widehat{\mu}\colon \widehat{F} \to \widehat{G}$ we have $\widehat{\mu}(f) = 1$.

Definition 4.10. (see also [AV], [Al1]) A variety of pro-finite groups (for brevity, a *pro-variety*) is a class of pro-finite groups defined by some set of pro-finite identities.

An analogue of Birkhoff's theorem for pro-finite groups says that a class of pro-finite groups is a pro-variety if and only if it is closed under taking closed subgroups, images under continuous homomorphisms, and direct products. This implies that for an arbitrary pseudo-variety V of finite groups, the class of all pro-V-groups is a pro-variety. The converse statement is also true. For any pro-variety C there exists a pseudo-variety of finite groups V such that the class of all pro-V-groups coincides with C. In the case where V is a correct pseudo-variety of finite groups (that is, is defined by a correct sequence), one can construct identities defining the pro-variety of pro-V-groups in an explicit form.

Let X be a finite set. Let $u = u_1, \dots, u_n, \dots$ be a sequence of elements of a free group $F = F(X)$. Since $\widehat{F}$ is a compact group, there exists a convergent subsequence $v = v_1, \dots, v_m, \dots$ of u.

Proposition 4.11. *Let $v = v_1, v_2, \dots, v_n, \dots$ be a convergent sequence of elements of F with $\lim \bar{v}_n = f$. Let $\widehat{G}$ be a pro-finite group. Then the identity $f \equiv 1$ holds in $\widehat{G}$ if and only if $v \equiv 1$ in G (that is, v identically converges to 1 in G, see Definition 4.7).*

Proof. First of all the sequence $\bar{g}_1, \dots, \bar{g}_n, \dots$ converges to 1 in $\widehat{G}$ if and only if $g_1, \dots, g_n, \dots$ converges to 1 in G.

Let the identity $f \equiv 1$ be fulfilled in $\widehat{G}$. Then

$$\widehat{\mu}(f) = \lim \widehat{\mu}(\bar{v}_n) = \lim \overline{\mu(v_n)} = 1\,.$$

Thus, $\lim \mu(v_n) = 1$ in G. This means that $v \equiv 1$ in G. Conversely, let $v \equiv 1$ in G. Then for every $\mu \colon F \to G$ the sequence $\mu(v)$ converges to 1 in G. The sequence $\overline{\mu(v)}$ converges to 1 in $\widehat{G}$. Using

$$\lim \widehat{\mu}(\bar{v}_n) = \lim \overline{\mu(v_n)} = 1 = \widehat{\mu}(f)\,,$$

we conclude that $\widehat{\mu}(f) = 1$ for arbitrary μ. This means that the identity $f \equiv 1$ holds in $\widehat{G}$. ☐

Let V be a pseudo–variety of finite groups defined by a correct sequence $u = u_1, u_2, \dots, u_n, \dots$, and let $v = v_1, v_2, \dots, v_n, \dots$ be a convergent subsequence of u. Denote the limit of v by f. Since u is a correct sequence, v determines the same class V as u.

Theorem 4.12. *With the above notation, the class of all pro-V-groups is the pro-variety defined by the pro-finite identity $f \equiv 1$.*

Proof. Let the pro-finite identity $f \equiv 1$ hold in a pro-finite group $\widehat{G}$. Then by Proposition 4.11, $v \equiv 1$ in G. Proposition 4.8 implies that for every neighbourhood H in G and all sufficiently large n the identity $v_n \equiv 1$ holds in G/H. This means that G/H lies in V and $\widehat{G}$ is a pro-V-group.

Conversely, let G/H lie in V. By the definition of V, this means that v identically converges to 1 in G. Therefore, the identity $f \equiv 1$ holds in $\widehat{G}$. ☐

Remark 4.13. Although all convergent subsequences of a correct sequence define the same pseudo-variety, their limits may be different. For example, consider a correct sequence of the form $u = v_1, av_1a^{-1}, v_2, av_2a^{-1}, \dots, v_n, av_na^{-1}, \dots$, where $a \in F$ and $v = v_1, v_2, \dots, v_n, \dots$ is a correct convergent sequence. If the limit of the subsequence v is f, we get a new convergent subsequence $v' = av_1a^{-1}, av_2a^{-1}, \dots, av_na^{-1}, \dots$ with limit afa^{-1}. However, the elements f and afa^{-1} define the same variety.

Corollary 4.14. *Let $F = F(x, y)$, and let u_n be defined by*

$$u_1 = w, \quad u_{n+1} = [u_n, y], \dots \tag{4.3}$$

where $w = [x, y]$ or w is any word satisfying the conditions the hypotheses of Proposition 4.1.

Let $v_1, v_2, \dots, v_m, \dots$ be any convergent subsequence of (4.3) *with limit f from $\widehat{F}$. Then the identity $f \equiv 1$ defines the pro-finite variety of pro-nilpotent groups.*

Proof. The corollary immediately follows from Proposition 4.1, Zorn's theorem, and Theorem 4.12. ☐

Theorem 4.15. *Let $F = F(x, y)$, let*

$$u_1 = w = x^{-2}yx^{-1}, \quad u_{n+1} = [xu_nx^{-1}, yu_ny^{-1}], \dots \tag{4.4}$$

be our main sequence, and let $v_1, v_2, \dots, v_m, \dots$ be any convergent subsequence of (4.4) *with limit f from $\widehat{F}$. Then the identity $f \equiv 1$ defines the pro-finite variety of pro-solvable groups.*

Proof. The theorem immediately follows from Theorems 1.1 and 4.12. □

We can now state the pro-finite analogue of the Thompson–Flavell theorem.

Corollary 4.16. *A pro-finite group G is pro-solvable if and only if every closed two-generator subgroup of G is pro-solvable.*

Proof. Let every two-generator subgroup of $\widehat{G}$ be pro-solvable. Take an element $f \in \widehat{F}(x, y)$ which is the limit of a convergent subsequence of our sequence u. Let μ be an arbitrary continuous homomorphism $\widehat{F}(x, y) \to \widehat{G}$. Then $\mu(f) = 1$ since $\mu(f)$ lies in a two-generator subgroup of $\widehat{G}$. This is true for arbitrary μ and, therefore, $f \equiv 1$. According to Theorem 4.15, $\widehat{G}$ is pro-solvable. □

Corollary 4.14 and Theorem 4.15 should be compared with results of J. Almeida [Al2]. He used the language of implicit operations and the notion of $n!$-type convergent subsequence to get nice proofs of theorems of similar type. He also noticed that if our main theorem about solvable groups is true for the sequence u_n^w with initial term $w = [x, y]$, then the $n!$ version of the corresponding statement is also true.

References

[AS] A. Adolphson and S. Sperber, *On the degree of the L-functions associated with an exponential sum*, Compositio Math. **68** (1998), 125–159.

[Al1] J. Almeida, *Finite Semigroups and Universal Algebra*, World Scientific, 1994.

[Al2] J. Almeida, *Dynamics of implicit operations and tameness of pseudovarieties of groups*, Trans. Amer. Math. Soc. **354** (2002), 387–411.

[AV] J. Almeida and M. Volkov, *Profinite methods in semigroup theory*, Preprint CMUP 2001-02.

[AP] Y. Aubry and M. Perret, *A Weil theorem for singular curves*, In: "Arithmetic, Geometry and Coding Theory", R. Pellikaan, M. Perret, and S. G. Vlăduţ (eds.), Walter de Gruyter, Berlin–New York, 1996, pp. 1–7.

[BP] M. Boffa and F. Point, *Identités de Thue–Morse dans les groupes*, C.R. Acad. Sci. Paris, Sér. I **312** (1991), 667–670.

[Bo] E. Bombieri, *Thompson's problem* $\sigma^2 = 3$, Invent. Math. **58** (1980), 77–100.

[Br] R. Brandl, *Zur Theorie der untergruppenabgeschlossenen Formationen: endliche Varietäten*, J. Algebra **73** (1981), 1–22.

[BN] R. Brandl and D. Nikolova, *Simple groups of small Engel depth*, Bull. Austral. Math. Soc. **33** (1986), 245–251.

[BW] R. Brandl and J. S. Wilson, *Characterization of finite soluble groups by laws in a small number of variables*, J. Algebra **116** (1988), 334–341.

[Bu] B. Buchberger, *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*, PhD Thesis, Univ. of Innsbruck, Austria, 1965.

[BM] R. G. Burns and Yu. Medvedev, *A note on Engel groups and local nilpotence*, J. Austral. Math. Soc. Ser. A **64** (1998), 92–100.

[De] P. Deligne, *La conjecture de Weil II*, Inst. Hautes Études Scient. Publ. Math. **52** (1981), 313–428.

[DL] P. Deligne and G. Lusztig, *Representations of reductive groups over finite fields*, Ann. of Math. (2) **103** (1976), 103–161.

[ES] S. Eilenberg and S. Schützenberger, *On pseudovarieties*, Adv. Math. **19** (1976), 413–418.

[ESt] S. Eilenberg and N. Steenrod, *Foundations of algebraic topology*, Princeton Univ. Press, 1952.

[Fl] P. Flavell, *Finite groups in which every two elements generate a soluble group*, Invent. Math. **121** (1995), 279–285.

[FJ] M. Fried and M. Jarden, *Field Arithmetic*, Springer-Verlag, Berlin, 1986.

[Fu] K. Fujiwara, *Rigid geometry, Lefschetz–Verdier trace formula and Deligne's conjecture*, Invent. Math. **127** (1997), 480–533.

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

[GP1] G.-M. Greuel and G. Pfister, *Advances and improvements in the theory of standard bases and syzygies*, Arch. Math. **66** (1996), 163–176.

[GP2] G.-M. Greuel and G. Pfister, *Gröbner bases and algebraic geometry*, In: "Gröbner Bases and Applications", B. Buchberger and F. Winkler (eds.), Lecture Notes Ser. **251**, Cambridge Univ. Press, 1998, pp. 109–143.

[GP3] G.-M. Greuel and G. Pfister, *A* SINGULAR *Introduction to Commutative Algebra*, Springer-Verlag, Berlin et al., 2002.

[GP4] G.-M. Greuel and G. Pfister, *Computer algebra and finite groups*, In: Proc. First Intern. Congr. Math. Software, Beijing 2002, A. Cohen, X.-S. Gao, and N. Takayama (eds.), World Scientific, 2002.

[GPS] G.-M. Greuel, G. Pfister, and H. Schönemann, SINGULAR 2.0. *A Computer Algebra System for Polynomial Computations*, Centre for Computer Algebra, Univ. of Kaiserslautern, 2001, `http://www.singular.uni-kl.de` .

[Gr] K. Gruenberg, *Two theorems on Engel groups*, Proc. Camb. Phil. Soc. **49** (1953), 377–380.

[GKNP] F. Grunewald, B. Kunyavskiĭ, D. Nikolova, and E. Plotkin, *Two-variable identities in groups and Lie algebras*, Zap. Nauch. Semin. POMI **272** (2000), 161–176; J. Math. Sci. (New York) **116** (2003), 2972–2981.

[Gu] N. D. Gupta, *Some group laws equivalent to the commutative law*, Arch. Math. (Basel) **17** (1966), 97–102.

[GH] N. D. Gupta and H. Heineken, *Groups with a two-variable commutator identity*, Math. Z. **95** (1967), 276–287.

[H] B. Huppert, *Endliche Gruppen*, I, Springer-Verlag, Berlin–Heidelberg–New York, 1979.

[HB] B. Huppert and N. Blackburn, *Finite Groups*, III, Springer-Verlag, Berlin–Heidelberg–New York, 1982.

[Ka1] N. M. Katz, *Affine cohomological transforms, perversity, and monodromy*, J. Amer. Math. Soc. **6** (1993), 149–222.

[Ka2] N. M. Katz, *Sums of Betti numbers in arbitrary characteristics*, Finite Fields and Their Applications **7** (2001), 29–44.

[Ka3] N. M. Katz, *L-functions and monodromy: four lectures on Weil II*, Preprint, `http://www.math.princeton.edu/~nmk/arizona34.pdf` .

[Ko] A. N. Kostrikin, *Around Burnside*, Nauka, Moscow, 1986; English transl. Springer-Verlag, Berlin–New York, 1990.

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

[LY] D. Leep and C. Yeomans, *The number of points on a singular curve over a finite field*, Arch. Math. (Basel) **63** (1994), 420–426.

[Lu] A. Lubotzky, Pro-finite presentations, J. Algebra **242** (2001), 672–690.

[MSW] S. Margolis, M. Sapir, and P. Weil, *Closed subgroups in pro-V topologies and the extension problem for inverse automata*, Intern. J. Algebra and Computation **11** (2001), 405–445.

[Ne] H. Neumann, *Varieties of Groups*, Springer-Verlag, New York, 1967.

[Ni1] D. Nikolova, *Groups with a two-variable commutator identity*, C. R. Acad. Bulgare Sci. **36** (1983), 721–724.

[Ni2] D. Nikolova, *Solubility of finite groups with a two-variable commutator identity*, Serdica **11** (1985), 59–63.

[Pi] R. Pink, *On the calculation of local terms of the Lefschetz–Verdier trace formula and its application to a conjecture of Deligne*, Ann. of Math. (2) **135** (1992), 483–525.

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

[Plo1] B. I. Plotkin, *On nilgroups*, Dokl. Akad. Nauk SSSR **94** (1954), 999–1001. (Russian)

[Plo2] B. I. Plotkin, *Radical groups*, Mat. Sb. N.S. **37(79)** (1955), 507–526; English transl. in Amer. Math. Soc. Transl. (2) **17** (1961), 9–28.

[Plo3] B. I. Plotkin, *Generalized soluble and generalized nilpotent groups*, Uspekhi Mat. Nauk **13** (1958), *no.*4, 89–172; English transl. in Amer. Math. Soc. Transl. (2) **17** (1961), 29–115.

[PPT] B. Plotkin, E. Plotkin, and A. Tsurkov, *Geometrical equivalence of groups*, Comm. Algebra **27** (1999), 4015–4025.

[Re] L. Rédei, *Die endlichen einstufig nichtnilpotenten Gruppen*, Publ. Math. Debrecen **4** (1956), 303–324.

[RZ] L. Ribes and P. Zalesskii, *Profinite Groups*, Springer-Verlag, Berlin, 2000.

[Sch] O. J. Schmidt, *Groups all of whose subgroups are special*, Mat. Sbornik **31** (1924), 366–372. (Russian)

[SGA4] *Séminaire de Géometrie Algébrique du Bois-Marie* SGA $4\frac{1}{2}$, *Cohomologie Étale* (P. Deligne et al.), Lecture Notes Math. **569**, Springer-Verlag, Berlin et al., 1977.

[SGA5] *Séminaire de Géometrie Algébrique du Bois-Marie 1965–66* SGA 5, *Cohomologie ℓ-adique et Fonctions L* (A. Grothendieck et al., L. Illusie, ed.), Lecture Notes Math. **589**, Springer-Verlag, Berlin et al., 1977.

[Sh] I. R. Shafarevich, *Basic Algebraic Geometry*, 2nd ed., Springer-Verlag, Berlin et al., 1994.

[Th] J. Thompson, *Non-solvable finite groups all of whose local subgroups are solvable*, Bull. Amer. Math. Soc. **74** (1968), 383–437.

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

[We] P. Weil, *Profinite methods in semigroups*, Intern. J. Algebra and Computation **12** (2002), 137–178.

[Wi] J. S. Wilson, *Two-generator conditions for residually finite groups*, Bull. London Math. Soc. **23** (1991), 239–248.

[WZ] J. S. Wilson and E. Zelmanov, *Identities for Lie algebras of pro-p groups*, J. Pure Appl. Algebra **81** (1992), 103–109.

[Ze1] E. I. Zelmanov, *Engel Lie algebras*, Sibirsk. Mat. Zh. **29** (1988), *no.* 5, 112–117, 238; English transl. in Siberian Math. J. **29** (1988), 777–781.

[Ze2] E. I. Zelmanov, *Solution of the restricted Burnside problem for groups of odd exponent*, Izv. Akad. Nauk SSSR Ser. Mat. **54** (1990), 42–59; English transl. in Math. USSR Izv. **36** (1991), 41–60.

[Ze3] E. I. Zelmanov, *Solution of the restricted Burnside problem for* 2*-groups*, Mat. Sb. **182** (1991), 568–592; English transl. in Math. USSR Sb. **72** (1992), 543–565.

[Zi] T. Zink, *The Lefschetz trace formula for an open algebraic surface*, In: “Automorphic Forms, Shimura Varieties and *L*-Functions”, Proc. Conf. Ann Arbor 1988 (L. Clozel and J. S. Milne, eds.), Perspectives in Math. **11**, Academic Press, Boston, 1990, pp. 337–376.

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

Bandman, Kunyavskiĭ and Plotkin: Department of Mathematics and Statistics, Bar-Ilan University, 52900 Ramat Gan, ISRAEL

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

Greuel and Pfister: Fachbereich Mathematik, Universität Kaiserslautern, Postfach 3049, 67653 Kaiserslautern, GERMANY

E-mail address: `greuel@mathematik.uni-kl.de, pfister@mathematik.uni-kl.de`

Grunewald: Mathematisches Institut der Universität Heinrich Heine Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, GERMANY

E-mail address: `grunewald@math.uni-duesseldorf.de`