

On Brandl's Question on Finite Simple Groups of Bounded Engel Depth: A Reduction to the Suzuki Groups

Alper Ferudun*

October 22, 2026

Abstract

For $r \geq 1$ let V_r be the class of finite groups G that satisfy a law $[x, {}_r y] = [x, {}_s y]$ for some $s = s(G) > r$. R. Brandl asked (Kourovka Notebook, Problem 11.16 b)) whether V_r contains only finitely many nonabelian simple groups. Brandl proved that V_2 contains none (1983), and Brandl and Nikolova proved that for $q \geq 4$ the group $\text{PSL}(2, q)$ lies in V_3 only for $q \in \{4, 5, 8\}$ (1986). We give a *partial answer to part b): for every r , all but finitely many nonabelian finite simple groups in V_r are Suzuki groups, so b) holds if and only if the Engel depth of $\text{Sz}(q)$ tends to infinity; this last statement and part a) stay open*. Here the Engel depth of G is the least r with $G \in V_r$; part a), on the Fitting length of soluble groups in V_r , is not treated. Extending the method of Brandl and Nikolova, we prove that $\text{PSL}(2, q) \notin V_r$ for all $q > Q_r$, with an explicit Q_r : the Fricke trace recursion turns long Engel chains into rational points of an absolutely irreducible plane curve, which the Weil bound provides. Alternating, classical and exceptional groups other than $\text{Sz}(q)$ have alternating sections of unbounded degree or sections $\text{PSL}(2, q)$, and the classification of finite simple groups completes the reduction. For an odd prime p , every proper section of $\text{Sz}(2^p)$ has Engel depth at most 3, so no argument through sections can reach these groups. As evidence only, we compute the Engel depths 62 of $\text{Sz}(8)$ and 527 of $\text{Sz}(32)$, and lower bounds from explicit pairs up to 792,500 for $\text{Sz}(8192)$. No printed answer to part b) was found in the sources searched; we claim no priority. This is an unrefereed note.

1 Introduction

1.1 The problem

For elements x, y of a group put $[x, y] = x^{-1}y^{-1}xy$, $[x, {}_1 y] = [x, y]$ and $[x, {}_{i+1} y] = [[x, {}_i y], y]$. Every finite group satisfies a law $[x, {}_r y] = [x, {}_s y]$ for some integers $1 \leq r < s$; Brandl called the least such r the *depth* of the group [Bra83b], and Brandl and Nikolova the *Engel depth* [BN86]. Problem 11.16 of the Kourovka Notebook [KM26], posed by R. Brandl in the 11th issue (1990), concerns the class V_r of all finite groups G that satisfy such a law with the given r and some $s = s(G) > r$. Part a) asks whether the soluble groups in V_r have Fitting length bounded by a function of r . Part b) asks: “Is it true that V_r contains only finitely many nonabelian simple groups?”, and the Notebook adds that this is true for $r < 4$. The text is the same in versions 48 and 49 of [KM26], with no comment; the problem is record KOU-11.16 of the UnsolvedMath collection [ula26], where it is listed as open. *This note treats part b) only; part a) is not treated.*

We write $\mu(G)$ for the Engel depth, normalised as in Section 2: $\mu(G)$ is the largest pre-period of the sequences $x, [x, y], [x, {}_2 y], \dots$, and $G \in V_r$ if and only if $\mu(G) \leq r$ (Lemma 2.1); for $G \neq 1$ this is

*Mercury Software GmbH. Email: alper@mercurycodelab.com. GitHub: <https://github.com/AlperTheKing>.
Unrefereed preprint. CC BY 4.0.

Brandl’s depth. Brandl and Nikolova recall a theorem of Heineken and Neumann [HN67]: a nontrivial variety contains only finitely many of the groups $\text{PSL}(2, q)$ and $\text{Sz}(q)$. Since s may depend on G , the classes V_r are not varieties, and Brandl and Nikolova conjectured that every V_r contains only finitely many nonabelian simple groups [BN86, p. 246]; this is part b).

1.2 Known results

Groups in V_1 are abelian, by a theorem of Gupta [Gup66] that Brandl recalls [Bra83b, Section 1], and every group in V_2 is soluble [Bra83b, Sections 1 and 2]; for the latter Brandl shows that $\text{Sz}(q)$ and $\text{SL}(3, 3)$ are not in V_2 [Bra83b, Lemma 2], nor $\text{PSL}(2, q)$ for $q > 4$ [Bra83b, Lemma 3], and uses Thompson’s list of the minimal simple groups (he notes that the solubility was found independently by D. Nikolova). Brandl and Nikolova proved that for $q \geq 4$ the group $\text{PSL}(2, q)$ lies in V_3 if and only if $q \in \{4, 5, 8\}$; they list the minimal laws of $\text{PSL}(2, q)$ for $q \in \{4, 7, 8, 9, 11, 13\}$ and remark that $\text{Sz}(8)$ has Engel depth at least 11 [BN86]. Nikolova computed the minimal laws $[x, {}_3y] = [x, {}_{63}y]$ of A_5 and $[x, {}_4y] = [x, {}_{124}y]$ of A_6 [Nik84]. So the remark “true for $r < 4$ ” is proved in [Bra83b] for $r \leq 2$, and in [BN86] for $r = 3$ and the groups $\text{PSL}(2, q)$. For $r = 3$ and the other families we found no published proof; the results below settle $r = 3$ for all families except the Suzuki groups, and we do not prove the case $r = 3$ for the Suzuki groups.

1.3 Results

Let

$$Q_r = ((d-1)(d-2)+2)^2, \quad d = 3 \cdot 2^{r-1} \quad (r \geq 1); \quad (1)$$

for even q the smaller value with $d = 2^r$ suffices (Theorem 3.5).

Theorem 1.1. *For every $r \geq 1$, all but finitely many nonabelian finite simple groups in V_r are isomorphic to Suzuki groups $\text{Sz}(2^{2m+1})$, $m \geq 1$. More precisely, a nonabelian finite simple group in V_r that is not a Suzuki group is a sporadic group, the Tits group, an alternating group A_m with $m \leq 2^{r+1} + 1$, or a group of Lie type with parameter $q \leq Q_r$ (the q of the usual notation, as in $\text{PSU}_N(q)$ or ${}^3D_4(q)$) whose natural module, if it is classical, has dimension at most $2^{r+2} + 5$.*

Corollary 1.2. *Part b) of Problem 11.16 has a positive answer if and only if $\mu(\text{Sz}(q)) \rightarrow \infty$ as $q = 2^{2m+1} \rightarrow \infty$.*

So the note gives a *partial answer to part b)*: for every r , all but finitely many nonabelian finite simple groups in V_r are Suzuki groups, so b) holds if and only if the Engel depth of $\text{Sz}(q)$ tends to infinity; this last statement and part a) stay open. The main step, Theorem 3.5, extends the finiteness part of the theorem of Brandl and Nikolova ($r = 3$) to all r , by their method: if $q > Q_r$, then $\text{PSL}(2, q) \notin V_r$; in particular $\mu(\text{PSL}(2, q)) \geq \lfloor (\log_2 q)/4 \rfloor$ for $q \geq 256$. The other families are reduced to it and to alternating groups through sections (Theorem 4.1). For the Suzuki groups this is impossible: for an odd prime p , every proper section of $\text{Sz}(2^p)$ has Engel depth at most 3 (Proposition 5.1). As evidence only, we computed $\mu(\text{Sz}(8)) = 62$ and $\mu(\text{Sz}(32)) = 527$, and certified lower bounds from explicit pairs for $q = 128, 512, 2048, 8192$ (Table 2). Since $\text{Sz}(2^d) \leq \text{Sz}(2^n)$ for $d \mid n$, they show that for $3 \leq r \leq 61$ the Suzuki groups whose membership in V_r is not decided here are exactly the $\text{Sz}(2^n)$ with $\gcd(n, 15015) = 1$ (Corollary 5.4).

1.4 Method and related work

The proof of Theorem 3.5 follows that of Brandl and Nikolova for $r = 3$ [BN86, Section 2]: an explicit pair, an absolutely irreducible curve (by Eisenstein’s criterion) and Weil’s bound. Their

explicit matrices are replaced by the trace map of [BGG⁺06, BGKJ10]: for y of trace s the traces of $[x, {}_k y]$ obey $t \mapsto t^2 - s^2 t + 2s^2 - 2$ (Lemma 3.1), the recursion of [BGG12, (3.3)] for the other commutator convention; an Engel chain that ends at a non-periodic trace is a point of an absolutely irreducible curve, and the Weil bound for singular plane curves [AP96], in the form [BGG12, (5.8)], provides such points. Bandman, Garion and Grunewald prove that Engel words are surjective on $\mathrm{PSL}(2, q)$ for large q [BGG12]; their Proposition 4.9 uses the pre-period $n_0(q)$ of the trace map over a fixed field, but neither [BGG12] nor [BGKJ10] studies how such pre-periods grow with q , or Brandl's problem, and neither cites Brandl's papers on Engel depth ([BGKJ10] cites two unrelated papers of Brandl, on polynomials with roots modulo every prime). For the Suzuki groups there is no trace map [BGKJ10, Section 4] (see Section 5.2). The Engel sinks of Khukhro and Shumyatsky [KS16] contain all sufficiently long commutators $[x, {}_n g]$; results on them bound the periodic part of the Engel map, not its pre-period. Brandl also studied Engel cycles [Bra83a].

2 Engel tails and the classes V_r

Let G be a finite group. For $y \in G$ let $\varphi_y: G \rightarrow G$, $\varphi_y(z) = [z, y]$, and for $x \in G$ let $a_k = a_k(x, y) = \varphi_y^k(x)$, so that $a_0 = x$ and $a_k = [x, {}_k y]$ for $k \geq 1$. An element z is φ_y -periodic if $\varphi_y^m(z) = z$ for some $m \geq 1$. The *tail* $\mathrm{tail}(x, y)$ is the least $k \geq 0$ such that a_k is φ_y -periodic; then a_j is periodic for every $j \geq \mathrm{tail}(x, y)$. The *Engel depth* of G is

$$\mu(G) = \max_{x, y \in G} \mathrm{tail}(x, y).$$

Since $a_k(x^\alpha, y^\alpha) = a_k(x, y)^\alpha$ for $\alpha \in \mathrm{Aut}(G)$, the tail is invariant under automorphisms applied to both entries; so $\mu(G)$ is the maximum over y in a set of representatives of the conjugacy classes.

Lemma 2.1. *Let $r \geq 1$. A finite group G lies in V_r if and only if $a_r(x, y)$ is φ_y -periodic for all $x, y \in G$, that is, if and only if $\mu(G) \leq r$. Hence $V_1 \subseteq V_2 \subseteq \dots$, and for $G \neq 1$ the number $\mu(G)$ is the depth of [Bra83b].*

Proof. If $[x, {}_r y] = [x, {}_s y]$ for all x, y with $s > r$, then $\varphi_y^{s-r}(a_r) = a_s = a_r$, so a_r is periodic. Conversely, if each $a_r(x, y)$ lies on a cycle of length $p(x, y)$, let L be the least common multiple of the finitely many $p(x, y)$ and $s = r + L$; then $a_s = \varphi_y^L(a_r) = a_r$ for all pairs. For $G \neq 1$ and $x \neq 1$ we have $a_1(x, 1) = 1 \neq x$, so $\mathrm{tail}(x, 1) = 1$ and $\mu(G) \geq 1$; the least $r \geq 1$ with $G \in V_r$ is then $\mu(G)$. $\square$

Lemma 2.2. *Let V'_r be defined as V_r with the commutator $[x, y]' = xyx^{-1}y^{-1}$. Then $V'_r = V_r$.*

Proof. Fix y , let $\varphi'_y(z) = [z, y]'$, $\iota(z) = z^{-1}$ and $\kappa(z) = y^{-1}zy$. Then $\iota\varphi'_y\iota(z) = (z^{-1}yz y^{-1})^{-1} = yz^{-1}y^{-1}z = \kappa^{-1}(\varphi_y(z))$. As κ commutes with φ_y , the map $\psi = \kappa^{-1}\varphi_y$ satisfies $\psi^n = \kappa^{-n}\varphi_y^n$. If N is the order of κ , then $\varphi_y^m(z) = z$ implies $\psi^{mN}(z) = z$ and $\psi^m(z) = z$ implies $\varphi_y^{mN}(z) = z$; so ψ and φ_y have the same periodic points, κ permutes them, and $\psi^n(z)$ is periodic if and only if $\varphi_y^n(z)$ is. Hence $\mathrm{tail}_{\varphi'_y}(z) = \mathrm{tail}_\psi(z^{-1}) = \mathrm{tail}_{\varphi_y}(z^{-1})$, and the maxima over z and y agree. By Lemma 2.1 for both conventions, $V'_r = V_r$. $\square$

Lemma 2.3. *V_r is closed under subgroups and homomorphic images. Hence $\mu(K/L) \leq \mu(G)$ whenever $L \trianglelefteq K \leq G$.*

Proof. A law of G holds in every subgroup and in every quotient of a subgroup. Apply this with $r = \mu(G)$ and use Lemma 2.1. $\square$

Lemma 2.4. (a) *If G is nilpotent, then 1 is the only φ_y -periodic element, for every y ; so $\mathrm{tail}(x, y) = \min\{k : [x, {}_k y] = 1\}$.*

(b) In the dihedral group $D = \langle \rho, \tau \mid \rho^{2^n}, \tau^2, \tau\rho\tau = \rho^{-1} \rangle$ of order 2^{n+1} , $n \geq 2$, we have $\text{tail}(\rho, \tau) = n$; so $\mu(D) \geq n$.

(c) For $m \geq 6$, $\mu(A_m) \geq \lfloor \log_2(m-2) \rfloor$.

Proof. (a) If G has class c , then $\varphi_y^c(z) = 1$ for all z , and 1 is fixed; a periodic z satisfies $z = \varphi_y^{cM}(z) = 1$ for a suitable multiple cM of its period. (b) For $j \in \mathbb{Z}$, $[\rho^j, \tau] = \rho^{-j}\tau\rho^j\tau = \rho^{-2j}$, so $a_k(\rho, \tau) = \rho^{(-2)^k}$, which is 1 if and only if $k \geq n$. By (a), $\text{tail}(\rho, \tau) = n$. (c) Let $k = \lfloor \log_2(m-2) \rfloor \geq 2$, so $2^k + 2 \leq m$. The group D of (b) with $n = k$ acts faithfully on the vertices $\mathbb{Z}/2^k$ of a regular 2^k -gon, and $g \mapsto g \cdot (2^k+1, 2^k+2)^{e(g)}$, with $e(g) \in \{0, 1\}$ the parity of g on $\mathbb{Z}/2^k$, is an injective homomorphism into $A_{2^k+2} \leq A_m$. By Lemma 2.3 and (b), $\mu(A_m) \geq k$. $\square$

The bound in (c) is far from sharp: $\mu(A_5) = 3$, $\mu(A_6) = 4$ (Table 1, [Nik84]), and a verification run computed $\mu(A_7) = 15$ (with the minimal law $[x,_{15} y] = [x,_{35295} y]$).

3 The groups $\text{PSL}(2, q)$

Lemma 3.1. *Let K be a field, $x, y \in \text{SL}(2, K)$, $s = \text{tr } y$, $a_k = a_k(x, y)$ and $t_k = \text{tr } a_k$.*

(a) (Fricke) *For $A, B \in \text{SL}(2, K)$, $\text{tr}[A, B] = (\text{tr } A)^2 + (\text{tr } B)^2 + (\text{tr } AB)^2 - \text{tr } A \text{tr } B \text{tr } AB - 2$. In particular $t_1 = \alpha^2 + s^2 + \gamma^2 - \alpha s \gamma - 2$ with $\alpha = \text{tr } x$ and $\gamma = \text{tr } xy$.*

(b) *For every $k \geq 1$, $\text{tr}(a_k y) = s(t_k - 1)$ and $t_{k+1} = f_s(t_k)$, where $f_s(t) = t^2 - s^2 t + 2s^2 - 2$.*

(c) *With $h = t - 2$ and $c = 4 - s^2$: $f_s(2 + h) = 2 + g_c(h)$, where $g_c(h) = h^2 + ch$.*

Proof. (a) For $X, Y \in \text{SL}(2, K)$ the Cayley–Hamilton theorem gives $Y + Y^{-1} = (\text{tr } Y)I$, hence $\text{tr}(XY) + \text{tr}(XY^{-1}) = \text{tr } X \text{tr } Y$. Let $a = \text{tr } A$, $b = \text{tr } B$, $c' = \text{tr } AB$, and apply this to $X = A^{-1}B^{-1}$, $Y = AB$, both of trace c' : $\text{tr}[A, B] = c'^2 - \text{tr}(A^{-1}B^{-2}A^{-1})$. Now $\text{tr}(A^{-1}B^{-2}A^{-1}) = \text{tr}(A^{-2}B^{-2})$, and $A^{-2} = aA^{-1} - I$, $B^{-2} = bB^{-1} - I$ give $\text{tr}(A^{-2}B^{-2}) = abc' - a^2 - b^2 + 2$. (b) For $k \geq 1$, $a_k y = a_{k-1}^{-1} y^{-1} a_{k-1} y^2$ and $y^2 = sy - I$, so $a_k y = s a_k - a_{k-1}^{-1} y^{-1} a_{k-1}$ and $\text{tr}(a_k y) = s t_k - \text{tr } y^{-1} = s(t_k - 1)$. By (a) with $(A, B) = (a_k, y)$ and $u_k = \text{tr}(a_k y) = s(t_k - 1)$, $t_{k+1} = t_k^2 + s^2 + u_k^2 - s t_k u_k - 2$, and $u_k^2 - s t_k u_k = -s^2(t_k - 1)$. (c) $(2 + h)^2 - s^2(2 + h) + 2s^2 - 2 = 2 + (4 - s^2)h + h^2$. $\square$

The recursion (b) is [BG12, (3.3)], where the other commutator convention is used.

Lemma 3.2. *Let $x, y \in \text{SL}(2, q)$ with images $\bar{x}, \bar{y}$ in $\text{PSL}(2, q)$, and keep the notation of Lemma 3.1.*

(a) *If $k \geq 1$ and a_k is φ_y -periodic, then t_k is f_s -periodic.*

(b) *If $a_k(\bar{x}, \bar{y})$ is periodic in $\text{PSL}(2, q)$, then a_{k+1} is periodic in $\text{SL}(2, q)$.*

(c) *If q is even, or if $s \neq 0$, then for $k \geq 1$ the element $a_k(\bar{x}, \bar{y})$ is periodic if and only if a_k is.*

Hence, if t_{r+1} is not f_s -periodic, then $\text{tail}(\bar{x}, \bar{y}) \geq r + 1$, and $\text{tail}(\bar{x}, \bar{y}) \geq r + 2$ if q is even or $s \neq 0$.

Proof. (a) If $a_{k+m} = a_k$ with $m \geq 1$, then $t_k = t_{k+m} = f_s^m(t_k)$ by Lemma 3.1(b). (b) Since $[-a, y] = [a, -y] = [a, y]$, the sequence $(a_k(\bar{x}, \bar{y}))$ is the image of (a_k) . If $a_{k+m}(\bar{x}, \bar{y}) = a_k(\bar{x}, \bar{y})$, then $a_{k+m} = \pm a_k$ and $a_{k+m+1} = [\pm a_k, y] = a_{k+1}$. (c) For q even, $\text{PSL}(2, q) = \text{SL}(2, q)$. If q is odd, $k \geq 1$ and $a_{k+m} = -a_k$, then by Lemma 3.1(b) $s(-t_k - 1) = \text{tr}(a_{k+m} y) = -\text{tr}(a_k y) = -s(t_k - 1)$, so $2s = 0$. For the last statement, a_{r+1} is not periodic by (a); so $a_r(\bar{x}, \bar{y})$ is not periodic by (b), and $a_{r+1}(\bar{x}, \bar{y})$ is not periodic under the hypothesis of (c). $\square$

Lemma 3.3. *Let $s, t \in \mathbb{F}_q$ with $s^2 \neq 4$, and $y_s = \begin{pmatrix} 0 & -1 \\ 1 & s \end{pmatrix}$. There is $x \in \text{SL}(2, q)$ with $\text{tr}[x, y_s] = t$.*

This is a special case of Macbeath's theorem [Mac69] that every triple in $\mathbb{F}_q^3$ is $(\text{tr } x, \text{tr } xy, \text{tr } y)$ for some $x, y \in \text{SL}(2, q)$ (see [BGKJ10, Theorem 3.4] for an elementary proof); we include the short argument for the case needed.

Proof. By Lemma 3.1(a), $\text{tr}[x, y_s] = \Phi(\alpha, \gamma) + s^2 - 2$ with $\alpha = \text{tr } x$, $\gamma = \text{tr}(xy_s)$ and $\Phi(\alpha, \gamma) = \alpha^2 - s\alpha\gamma + \gamma^2$. So it suffices to find α, γ with $\Phi(\alpha, \gamma) = N := t + 2 - s^2$, and then x with these two traces. If q is even, take $\gamma = 0$, $\alpha = \sqrt{N}$ and $x = \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix}$, which has determinant 1 and $\text{tr}(xy_s) = 0$. Let q be odd. Then $\Phi = (\alpha - s\gamma/2)^2 + D\gamma^2$ with $D = (4 - s^2)/4 \neq 0$. The sets $\{X^2 : X \in \mathbb{F}_q\}$ and $\{N - D\gamma^2 : \gamma \in \mathbb{F}_q\}$ have $(q+1)/2$ elements each, so they meet; this gives α and γ . For $x = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we have $\text{tr}(xy_s) = b - c + ds$. Given d , put $a = \alpha - d$, $e = \gamma - ds$ and $c = b - e$; then $\det x = 1$ becomes $b^2 - eb - (\alpha - d)d + 1 = 0$, with discriminant $\Delta(d) = (s^2 - 4)d^2 + (4\alpha - 2s\gamma)d + \gamma^2 - 4$. As $s^2 \neq 4$, the quadratic polynomial Δ takes $(q+1)/2$ values, so $\Delta(d) = w^2$ for some $d, w \in \mathbb{F}_q$, and $b = (e + w)/2$ solves the equation. $\square$

Lemma 3.4. *For $r \geq 1$ let g_c^r be the r -fold iterate of $h \mapsto h^2 + ch$, and $R_r(h, c) = g_c^r(h) + c \in \mathbb{Z}[h, c]$.*

(a) *For every prime p , R_r is irreducible in $\overline{\mathbb{F}}_p[h, c]$; its total degree is 2^r .*

(b) *For every odd prime p , $R_r(h, 4 - s^2)$ is irreducible in $\overline{\mathbb{F}}_p[h, s]$; its total degree is $3 \cdot 2^{r-1}$.*

Proof. Let $k = \overline{\mathbb{F}}_p$. Since $g_0(h) = h^2$, we have $R_r \equiv h^{2^r} \pmod{c}$, and $R_r(0, c) = c$ because $g_c(0) = 0$. (a) As a polynomial in h over $k[c]$, R_r is monic of degree 2^r , its other coefficients are divisible by c , and its constant term c is not divisible by c^2 . By Eisenstein's criterion for the prime c of $k[c]$, R_r is irreducible over $k(c)$, and as it is monic in h , it is irreducible in $k[h, c]$ (Gauss's lemma). (b) Put $\tau = s - 2$, so $c = 4 - s^2 = -\tau(\tau + 4)$. Then $R_r \equiv h^{2^r} \pmod{\tau}$, and the constant term $-\tau(\tau + 4)$ is not divisible by τ^2 , since $\tau + 4$ takes the value $4 \neq 0$ at $\tau = 0$ (p odd). Eisenstein's criterion for the prime τ of $k[\tau] = k[s]$ and Gauss's lemma give (b). For the degrees, let D_j be the total degree of g_c^j (in (h, c) , resp. in (h, s) after $c = 4 - s^2$). Then $D_1 = 2$, resp. 3, and $D_{j+1} = 2D_j$, because the top-degree part of $(g_c^j)^2$ is the square of that of g_c^j and $\deg c + D_j < 2D_j$. So R_r has degree D_r . $\square$

Equivalently, the Newton polygon of R_r with respect to c (resp. τ) is the single segment from $(0, 1)$ to $(2^r, 0)$: all roots of R_r over $k((c))$ (resp. $k((\tau))$) have valuation 2^{-r} , so the ramification index, and with it the degree of the extension, is 2^r . In characteristic 2 the argument of (b) fails, as $c = \tau^2$ there; case (a) is used instead.

Theorem 3.5. *Let $r \geq 1$ and let q be a prime power. Put $d = 2^r$ if q is even and $d = 3 \cdot 2^{r-1}$ if q is odd, and $Q = ((d-1)(d-2) + 2)^2$. If $q > Q$, there are $x, y \in \text{SL}(2, q)$ whose images in $\text{PSL}(2, q)$ have tail at least $r+1$, and at least $r+2$ if q is even or $\text{tr } y \neq 0$. Hence $\text{PSL}(2, q) \notin V_r$ for $q > Q_r$, and $\mu(\text{PSL}(2, q)) \geq \lfloor (\log_2 q)/4 \rfloor$ for $q \geq 256$.*

Proof. Let p be the characteristic. Let C be the affine plane curve $R_r(h, c) = 0$ if q is even, and $R_r(h, 4 - s^2) = 0$ if q is odd; by Lemma 3.4 it is defined over $\overline{\mathbb{F}}_p$ and absolutely irreducible of degree d , and so is its projective closure $\bar{C}$. By the Weil bound for absolutely irreducible, possibly singular, plane curves of degree d ,

$$|\#\bar{C}(\mathbb{F}_q) - (q+1)| \leq (d-1)(d-2)\sqrt{q}$$

(Aubry-Perret [AP96], in the form [BGG12, (5.8)], which also refers to [GL02, Corollary 7.4] and [LY94, Corollary 2]). At most d of these points lie on the line at infinity. Points of C with $c = 0$ satisfy $h^{2^r} = R_r(h, 0) = 0$; they are $(h, c) = (0, 0)$, resp. $(h, s) = (0, \pm 2)$. So C has at least $N_0 = q - 1 - (d-1)(d-2)\sqrt{q} - d$ points over $\mathbb{F}_q$ with $c \neq 0$. Let $q > Q$. If $d \geq 3$, then $\sqrt{q} > (d-1)(d-2) + 2 \geq d+1$ (as $(d-1)(d-2) + 2 - (d+1) = (d-1)(d-3)$), so $N_0 = \sqrt{q}(\sqrt{q} - (d-1)(d-2)) - d - 1 > 2\sqrt{q} - d - 1 > d+1$. If $d = 2$, that is $r = 1$ and q even, then $N_0 = q - 3 > 0$.

q	4	5	7	8	9	11	13	16	17	19	23	25	27
$\mu(\text{PSL}(2, q))$	3	3	4	3	4	6	7	5	7	9	8	9	9

q	29	31	32	37	41	43	47	49	53	59	61	64
$\mu(\text{PSL}(2, q))$	12	9	4	12	11	14	13	17	17	17	19	6

Table 1: Exact Engel depths of $\text{PSL}(2, q)$, $q \leq 64$ (all classes of y , the full functional graph of φ_y); $\mu(\text{SL}(2, q)) = \mu(\text{PSL}(2, q))$ for all these q . The values for $q = 4, 7, 8, 9, 11, 13$ are those of the table of [BN86, p. 246], which has no column $q = 5$; the value for $q = 5$ follows from $\text{PSL}(2, 5) \cong \text{PSL}(2, 4)$. For $q \leq 13$ the minimal laws found are $[x, r y] = [x, s y]$ with $s = 63, 172, 129, 124, 1986, 2191$ for $q = 4, 7, 8, 9, 11, 13$, as in [BN86], and those of $A_5 \cong \text{PSL}(2, 4)$ and $A_6 \cong \text{PSL}(2, 9)$ are the laws of [Nik84].

Take such a point. If q is even, it is (h, c) with $c \neq 0$, and we put $s = c^{q/2}$, so that $s^2 = c = 4 - s^2$; if q is odd, it is (h, s) and we put $c = 4 - s^2 \neq 0$. In both cases $s^2 \neq 4$ and $g_c^r(h) = -c$. By Lemma 3.3 there is x with $\text{tr}[x, y] = 2 + h$ for $y = y_s$. By Lemma 3.1(b),(c), $h_k = t_k - 2$ satisfies $h_1 = h$ and $h_{k+1} = g_c(h_k)$, so $h_{r+1} = g_c^r(h) = -c$. Since $g_c(-c) = 0 = g_c(0)$ and $-c \neq 0$, the point $-c$ is not g_c -periodic, so $t_{r+1} = 2 - c$ is not f_s -periodic, and Lemma 3.2 gives the tails. By Lemma 2.1, $\text{PSL}(2, q) \notin V_r$; as $Q \leq Q_r$, this holds for $q > Q_r$. Finally, let $q \geq 256$ and $r = \lfloor (\log_2 q)/4 \rfloor - 1 \geq 1$. Since $(d-1)(d-2) + 2 < d^2$, we have $Q < d^4 \leq 81 \cdot 2^{4r-4} < 2^{4r+4} \leq q$, so $\mu(\text{PSL}(2, q)) \geq r + 1$. $\square$

Remark 3.6. (1) Q_r is far from sharp. For $r = 1, \dots, 7$ the least primes p for which C has an $\mathbb{F}_p$ -point with $s \neq 0$ and $c \neq 0$ are 7, 7, 11, 11, 13, 19, 19; with $s = 0$ allowed, the least prime for $r = 1$ is 3 (the point $(h, s) = (1, 0)$, where $c = 1$), and the other six values do not change. Every prime $1000 < p < 3000$ gives such a point for $r \leq 4$, and the verification runs found that for $r = 3$ the odd prime powers $q \leq 30,000$ for which C has no $\mathbb{F}_q$ -point with $c \neq 0$ are 3, 5, 7, 9. (2) For $r = 3$, Theorem 3.5 gives only finiteness; the exact list $q \in \{4, 5, 8\}$ of [BN86] rests on their sharper count ($|N - q| < 12\sqrt{q} + 5$, via Schmidt), used for $q > 169$, on direct checks for the primes $7 < p < 168$, and on separate arguments for the powers of 3 and 5 and, with Hilbert's Theorem 90, for the powers of 2. (3) For q even, g_c is additive, and $\mu(\text{PSL}(2, 2^k))$ grows slowly (Table 1).

4 The other families

By the classification of finite simple groups (see [Wil09, Section 1.2] for the statement), a nonabelian finite simple group is an alternating group, a classical group, an exceptional group of Lie type, or one of the 26 sporadic groups, the Tits group ${}^2F_4(2)'$ being counted with the sporadic groups below. Up to isomorphism the classical groups are $\text{PSL}_N(q)$ ($N \geq 2$), $\text{PSU}_N(q)$ ($N \geq 3$), $\text{PSp}_N(q)$ ($N \geq 4$ even) and $\text{P}\Omega_N^\epsilon(q)$ ($N \geq 7$), where N is the dimension of the natural module. We use the following standard facts on subgroups. They are stated with references for the reader; *the theorem and section numbers in these references were not checked against the books and papers.*

- (C1) The even permutation matrices in $\text{SL}_N(q)$ form a subgroup A_N that meets the scalars trivially; so A_N embeds in $\text{PSL}_N(q)$.
- (C2) A nondegenerate hermitian form over $\mathbb{F}_{q^2}$ has an orthonormal basis [KL90, Section 2.3]; the even permutation matrices in this basis lie in $\text{SU}_N(q)$, and A_N embeds in $\text{PSU}_N(q)$.
- (C3) In a symplectic basis $e_1, f_1, \dots, e_n, f_n$ the maps $e_i \mapsto e_{\pi(i)}$, $f_i \mapsto f_{\pi(i)}$ ($\pi \in S_n$) are isometries; for $n \geq 5$, A_n embeds in $\text{PSp}_{2n}(q)$.

- (C4) An orthogonal space V of dimension N has Witt index $l \geq \lfloor N/2 \rfloor - 1$. Permuting l hyperbolic pairs and fixing their orthogonal complement gives $S_l \leq O(V)$. For $l \geq 5$ the group A_l is perfect, so $A_l \leq [O(V), O(V)] \leq \Omega(V)$, as $O(V)/\Omega(V)$ is abelian [KL90, Section 2.5]; and A_l embeds in $\mathrm{P}\Omega(V)$.
- (C5) $A \mapsto \mathrm{diag}(A, I_{N-2})$ embeds $\mathrm{SL}_2(q)$ in $\mathrm{SL}_N(q)$; for $N \geq 3$ its image meets the scalars trivially. So $\mathrm{PSL}_N(q)$ has a section $\mathrm{PSL}_2(q)$.
- (C6) $\mathrm{Sp}_2(q) = \mathrm{SL}_2(q)$, acting on a hyperbolic plane and trivially on its orthogonal complement, is a subgroup of $\mathrm{Sp}_N(q)$; so $\mathrm{PSp}_N(q)$ has a section $\mathrm{PSL}_2(q)$.
- (C7) $\mathrm{SU}_2(q) \cong \mathrm{SL}_2(q)$ [KL90, Proposition 2.9.1]; acting on a nondegenerate 2-space and trivially on its orthogonal complement, it is a subgroup of $\mathrm{SU}_N(q)$, $N \geq 3$; so $\mathrm{PSU}_N(q)$ has a section $\mathrm{PSL}_2(q)$.
- (C8) For $N \geq 7$, V contains a nondegenerate 4-space W of plus type, and $\Omega(W) \cong \Omega_4^+(q)$ is a central product $\mathrm{SL}_2(q) \circ \mathrm{SL}_2(q)$ [KL90, Proposition 2.9.1]. For $q \geq 4$ a factor $\mathrm{SL}_2(q)$ is perfect, so it lies in $[O(V), O(V)] \leq \Omega(V)$; hence $\mathrm{P}\Omega(V)$ has a section $\mathrm{PSL}_2(q)$.
- (E1) In the untwisted groups $G_2(q)$, $F_4(q)$, $E_6(q)$, $E_7(q)$, $E_8(q)$, the root subgroups X_α , $X_{-\alpha}$ of a root α generate a homomorphic image of $\mathrm{SL}_2(q)$ [Car72, Chapter 6]; so these groups have a section $\mathrm{PSL}_2(q)$.
- (E2) ${}^3D_4(q)$ contains $G_2(q)$ [Kle88b]; use (E1).
- (E3) ${}^2E_6(q)$ contains $F_4(q)$: F_4 is the centralizer of the graph automorphism σ of E_6 , and the Steinberg endomorphism $\sigma \circ \mathrm{Frob}_q$ acts on it as Frob_q ; use (E1).
- (E4) ${}^2F_4(q)$, $q = 2^{2m+1} \geq 8$, contains $\mathrm{Sp}_4(q)$ [Mal91] (alternatively, a parabolic subgroup with Levi factor $\mathrm{GL}_2(q)$); use (C6).
- (E5) ${}^2G_2(q)$, $q = 3^{2m+1} \geq 27$, contains $C_2 \times \mathrm{PSL}_2(q)$, the centralizer of an involution [War66, Kle88a].

Theorem 4.1. *For every $r \geq 1$ only finitely many nonabelian finite simple groups that are not isomorphic to a Suzuki group ${}^2B_2(q) = \mathrm{Sz}(q)$ lie in V_r . More precisely, such a group in V_r is sporadic, or A_m with $m \leq 2^{r+1} + 1$, or of Lie type with parameter $q \leq Q_r$, of natural dimension $N \leq 2^{r+2} + 5$ if it is classical.*

Proof. Let S be such a group with $\mu(S) \leq r$ (Lemma 2.1). (i) If $S = A_m$, then $\lfloor \log_2(m-2) \rfloor \leq r$ by Lemma 2.4(c), so $m-2 < 2^{r+1}$. (ii) Let S be classical with natural dimension N . By (C1)–(C4), S contains A_l with $l = N$, N , $N/2$, resp. $l \geq \lfloor N/2 \rfloor - 1$, in each case $l \geq (N-3)/2$ (for (C3) and (C4) provided $l \geq 5$). If $l \geq 6$, then $l \leq 2^{r+1} + 1$ by Lemma 2.3 and Lemma 2.4(c), so $N \leq 2l+3 \leq 2^{r+2} + 5$; if $l \leq 5$, then $N \leq 13 \leq 2^{r+2} + 5$. By (C5)–(C8), S has a section $\mathrm{PSL}_2(q)$ (for $\mathrm{PSL}_2(q)$ itself trivially; in (C8) we may assume $q \geq 4$), so $q \leq Q_r$ by Lemma 2.3 and Theorem 3.5. There are finitely many such S . (iii) Let S be exceptional, with parameter q , and not of type ${}^2B_2$. There are finitely many types, and S is determined by its type and q ; by (E1)–(E5) and (C6), S has a section $\mathrm{PSL}_2(q)$, so $q \leq Q_r$. (iv) There are finitely many sporadic groups. $\square$

Proof of Theorem 1.1 and Corollary 1.2. Theorem 1.1 is Theorem 4.1. If $\mu(\mathrm{Sz}(q)) \rightarrow \infty$, then each V_r contains only finitely many Suzuki groups, and by Theorem 1.1 only finitely many nonabelian simple groups. Otherwise there are r and infinitely many $q = 2^{2m+1}$ with $\mu(\mathrm{Sz}(q)) \leq r$, and V_r contains these pairwise non-isomorphic simple groups. $\square$

The Suzuki groups are missing from (E1)–(E5) for a reason: Proposition 5.1 below shows that for prime exponents they have no proper section of Engel depth larger than 3.

5 The Suzuki groups

Let $q = 2^n$, $n = 2m + 1 \geq 3$, and $\pi(a) = a^{2^{m+1}}$, so that $\pi^2(a) = a^2$. We use the matrices of [BGG⁺06, Section 4.1] (arXiv version):

$$U(a, b) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ a\pi(a) + b & \pi(a) & 1 & 0 \\ a^2\pi(a) + ab + \pi(b) & b & a & 1 \end{pmatrix}, \quad M(c) = \text{diag}(c^{1+2^m}, c^{2^m}, c^{-2^m}, c^{-1-2^m}),$$

and the antidiagonal permutation matrix T ; then $\text{Sz}(q) = \langle U(a, b), M(c), T \rangle \leq \text{SL}(4, q)$ [BGG⁺06, Section 4.1], and these matrices preserve the alternating form with Gram matrix T . The same matrices appear in [BGKJ10, Section 4], with the Bruhat decomposition $\text{Sz}(q) = B \cup BTQ$, where $Q = \{U(a, b)\}$ has order q^2 , $[Q, Q] = Z(Q)$ has order q , and $B = QH$ with $H = \{M(c)\}$ cyclic of order $q - 1$. For $a \in \mathbb{F}_{2^d}$ with $d \mid n$ and $d = 2m' + 1$ we have $\pi(a) = a^{2^{m'+1}}$, because $m - m' = d(n/d - 1)/2$ is divisible by d ; so the same matrices over $\mathbb{F}_{2^d}$ give $\text{Sz}(2^d) \leq \text{Sz}(2^n)$, and $\mu(\text{Sz}(2^n)) \geq \mu(\text{Sz}(2^d))$ by Lemma 2.3. By Suzuki's theorem [Suz62, Theorem 9] (the theorem number was not checked against the paper), every maximal subgroup of $\text{Sz}(q)$ is conjugate to B (a Frobenius group with kernel Q), to the dihedral group $D_{2(q-1)}$, to a Frobenius group $C_{q \pm 2^{m+1}+1} : C_4$, or to a subfield group $\text{Sz}(q_0)$ with $q = q_0^k$, k prime and $q_0 \geq 8$.

Proposition 5.1. *Let p be an odd prime and $q = 2^p$. Every proper section of $\text{Sz}(q)$ has Engel depth at most 3. Hence, if $x, y \in \text{Sz}(q)$ and $\text{tail}(x, y) \geq 4$, then $\langle x, y \rangle = \text{Sz}(q)$.*

Proof. As $\text{Sz}(q)$ is simple, a proper section is a section of a maximal subgroup, and for $n = p$ prime there are no subfield groups with $q_0 \geq 8$. By Lemma 2.3 it suffices to show $\mu \leq 3$ for the three types. *The group $B = QH$.* If $y \notin Q$, then y is conjugate in B to some $t \in H \setminus 1$, and we may assume $y = t$. Then $\varphi_t(B) \leq [B, B] \leq Q$, and φ_t is injective on Q : $z^{-1}t^{-1}zt = w^{-1}t^{-1}wt$ gives $wz^{-1} \in C_Q(t) = 1$. So φ_t permutes Q , every element of Q is periodic, and $\text{tail}(x, t) \leq 1$. If $y \in Q$, then $a_1 \in Q$, $a_2 \in [Q, Q] = Z(Q)$ and $a_3 = 1$, so $\text{tail}(x, y) \leq 3$. *The group D_{2N} , $N = q - 1$ odd,* with rotation subgroup R . For a reflection y , $a_1 \in R$ and $\varphi_y(\rho) = \rho^{-2}$ is a bijection of R , so $\text{tail} \leq 1$; for $y \in R$, $a_1 \in \{1, y^2\}$ and $a_2 = 1$, so $\text{tail} \leq 2$. *A Frobenius group $A : C_4$ with A cyclic of odd order.* If $y \in A$, then $a_1 \in A$ and $a_2 = 1$. Otherwise y acts on A as a fixed-point-free automorphism, so $u \mapsto [u, y]$ is injective on the abelian group A , and $a_1 \in A$ is periodic. If $\langle x, y \rangle$ is proper, then $\text{tail}(x, y) \leq \mu(\langle x, y \rangle) \leq 3$. $\square$

So for $r \geq 4$ the groups $\text{Sz}(2^p)$ can be shown not to lie in V_r only by generating pairs, and an argument as in Section 4 is impossible. The following two reductions are used in the computations.

Lemma 5.2. *Let $y \in G$, $\Psi_y(u) = y^{-1}uyu^{-1}y$ for u in the conjugacy class $\text{Cl}(y)$, and $\beta(u) = u^{-1}y$. Then Ψ_y maps $\text{Cl}(y)$ to itself, $[x, y] = \beta(x^{-1}yx)$, and $\varphi_y \circ \beta = \beta \circ \Psi_y$. Hence $\text{tail}(x, y) = 1 + \text{tail}_{\Psi_y}(x^{-1}yx)$ whenever the right-hand side is at least 2, and if $\tau_y = \max_u \text{tail}_{\Psi_y}(u) \geq 1$, then $\max_x \text{tail}(x, y) = 1 + \tau_y$.*

Proof. $[x, y] = (x^{-1}yx)^{-1}y$. Both $\varphi_y(\beta(u))$ and $\beta(\Psi_y(u))$ equal $y^{-1}uy^{-1}u^{-1}y^2$, and for $u = g^{-1}yg$ one has $\Psi_y(u) = k^{-1}yk$ with $k = [g, y]$. So β , which is injective, conjugates Ψ_y to the restriction of φ_y to $\varphi_y(G) = \beta(\text{Cl}(y))$, which contains all cycles of φ_y . Hence, for $k \geq 0$, $a_{k+1}(x, y)$ is periodic if and only if $\Psi_y^k(x^{-1}yx)$ is. If u has Ψ_y -tail $\tau \geq 1$ and $x^{-1}yx = u$, then a_1 is not periodic, so neither is a_0 , and $\text{tail}(x, y) = 1 + \tau$. $\square$

Lemma 5.3. *Let $C \leq C_G(y)$. Then φ_y induces a map on the orbits of C acting by conjugation, and the sequence of orbits of $a_k(x, y)$ has the same tail as $(a_k(x, y))$. In particular, if the orbits O_k of a_k satisfy $O_{\mu+\lambda} = O_\mu$ and $O_{\mu-1+\lambda} \neq O_{\mu-1}$ for some $\mu, \lambda \geq 1$, then $\text{tail}(x, y) = \mu$.*

q	y	pairs tried	best tail(x, y)	λ	tail / $q^{3/2}$
8	all classes	exhaustive	62	90	2.74
32	all classes	exhaustive	527	1275	2.91
128	three tori, full storage	45	3759	52324	≥ 2.59
128	$y = M(c)$, H -orbits	150	3554	412	≥ 2.45
512	$y = M(c)$, H -orbits	40	24814	291	≥ 2.14
2048	$y = M(c)$, H -orbits	8	183810	10555	≥ 1.98
8192	$y = M(c)$, H -orbits	5	792500	279404	≥ 1.06

Table 2: Engel depths of $\text{Sz}(q)$: exact values for $q = 8, 32$ and lower bounds from explicit pairs. λ is the period of the cycle reached (of the H -orbits in the rows with $y = M(c)$). The two rows for $q = 128$ come from two runs: the bound 3759 is from the run with y in all three types of tori and full storage of the sequence, the pair with tail 3554 from the run with the H -orbit method (its period in G is 52324). The pairs are printed in the recorded outputs and in `check_note.py`; there an element of $\mathbb{F}_{2^n}$ is an integer whose binary digits are its coordinates in the basis $1, X, \dots, X^{n-1}$ of $\mathbb{F}_2[X]/(f)$, with $f = X^3 + X + 1, X^5 + X^2 + 1, X^7 + X + 1, X^9 + X^4 + 1, X^{11} + X^2 + 1, X^{13} + X^4 + X^3 + X + 1$ for $n = 3, 5, \dots, 13$.

Proof. $\varphi_y(c^{-1}zc) = c^{-1}\varphi_y(z)c$ for $c \in C$. If $a_{k+m} = c^{-1}a_kc$, then $a_{k+jm} = c^{-j}a_kc^j$ for all j , so $a_{k+|c|m} = a_k$; conversely periodic elements have periodic orbits. If $O_{\mu+\lambda} = O_\mu$, then O_μ is periodic with period dividing λ ; if $O_{\mu-1}$ were periodic, it would lie on the same cycle, and $O_{\mu-1+\lambda} = O_{\mu-1}$ would follow. $\square$

For $y = M(c)$ we take $C = H$, which is abelian and contains y . In the programs an H -orbit is represented by the normal form in which the first nonzero off-diagonal entry is 1 (conjugation by $M(t)$ multiplies the entries by t^e with $e \in \{\pm 1, \pm 2^{m+1}, \pm(1 + 2^{m+1}), \pm 2(1 + 2^m)\}$, all prime to $q - 1$), and tails are found by Brent's cycle detection [Bre80], with constant memory.

5.1 Computations

Exact values. Using Lemma 5.2 for all conjugacy classes, the programs of Section 6 give $\mu(\text{Sz}(8)) = 62$ and $\mu(\text{Sz}(32)) = 527$. For $q = 8$ the maxima of $\max_x \text{tail}(x, y)$ over the classes of y are 2 (involutions), 11 and 8 (order 4), 62 (order 5), 56 (order 7) and 27 (order 13), in line with the bound 11 of [BN86]. For $q = 32$ they are 2; 101, 103 (order 4); 303, 323, 442 (order 31); 293, 334 (order 41); 527 (order 25); 249 (order 5); the class sizes add up to $|\text{Sz}(32)| - 1$. Each maximum is attained by an explicit pair, which was re-checked by direct iteration.

Certified lower bounds. For larger q we computed the exact tails of random pairs; each pair in Table 2 certifies $\mu(\text{Sz}(q)) \geq \text{tail}(x, y)$. The certificate consists of a word $x = U(a, b)M(k)$ or $x = U(a, b)M(k)TU(a', b')$, and the same for y , which shows $x, y \in \text{Sz}(q)$ by multiplying out, and of the two comparisons $a_{\mu+\lambda} = a_\mu$, $a_{\mu-1+\lambda} \neq a_{\mu-1}$, for H -orbits in the case $y = M(c)$ (Lemma 5.3). The program `check_note.py` checks these certificates for all pairs of Table 2 in about 20 seconds.

Corollary 5.4. *Let $n \geq 3$ be odd. Then $\text{Sz}(2^n) \notin V_r$ if n has a divisor $d \in \{3, 5, 7, 9, 11, 13\}$ with $r < B_d$, where $(B_3, B_5, B_7, B_9, B_{11}, B_{13}) = (62, 527, 3759, 24814, 183810, 792500)$. For $3 \leq r \leq 61$ the Suzuki groups whose membership in V_r is not decided by these results are exactly the $\text{Sz}(2^n)$ with $\gcd(n, 15015) = 1$, the smallest being $\text{Sz}(2^{17})$; for $r \leq 2$ there are none [Bra83b, Lemma 2].*

Proof. $\text{Sz}(2^d) \leq \text{Sz}(2^n)$ and Table 2, with Lemma 2.1; note $15015 = 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13$ and $B_d \geq 62$ for all six d . $\square$

For $r \geq 62$ the group $\text{Sz}(8)$ itself lies in V_r , and the undecided set is larger. The ratios in Table 2 are consistent with the following heuristic, which we do not claim to be more than that: by Lemmas 5.2 and 5.3 the dynamics lives on the about q^3 orbits of H on $\text{Cl}(y)$, and the largest tail of a random map on N points is of order $\sqrt{N}$; this suggests $\mu(\text{Sz}(q)) \approx c q^{3/2}$. For $\text{PSL}(2, p)$ the same heuristic, with the p values of the trace, suggests $\mu \approx c\sqrt{p}$, in line with Table 1. The bounds for $q \geq 128$ come from few random pairs and are certainly not optimal.

5.2 What a proof for the Suzuki groups would need

A proof in the style of Theorem 3.5 would need, for every r , rational points on a variety of Engel chains of length r that end at a non-periodic element. For $\text{Sz}(q)$ there is no trace map [BGKJ10, Section 4]; the reduced dynamics of Lemmas 5.2 and 5.3 lives on a variety of dimension 6 over $\overline{\mathbb{F}}_2$, with the twisted Frobenius of the Suzuki groups instead of Frob_q . One would need an absolutely irreducible component of the chain variety that is stable under this Frobenius, and a Lang–Weil estimate for the twisted Frobenius, as Bandman, Greuel, Grunewald, Kunyavskii, Pfister and Plotkin obtain for their equations from the trace formula of Deligne’s conjecture (proved by Pink and Fujiwara) [BGG⁺06, Section 2.4]. We did not find a way to control the irreducibility for all r . Concretely, let $y = M(c) \neq 1$, let 0 and ∞ be its two fixed points on the Suzuki–Tits ovoid, and $B_\infty = U_\infty H$ the stabilizer of ∞ , with normal Sylow 2-subgroup U_∞ . If $g = [x, y]$ maps ∞ to 0, then $a_{j+1} = g^{-1}y^{-1}gy$ fixes ∞ , and its eigenvalue on the line ∞ is the product of those of y^{-1} on 0 and of y on ∞ , that is $c^{\pm 2(1+2^m)}$; this is $\neq 1$, as $c \neq 1$ and $2(1+2^m)$ is prime to $q-1$. So $a_{j+1} \in B_\infty \setminus U_\infty$, while $a_{j+2} \in [B_\infty, B_\infty] \leq U_\infty$, and φ_y maps U_∞ bijectively onto itself, since y acts on U_∞ without nontrivial fixed points (as in the proof of Proposition 5.1). Hence a_{j+1} is not periodic and $\text{tail}(x, y) = j+2$. So it would suffice that the equation $[x, y](\infty) = 0$ has a solution $x \in \text{Sz}(q)$ for every j once q is large. We do not know whether it has.

6 Computations and reproducibility

Theorems 1.1, 3.5 and 4.1, Corollary 1.2 and Proposition 5.1 do not depend on computations. Tables 1 and 2 and Corollary 5.4 are results of exact finite computations; they serve as checks of the normalisation and as evidence on the open case. The package (folder `reproducibility/` of the source archive) has four parts.

(1) `check_note.py`, written with this note (standard library only, about 20 seconds, 28 checks, all passed): Fricke’s identity and Lemma 3.1 on 2850 random pairs over 19 fields; the construction of Lemma 3.3 for all 5648 pairs (s, t) over the 16 fields with $q \leq 32$; $\mu(\text{PSL}(2, q))$ and $\mu(\text{SL}(2, q))$ for $q \leq 16$, with the minimal laws of Table 1; Lemma 2.4(b),(c); the Eisenstein conditions, Newton polygons and degrees of Lemma 3.4 for small r ; the constants and the counting inequality of Theorem 3.5, and its construction in 38 cases (r, q) , 22 of them with $q > Q_r$; the order 29120 of the group of Section 5 for $q = 8$; and the certificates of all pairs of Table 2, which are compared with the recorded outputs. (2) The programs with which the results were first obtained (`original/`; Python 3, with `numpy` for the exact computations; from under a second to five minutes each). (3) The programs and recorded outputs of the first two verification runs (`independent_runs/`). (4) Those of the third verification run (`independent_run_2/`; Python 3 with `numpy`, `scipy` and `sympy`; about nine minutes in all).

Conclusion. The note gives a *partial answer to part b)*: for every r , all but finitely many nonabelian finite simple groups in V_r are Suzuki groups, so *b)* holds if and only if the Engel depth of $\text{Sz}(q)$ tends to infinity; this last statement and part a) stay open. The Suzuki groups are not reached by sections

(Proposition 5.1), and the computations of Table 2 are evidence only: they do not decide the Suzuki case.

Verification. The results were first obtained, with proofs and with the programs in `original/`. Two independent verification runs, both AI-assisted, followed; each re-derived the proofs and wrote its own programs. The first run compared the question and the definitions with [KM26, Bra83b, BN86] and the cited statements with [BGG12, (3.3), (5.8)] and [BGKJ10, Theorem 3.4]; its programs recomputed Table 1 for $q \leq 41$, Lemma 3.3 for non-prime fields, the construction of Theorem 3.5 for five pairs (r, p) with p up to 10,007 (tails exactly $r + 2$), $\text{Sz}(8)$ from scratch, all 34 nontrivial classes of $\text{Sz}(32)$ with its own representatives, and every pair of Table 2, with checks that the matrices preserve the form and the Suzuki–Tits ovoid. The second run wrote its programs before reading the original ones; it recomputed all of Table 1, the minimal laws of [BN86, Nik84], $\mu(A_7) = 15$, $\text{Sz}(8)$ and $\text{Sz}(32)$ in its own construction, the pairs of Table 2 with membership certificates in Bruhat form, and Proposition 5.1 for $q = 8, 32$; its programs reject eleven deliberately wrong inputs, and it re-ran the original programs from a copy, with identical outputs. All values agree. Neither run found an error in a proof. Their required corrections were made: the description of the table of [BN86], the unchecked numbering in Section 4 and the proviso $l \geq 5$, the two runs for $q = 128$ in Table 2, the condition $s \neq 0$ in Remark 3.6(1), the sentences on [BGKJ10, BGG12], and the references [Mal91, AP96, Nik84, Bra83a]. This text was written after these runs, with these changes: Lemma 3.3 has a self-contained proof (in the runs it rested on Macbeath’s theorem and on a citation for binary quadratic forms), Lemma 3.4 is proved by Eisenstein’s criterion (the runs checked the equivalent Newton-polygon argument), Lemma 3.2(c) adds the case of even q , and the bound $\mu(\text{PSL}(2, q)) \geq \lfloor (\log_2 q)/4 \rfloor$ of Theorem 3.5 and the eigenvalue argument in Section 5.2 were added. All but the last are checked by `check_note.py`.

A third independent verification run, also AI-assisted, then examined this text, the five changes first; it used none of the earlier programs. It re-derived every proof line by line, read [Bra83b, BN86] and the cited parts of [BGG12, BGKJ10, BGG⁺06] again, checked all DOIs of the bibliography against Crossref and repeated the literature search. Its programs (`independent_run_2/`) ran the construction of Lemma 3.3 as written for all 28,425 pairs (s, t) over the 27 fields with $q \leq 64$; checked the Eisenstein conditions and degrees of Lemma 3.4 for $r \leq 6$, and Lemma 3.2 on all pairs of $\text{SL}(2, q)$ for $q \leq 13$; recomputed Table 1 with its minimal laws, $\mu(A_7)$ and the numbers of Remark 3.6(1); ran the construction of Theorem 3.5 in 24 cases with $q > Q$ and $r \leq 5$, and with $r = \lfloor (\log_2 q)/4 \rfloor - 1$ in 11 fields with $256 \leq q \leq 2^{20} + 7$, finding the exact tails $r + 2$ by iteration; computed $\text{Sz}(8)$ in full ($\mu = 62$, with the class maxima of Section 5.1); tested Proposition 5.1 for $q = 8$ by computing $\langle x, y \rangle$ for all pairs, without Suzuki’s list, and the argument of Section 5.2 on all pairs with $y = M(c)$ in $\text{Sz}(8)$ and on random pairs in $\text{Sz}(32)$ and $\text{Sz}(128)$; and determined the field representation of the printed pairs from the data and re-checked every pair of Table 2, and the recorded class witnesses of $\text{Sz}(8)$ and $\text{Sz}(32)$, with its own membership certificates and exact tails. It also re-ran the whole package from the source archive, with identical outputs. It found no error in a theorem, proposition or lemma. Its required corrections were made: the least prime for $s = 0$ in Remark 3.6(1) (it is 3, not 5), the condition $q \geq 4$ in the abstract, the priority statement for $r \leq 2$, two ratios in Table 2 that had been rounded upwards, the credit to Gupta, the sentence on the citations of Brandl in [BGKJ10], the unchecked theorem number of [Suz62], two justifications in Section 5.2, the description of the proof of [BN86] in Remark 3.6(2), and the field representation in Table 2.

Scope and priority. The note answers part b) of Problem 11.16 for all families except the Suzuki groups (Theorem 1.1), reduces it to the growth of $\mu(\text{Sz}(q))$ (Corollary 1.2) and leaves that and part a) open. The problem is due to R. Brandl [KM26]; the case $r \leq 2$ is his [Bra83b]; the case $r = 3$ for $\text{PSL}(2, q)$, the conjecture and the table of minimal laws are due to Brandl and Nikolova [BN86], and

the depths of A_5 and A_6 to Nikolova [Nik84]. Theorem 3.5 extends the method of [BN86] with the trace map of [BGG⁺06, BGKJ10, BGG12]; the other sources are named where they are used. For $r \leq 2$, Theorems 3.5 and 4.1 follow from [Bra83b], since V_2 contains no nonabelian simple group, and for $r = 3$ the conclusion $\text{PSL}(2, q) \notin V_3$ of Theorem 3.5 is contained in the theorem of [BN86]. To our knowledge, Theorem 3.5 for $r \geq 4$, Theorem 4.1 for $r \geq 3$, Corollary 1.2, Proposition 5.1 and Table 2 are not in print; Theorem 3.5 is a direct extension of [BN86], and experts may regard it as routine.

We read [Bra83b, BN86] in full (the publisher’s open files), Problem 11.16 in versions 48 and 49 of [KM26], and, in the arXiv versions, of [BGG12] (v2) Sections 3–5 in part, with (3.3), Proposition 4.9 and (5.8), of [BGKJ10] (v2) Theorem 3.4 with its proof and the beginning of Section 4, and of [BGG⁺06] (math/0303165v1, 2003) Sections 1, 2 and the beginning of 4.1, to which our section numbers refer. Of [Nik84] we read the zbMATH review; of [GKNP03, GKP13] the verification runs read the abstracts; [Bra83a, KS16, HN67, Gup66] are cited from their bibliographic data (and [HN67] from [BN86], [Gup66] from [Bra83b]); [Mac69, AP96, GL02, LY94] are cited as stated in [BGKJ10, BGG12]. The books and the papers on maximal subgroups were not opened. The searches (21 and 22 October 2026, by the verification runs and at the writing) used zbMATH Open (“Engel depth” gives only [Bra83b, BN86]), OpenAlex (the 7 works citing [Bra83b] and the 6 citing [BN86]), Semantic Scholar, the arXiv API, Crossref and a few web searches; mathnet.ru refused our automated requests. No printed answer to the question was found in the sources searched; a search that finds nothing is not a proof of novelty, and we claim no priority.

AI-assisted tools were used to search the problem corpus and the literature, to find the arguments, to write and run the verification programs, and to help prepare this manuscript. The author has reviewed the proofs, the computations and the sources and is responsible for all claims. The note has not been peer reviewed.

References

- [AP96] Yves Aubry and Marc Perret. A Weil theorem for singular curves. In R. Pellikaan, M. Perret, and S. G. Vlăduţ, editors, *Arithmetic, Geometry, and Coding Theory*, pages 1–8. Walter de Gruyter, Berlin, New York, 1996. doi:10.1515/9783110811056.1.
- [BGG⁺06] Tatiana Bandman, Gert-Martin Greuel, Fritz Grunewald, Boris Kunyavskii, Gerhard Pfister, and Eugene Plotkin. Identities for finite solvable groups and equations in finite simple groups. *Compositio Mathematica*, 142(3):734–764, 2006. arXiv version: math/0303165 (v1, 2003), titled “Engel-like identities characterizing finite solvable groups”. doi:10.1112/S0010437X0500179X.
- [BGG12] Tatiana Bandman, Shelly Garion, and Fritz Grunewald. On the surjectivity of Engel words on $\text{PSL}(2, q)$. *Groups, Geometry, and Dynamics*, 6(3):409–439, 2012. arXiv:1008.1397, version 2. doi:10.4171/GGD/162.
- [BGKJ10] Tatiana Bandman, Fritz Grunewald, Boris Kunyavskii, and Nathan Jones. Geometry and arithmetic of verbal dynamical systems on simple groups. *Groups, Geometry, and Dynamics*, 4(4):607–655, 2010. arXiv:0809.0369, version 2. doi:10.4171/GGD/98.
- [BN86] Rolf Brandl and Daniela B. Nikolova. Simple groups of small Engel depth. *Bulletin of the Australian Mathematical Society*, 33(2):245–251, 1986. doi:10.1017/S0004972700003117.
- [Bra83a] Rolf Brandl. Engel cycles in finite groups. *Archiv der Mathematik*, 41(2):97–102, 1983. doi:10.1007/BF01196864.

- [Bra83b] Rolf Brandl. On groups with small Engel depth. *Bulletin of the Australian Mathematical Society*, 28(1):101–110, 1983. doi:10.1017/S0004972700026150.
- [Bre80] Richard P. Brent. An improved Monte Carlo factorization algorithm. *BIT*, 20(2):176–184, 1980. doi:10.1007/BF01933190.
- [Car72] Roger W. Carter. *Simple Groups of Lie Type*. John Wiley & Sons, London, 1972.
- [GKNP03] F. Grunewald, B. Kunyavskii, D. Nikolova, and E. Plotkin. Two-variable identities in groups and Lie algebras. *Journal of Mathematical Sciences*, 116(1):2972–2981, 2003. doi:10.1023/A:1023450709743.
- [GKP13] Fritz Grunewald, Boris Kunyavskii, and Eugene Plotkin. Characterization of solvable groups and solvable radical. *International Journal of Algebra and Computation*, 23(5):1011–1062, 2013. doi:10.1142/S0218196713300016.
- [GL02] Sudhir R. Ghorpade and Gilles Lachaud. Étale cohomology, Lefschetz theorems and number of points of singular varieties over finite fields. *Moscow Mathematical Journal*, 2(3):589–631, 2002. doi:10.17323/1609-4514-2002-2-3-589-631.
- [Gup66] N. D. Gupta. Some group-laws equivalent to the commutative law. *Archiv der Mathematik*, 17(2):97–102, 1966. doi:10.1007/BF01899854.
- [HN67] Hermann Heineken and Peter M. Neumann. Identical relations and decision procedures for groups. *Journal of the Australian Mathematical Society*, 7(1):39–47, 1967. doi:10.1017/S1446788700005073.
- [KL90] Peter Kleidman and Martin Liebeck. *The Subgroup Structure of the Finite Classical Groups*, volume 129 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1990. doi:10.1017/CB09780511629235.
- [Kle88a] Peter B. Kleidman. The maximal subgroups of the Chevalley groups $G_2(q)$ with q odd, the Ree groups ${}^2G_2(q)$, and their automorphism groups. *Journal of Algebra*, 117(1):30–71, 1988. doi:10.1016/0021-8693(88)90239-6.
- [Kle88b] Peter B. Kleidman. The maximal subgroups of the Steinberg triality groups ${}^3D_4(q)$ and of their automorphism groups. *Journal of Algebra*, 115(1):182–199, 1988. doi:10.1016/0021-8693(88)90290-6.
- [KM26] E. I. Khukhro and V. D. Mazurov. Unsolved Problems in Group Theory. The Kurovka Notebook, No. 21. Sobolev Institute of Mathematics, Novosibirsk; versions 48 (6 October 2026) and 49 (20 October 2026) on arXiv, 2026. Problem 11.16 (R. Brandl). arXiv:1401.0300.
- [KS16] E. I. Khukhro and P. Shumyatsky. Almost Engel finite and profinite groups. *International Journal of Algebra and Computation*, 26(5):973–983, 2016. doi:10.1142/S0218196716500405.
- [LY94] David B. Leep and Charles C. Yeomans. The number of points on a singular curve over a finite field. *Archiv der Mathematik*, 63(5):420–426, 1994. doi:10.1007/BF01196671.
- [Mac69] A. M. Macbeath. Generators of the linear fractional groups. In *Number Theory*, volume 12 of *Proceedings of Symposia in Pure Mathematics*, pages 14–32. American Mathematical Society, Providence, RI, 1969. doi:10.1090/pspum/012/0262379.

- [Mal91] Gunter Malle. The maximal subgroups of ${}^2F_4(q^2)$. *Journal of Algebra*, 139(1):52–69, 1991. doi:10.1016/0021-8693(91)90283-E.
- [Nik84] Daniela B. Nikolova. Computation of a commutator identity in alternating groups on an electronic computer. *Serdica*, 10:28–40, 1984. In Russian; English title as in zbMATH, Zbl 0557.20022.
- [Suz62] Michio Suzuki. On a class of doubly transitive groups. *Annals of Mathematics (2)*, 75(1):105–145, 1962. doi:10.2307/1970423.
- [ula26] ulamai. UnsolvedMath. Hugging Face dataset, <https://huggingface.co/datasets/ulamai/UnsolvedMath>, 2026. Version 1.8.0; record KOU-11.16.
- [War66] Harold N. Ward. On Ree’s series of simple groups. *Transactions of the American Mathematical Society*, 121(1):62–89, 1966. doi:10.2307/1994333.
- [Wil09] Robert A. Wilson. *The Finite Simple Groups*, volume 251 of *Graduate Texts in Mathematics*. Springer, London, 2009. doi:10.1007/978-1-84800-988-2.