Abstract

We characterise the solvable groups in the class of finite groups by an inductively defined sequence of two-variable identities. Our main theorem is the analogue of a classical theorem of Zorn which gives a characterisation of the nilpotent groups in the class of finite groups by a sequence of two-variable identities. To cite this article: T. Bandman et al., C. R. Acad. Sci. Paris, Ser. I 337 (2003).

Résumé


Version française abrégée

Nous annonçons une caractérisation des groupes résolubles dans la classe des groupes finis par une suite d’identités en deux variables de type Engel définies par récurrence.

Théorème 0.1. Pour qu’un groupe fini $G$ soit résoluble, il faut et il suffit qu’il vérifie une des identités $u_n(x, y) = 1$.

E-mail addresses: bandman@macs.biu.ac.il (T. Bandman), greuel@mathematik.uni-kl.de (G.-M. Greuel), fritz@math.uni-duesseldorf.de (F. Grunewald), kunyav@macs.biu.ac.il (B. Kunyavskiǐ), Pfister@mathematik.uni-kl.de (G. Pfister), plotkin@macs.biu.ac.il (E. Plotkin).
Cet énoncé est l’exact analogue du théorème de Zorn qui caractérise les groupes nilpotents dans la classe des groupes finis par des identités en deux variables. L’existence des lois (non explicites) en deux variables pour les groupes résolubles finis a été démontrée dans [3,4]. B. Plotkin a proposé des identités (explicites et récursives) en deux variables pour caractériser cette classe de groupes. Notre Théorème 0.1 établit sa conjecture sous la forme modifiée.

Il est clair que tout groupe résoluble vérifie les identités un(x, y) = 1 à partir d’un certain n ∈ N. Réciproquement, si G vérifie l’identité un(x, y) = 1, on déduit qu’il est résoluble du résultat suivant:

**Théorème 0.2.** Soit G un groupe simple fini abélien. Alors il existe x, y ∈ G tels que u1(x, y) ≠ 1 et u1(x, y) = u2(x, y).

Nous allons donner ici une idée de la démonstration du Théorème 0.2 dans le cas particulier G = PSL(2, Fq).

Soit w un mot en x, x⁻¹, y, y⁻¹, et soit G un groupe. Pour x, y ∈ G, on définit uw(x, y) := w, et par récurrence u(w+1)(x, y) := [xuw(x, y)x⁻¹, yuw(x, y)y⁻¹]. Soit R := Z[t, u, v, w, z] l’anneau des polynômes en cinq variables à coefficients dans Z, et soient

\[ x = x(t) = \begin{pmatrix} t & -1 \\ 1 & 0 \end{pmatrix}, \quad y = y(u, v, w, z) = \begin{pmatrix} u & v \\ w & z \end{pmatrix} \]

des matrices sur R. Soit a l’idéal de R engendré par le déterminant de y et par 4 polynômes que l’on obtient à partir de l’équation matricielle uw(x, y) = u(w)(x, y), et soit \( \mathcal{V}^w \subset A^5 \) le sous-ensemble fermé de l’espace affine de dimension 5 qui correspond à a. Soit a0 l’idéal de R engendré par le déterminant de y et par les coefficients des matrices que l’on obtient à partir de l’équation uw(x, y) = 1, et soit \( \mathcal{V}^w_0 \subset A^5 \) l’ensemble fermé correspondant à a0. On va démontrer que pour tout q > 3, il y a des \( F_q \)-points sur \( \mathcal{V}^w \setminus \mathcal{V}^w_0 \).

Pour le mot w = x⁻²y⁻¹x, l’ensemble fermé \( \mathcal{V}^w \) a deux composantes irréductibles : l’une est \( \mathcal{V}^w_0 \) et l’autre, qu’on note \( \mathcal{C} \), est absolument irréductible de dimension 2. L’application \( \varphi : \mathcal{S} \to A^1 \setminus \{0\} \), \( \varphi(x, y) = u \), est une fibration en courbes de genre 8. On pose \( \mathcal{C} := \varphi^{-1}(1) \), et d’après la borne de Hasse–Weil on conclut que pour \( q > 593 \), il y a des \( F_q \)-rations sur \( \mathcal{C} \). Pour finir la démonstration du Théorème 0.2 dans le cas des groupes PSL(2), nous exhibons, pour tout 3 < q < 593, un \( F_q \)-point de \( \mathcal{C} \).

Nous allons décrire les composantes irréductibles de la variété \( \mathcal{V}^w \) pour le mot w = [x, y]. L’idéal correspondant a contient le polynôme \((- t + v + w)(v + w))\). Soit \( \mathcal{V}^w_1 \) l’ensemble fermé défini par l’idéal qui est engendré par a et v + w. Cette variété a 5 composantes parmi lesquelles l’une est de dimension 2 et coïncide avec \( \mathcal{V}^w_0 \), et les autres sont des composantes de dimension 0 et chacune se décompose en 4 composantes absolument irréductibles sur le corps de déploiement du polynôme \( 5x^4 + 20z^3 + 36z^2 + 32z + 16 \). Soit \( \mathcal{V}^w_2 \) l’ensemble fermé défini par l’idéal qui est engendré par a et \(- t + v + w\). Comme ci-dessus, cette variété a 5 composantes, toutes de dimension 1, parmi lesquelles l’une est contenue dans \( \mathcal{V}^w_0 \) et chacune des autres se décompose en 3 composantes absolument irréductibles sur le corps de déploiement du polynôme \( t^2 + t - 1 \). Comme aucune des composantes, sauf celle qui correspond aux solutions triviales de l’équation \( u_1 = u_2 \), n’est absolument irréductible, notre méthode ne marche pas pour le mot initial \( w = [x, y] \). En effet, pour ce mot, l’analogue du Théorème 0.2 n’est pas valable.

1. Introduction

We announce a characterisation of the solvable groups in the class of finite groups by an inductively defined, Engel like sequence of two-variable identities. Let G be a group and x, y ∈ G. Define

\[ u_1(x, y) := x^{-2}y^{-1}x, \quad \text{and inductively} \quad u_{n+1}(x, y) := \left[ xu_n(x, y)x^{-1}, yu_n(x, y)y^{-1} \right]. \]

The commutator of a, b ∈ G is defined as \([a, b] := aba^{-1}b^{-1}\). Our main result is

**Theorem 1.1.** A finite group G is solvable if and only if for some n the identity \( u_n(x, y) = 1 \) holds for all \( x, y \in G \).
Theorem 1.1 is analogous to a result of Zorn which characterises the nilpotent groups in the class of finite groups by two-variable identities, that is a finite group \( G \) is nilpotent if and only if it satisfies one of the identities \( e_n(x, y) = [x, y, y, \ldots, y] = 1 \) (here \( [x, y] = xyx^{-1}y^{-1} \), \( [x, y, y] = [[[x, y], y], y] \), etc.). The existence of two-variable (but non-explicit) laws for finite solvable groups is proved in [3,4]. B. Plotkin suggested explicit recursive two-variable identities for characterising this class of groups. Theorem 1.1 establishes his conjecture in a modified form.

Clearly in every solvable group the identities \( u_1(x, y) = 1 \) are satisfied from a certain \( n \in \mathbb{N} \) onward. The non-trivial direction of Theorem 1.1 follows immediately from

**Theorem 1.2.** 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 \( u_1(x, y) = u_2(x, y) \).

We remark that the equation \( u_1(x, y) = u_2(x, y) \) is equivalent to

\[
x^{-1}yx^{-1}y^{-1}x^2 = yx^{-2}y^{-1}xy^{-1}.
\]

Using Thompson’s list of the minimal simple non-solvable groups [11] we only need to prove Theorem 1.2 for the groups \( G \) in the following list: (1) \( G = \text{PSL}(2, \mathbb{F}_q) \) where \( q \geq 4 \) \( q = p^n \), \( p \) a prime, (2) \( G = \text{Sz}(2^n) \), \( n \in \mathbb{N} \), \( n \geq 3 \) and odd, (3) \( G = \text{PSL}(3, \mathbb{F}_3) \). Here \( \text{Sz}(2^n) \) \( n \in \mathbb{N} \), \( n \geq 3 \) denote the Suzuki groups (see [9, XI.3]). For small groups from this 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 \( \text{PSL}(3, \mathbb{F}_3) \).

The general idea of our proof can be roughly described as follows. For a group \( G \) in the above list, using a matrix representation over \( \mathbb{F}_q \) we interpret solutions of the equation \( u_1(x, y) = u_2(x, y) \) as \( \mathbb{F}_q \)-rational points of an algebraic variety. Lang–Weil type estimates for the number of rational points on a variety defined over a finite field guarantee in appropriate circumstances the existence of such points for big \( q \). Of course we are faced here with the extra difficulty of having to ensure that \( u_1(x, y) \neq 1 \) holds. This is achieved by taking the \( x, y \) from appropriate Zariski-closed subsets only. See Sections 2, 3 for the details.

The proof of Theorem 1.2 involves not only group-theoretic methods, but also algebraic geometry and computer algebra, particularly, the packages SINGULAR [7] and MAGMA. Not only proofs but even the precise statements of our results would hardly have been found without extensive computer experiments.

Theorem 1.1 has the following application in the profinite category (see [2]).

**Theorem 1.3.** Let \( F = F(x, y) \) denote the free group in two variables, and let \( \widehat{F} \) be its profinite completion. Let \( v_1, v_2, \ldots, v_m, \ldots \) be any convergent subsequence of the sequence (1) with limit \( f \) from \( \widehat{F} \). Then the identity \( f \equiv 1 \) defines the profinite variety of prosolvable groups.

Theorem 1.2 also generates short presentations of some finite simple groups. Let \( B \) be the group generated by \( x, y \) with the single relation (2), that is \( B = \langle x, y | u_1 = u_2 \rangle \). The solvable quotients of \( B \) are all cyclic, but \( B \) has at least all minimal simple groups from Thompson’s list as quotients. We for example found that \( \text{PSL}(2, \mathbb{F}_5) = \langle x, y | u_1 = u_2, x^3 = y^2 = 1 \rangle \) and \( \text{Sz}(8) = \langle x, y | u_1 = u_2, x^7 = y^5 = (xy)^2 = (x^{-1}y^{-1}xy)^2 = 1 \rangle \).

**2. The case \( G = \text{PSL}(2, \mathbb{F}_q) \)**

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).

Let \( w \) be a word in \( x, y \), \( x^{-1}, y, y^{-1} \). Let \( G \) be a group and \( x, y \in G \). Define \( u_1^n(x, y) := w \), and inductively \( u_1^{n+1}(x, y) := [xu_1^n(x, y)x^{-1}, yu_1^n(x, y)y^{-1}] \). (3)

Let \( R := \mathbb{Z}[t, u, v, w, z] \) be the polynomial ring over \( \mathbb{Z} \) in five variables. Consider further the two following \( 2 \times 2 \)-matrices over \( R \).
Let $\alpha$ be the ideal of $R$ generated by the determinant of $y$ and by the 4 polynomials arising from the matrix equation $u_i^w(x, y) = u_j^w(x, y)$ and let $V^w_0 \subset \mathbb{A}^5$ be the corresponding closed set of 5-dimensional affine space. Let further $a_0$ be the ideal of $R$ generated by the determinant of $y$ and by the matrix entries arising from the equation $u_i^w(x, y) = 1$ and let $V^w_0 \subset \mathbb{A}^5$ be the corresponding closed set. Our approach aims at showing that $V^w_0 \cap V^w_0$ has points over finite fields. We have therefore searched for words $w$ satisfying $\dim(V^w_0) - \dim(V^w_0) \geq 1$ and also $\dim(V^w_0) \geq 1$. We have only found the following words with this property.

\[
\begin{align*}
x^{-2}y^{-1}x, & \quad y^{-1}x y, \quad x^{-1}y^{-1}, \quad x^{-1}x y^{-1} x, \quad x^{-1} x x^{-1} y^{-1} x.
\end{align*}
\]

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

For the last 5 of the words in (4) the corresponding closed sets $V^w_0$ do not have absolutely irreducible components which are not contained in $V^w_0$ and in fact the analogue of Theorem 1.2 is not true for them. For the first word $w = x^{-2}y^{-1}x$ the closed set $V^w_0$ has 2 irreducible components. One of them is $V^w_0$, the second which we call $S$ has dimension 2 and is absolutely irreducible. The map $\psi : S \to \mathbb{A}^1 \setminus \{0\}$, $\psi(x, y) = u$, is a fibration with curves of genus 8 as fibers. We put $C := \psi^{-1}(1)$. According to the Hasse–Weil bound the number of $\mathbb{F}_q$-rational points of $C$ is at least $q + 1 - 2p_0 \sqrt{q} - d$ where $d$ is the degree and $p_0$ the arithmetic genus of $C$, the projective closure of $C$. Computations give $d = 10$ and $p_0 = 12$. 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 $\text{PSL}(2)$.

We conjecture after long computer experiments that Theorem 1.1 holds for any sequence formed like in (3) from words $u_i$ lying in $\mathbb{F}_q^*$ if $n$ is odd and in $\mathbb{F}_q^{2n}$ if $n$ is even.

3. The case of the Suzuki groups

To prove Theorem 1.2, the Suzuki groups $G = \text{Sz}(q)$ ($q = 2^n$, $n$ odd) provide the most difficult case. This is due to the fact that although $\text{Sz}(q)$ is contained in $\text{GL}(4, \mathbb{F}_q)$, it is not a Zariski-closed set. In fact the group $\text{Sz}(q)$ 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 $\text{Sz}(q)$ contains entries depending on $q$. We shall describe now how our problem can still be treated by methods of algebraic geometry.

Let $R := \mathbb{F}_2[\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta, \mu]$ be the polynomial ring over $\mathbb{F}_2$ in eight variables. Let $\pi : R \to R$ be its endomorphism defined by $\pi(\alpha) = a_0, \pi(\beta_0) := a^2, \ldots, \pi(\delta) := d_0, \pi(\epsilon_0) := d^2$. Let $\mathbb{F}$ be the algebraic closure of $\mathbb{F}_2$ and consider $\alpha, \ldots, \delta$ as the coordinates of eight dimensional affine space $\mathbb{A}^8$ over $\mathbb{F}$. The endomorphism $\pi$ defines an algebraic bijection $\alpha : \mathbb{A}^8 \to \mathbb{A}^8$. The square of $\alpha$ is the Frobenius automorphism on $\mathbb{A}^8$. Let $p \in \mathbb{A}^8$ be a fixed point of $\alpha^n$, then its coordinates are in $\mathbb{F}_{2^n}$ if $n$ is odd and in $\mathbb{F}_{2^{2n}}$ if $n$ is even.

Consider further the following two matrices in $\text{GL}(4, \mathbb{F})$

\[
\begin{align*}
x &= \begin{pmatrix}
a^2 & 0 & 1 & 0 \\
\frac{a}{\alpha} & 0 & 0 & 0 \\
1 & 0 & 0 & 0
\end{pmatrix}, & 
\quad y &= \begin{pmatrix}
c^2 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
1 & 0 & 0 & 0
\end{pmatrix}.
\end{align*}
\]

The matrices $x, y$ also define maps from $\mathbb{A}^8$ to $\text{GL}(4, \mathbb{F})$. It can easily be checked that the matrices corresponding to a fixed point of $\alpha^n$ ($n$ odd and $n \geq 3$) lie in $\text{Sz}(2^n)$.
Let $a$ be the ideal of $R$ generated by the 16 polynomials arising from the matrix equation (2) and let $\mathcal{V} \subset \mathbb{A}^8$ be the corresponding closed set. By a SINGULAR computation we find

**Proposition 3.1.** (1) $\dim(\mathcal{V}) = 2$, (2) $\pi(a) = a$.

Using Proposition 3.1 we see that $\alpha$ defines an algebraic map $\alpha : \mathcal{V} \to \mathcal{V}$. Our task now becomes to show that $\alpha^n$ ($n$ odd and $n \geq 3$) has a non-zero fixed point on the surface $\mathcal{V}$. Our basic tool is the Lefschetz trace formula resulting from Deligne’s conjecture proved by Fujiwara [5]. To apply the Lefschetz trace formula we need to study their degrees. Note that since $\mathcal{V}$ is affine, we have $b^3(\mathcal{V}) = b^4(\mathcal{V}) = 0$. Since $\mathcal{V}$ is non-singular, the ordinary and compact Betti numbers of $\mathcal{V}$ are related by Poincaré duality, and we have $b^1(\mathcal{V}) = b^{4-1}(\mathcal{V})$.

Let now $\mathcal{U} \subset \mathcal{V}'$ be the complement in $\mathcal{V}'$ of the closed set given by the equation $cc0 = 0$. We have

**Proposition 3.3.** $\mathcal{U}$ is a smooth $\alpha$-invariant absolutely irreducible affine surface with $b^1(\mathcal{U}) \leq 675$ and $b_2(\mathcal{U}) \leq 2^{22}$.

Here $b^1(\mathcal{U}) = \dim H^1_{et}(\mathcal{U}, \mathcal{Q}_l)$ are the $\ell$-adic Betti numbers ($\ell \not= 2$). The estimates contained in Proposition 3.3 are derived from results of Adolphson and Sperber [1] and Ghorpade and Lachaud [6] permitting to bound the Betti numbers of an affine variety in terms of the number of variables, the number of defining polynomials and their degrees. Note that since $\mathcal{U}$ is affine, we have $b^2(\mathcal{U}) = b^4(\mathcal{U}) = 0$. Since $\mathcal{U}$ is non-singular, the ordinary and compact Betti numbers of $\mathcal{U}$ are related by Poincaré duality, and we have $b^1(\mathcal{U}) = b^{4-1}(\mathcal{U})$.

Let $\text{Fix}(\mathcal{U}, n)$ be the number of fixed points of $\alpha^n$ acting on $\mathcal{U}$. The Lefschetz trace formula applied to $\mathcal{U}$ and from Deligne’s estimates for the eigenvalues of the endomorphism induced by $\alpha$ on étale cohomology we get

$$|\text{Fix}(\mathcal{U}, n) - 2^n| \leq b^1(\mathcal{U})2^{3n/4} + b^2(\mathcal{U})2^{n/2}. \quad (5)$$

An easy estimate shows that $\text{Fix}(\mathcal{U}, n) \neq 0$ for $n > 48$. The cases $n < 48$ are checked with the help of MAGMA. As a by-product of these computations we found the first terms of the zeta-function of the operator $\alpha$ acting on the set $\mathcal{U}$. This is a rational function defined by

$$Z(\alpha, T) := \exp \left( - \sum_{n=1}^{\infty} \frac{\text{Fix}(\mathcal{U}, n)}{n} T^n \right). \quad (6)$$

We have found that $Z(\alpha, T)$ is equal to

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

up to terms of order $T^{33}$. This formula suggests heuristic values $b_1(\mathcal{U}) = 1$, $b_2(\mathcal{U}) = 6$, $b_3(\mathcal{U}) = 43$.

4. Analogue, problems and generalizations

Theorem 1.1 leads to some natural problems. It would be very interesting to understand the special role of our initial word $u_1 = x^{-2}y^{-1}x$ (see Section 2). Further, there is the following analogue of van der Waerden’s problems for nilpotent groups. Fix $n \in \mathbb{N}$ and assume that a finite group $G$ satisfies the identity $u_n(x, y) = 1$. What can be said about the derived length of $G$?

In connection with Theorem 1.1, one may think of its analogues for Lie algebras and infinite groups.

For finite dimensional Lie algebras the following Engel type result is proved in [8].

**Theorem 4.1.** Let $L$ be a finite dimensional Lie algebra defined over a field $k$ of characteristic different from 2, 3, 5. Let $[\cdot, \cdot]$ denote its Lie bracket. Define
\[ v_1(x, y) := [x, y], \quad \text{and inductively} \quad v_{n+1}(x, y) := [[v_n(x, y), x], [v_n(x, y), y]]. \quad (7) \]

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

For infinite dimensional Lie algebras the following problem arises in the light of the Kostrikin–Zelmanov results. Let \( L \) be a Lie algebra over a field \( k \) and suppose that there is \( n \in \mathbb{N} \) 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?

For infinite groups much less is known. Let us call a group \( G \) Engel or quasi-Engel if there is an \( n \in \mathbb{N} \) such that the identity \( v_n(x, y) \equiv 1 \) or \( u_n(x, y) \equiv 1 \) holds in \( G \). Is every Engel (resp. quasi-Engel) group locally nilpotent (resp. locally solvable)? In the nilpotent case this is a long-standing open problem, cf. [10]. In the solvable case, residually finite and profinite groups deserve special interest.

The restricted Engel problem has a positive solution. One may pose the restricted quasi-Engel problem: 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 cosolvable normal subgroups in \( F \) is also cosolvable?

Finally, consider an interesting particular case of linear groups.

**Theorem 4.2.** Suppose that \( G \subset GL(n, K) \) where \( K \) is a field. Then \( G \) is solvable if and only if it is quasi-Engel.

**Acknowledgements**

Bandman, Kunyavskiï, 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. Kunyavskiï 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.

We are grateful to N. Gordeev, D. Grayson, L. Illusie, A. Lubotzky, A. Mann, S. Margolis, R. Pink, L. Rowen, D. Segal, Y. Segev, J.-P. Serre, M. Stoll, Y. Varshavsky and N. Vavilov for useful comments and advise. 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.

**References**