# NIELSEN EQUIVALENCE OF GENERATING PAIRS OF $\mathrm{PSL}(2,q)$

DARRYL MCCULLOUGH AND MARCUS WANDERLEY

ABSTRACT. For a pair of generators of $\mathrm{PSL}(2,q)$, the trace of their commutator is a well-defined element of the coefficient field $\mathbb{F}_q$. We determine, for all $q$, which elements of $\mathbb{F}_q$ occur as these traces. In most cases, including all $q > 11$, every element except 2 occurs. The orbit of this trace under the action of the Frobenius automorphism is an invariant of the weak Nielsen equivalence class of the generating pair. There is a simple formula to compute the number of orbits. This gives lower bounds for the number of distinct free actions of $\mathrm{PSL}(2,q)$ on the 3-dimensional handlebody of genus $1 + |\,\mathrm{PSL}(2,q)\,|$. These show that the groups $\mathrm{PSL}(2,q)$ have many more distinct actions, relative to their orders, than previously known examples.

## INTRODUCTION

For a two-generator group, two generating pairs are called Nielsen equivalent if one can be obtained from the other by a sequence of operations of replacing one generator by its inverse, or by its product with the other generator. The equivalence classes so obtained have been useful both in algebraic contexts [3, 9, 11, 30, 34, 36, 40] and in topological ones [18, 19, 20, 21, 22, 23, 27, 29]. The conjugacy classes of the commutator of the generators and its inverse form an invariant of the Nielsen equivalence class [32], which we call the *Nielsen invariant.* When the group is $\mathrm{SL}(2,q)$ or $\mathrm{PSL}(2,q)$, all elements in these conjugacy classes have the same trace, giving an invariant of the Nielsen equivalence class which is a single element of the coefficient field $\mathbb{F}_q$. Our main result tells precisely which elements occur as trace invariants:

**Main Theorem.** *The elements of $\mathbb{F}_q$ that occur as trace invariants of generating pairs of $\mathrm{PSL}(2,q)$ are as follows:*

i) *For $q = 2$, $q = 4$, $q = 8$, and all $q > 11$, all elements except 2 occur.*
ii) *For $q = 3$, $q = 9$, and $q = 11$, all elements except 1 and 2 occur.*
iii) *For $q = 5$, only 1 and 3 occur.*
iv) *For $q = 7$, all elements except 0, 1, and 2 occur.*

*Date*: October 22, 2003.

1991 *Mathematics Subject Classification.* Primary 57M60; Secondary 20G40.

*Key words and phrases.* Nielsen, Nielsen equivalence, special linear group, projective, generating set, Frobenius automorphism, orbit, handlebody, group action, free.

The first author was supported in part by NSF grant DMS-0102463.

For most applications, one needs not Nielsen equivalence but weak Nielsen equivalence, which means that the generating pairs are Nielsen equivalent after changing one of them by an automorphism of the group. In our case, the automorphisms are well-understood, by the following result due to Schreier and van der Waerden [37] (see also [6] and the appendix to [13]).

**Theorem.** *Every automorphism of* $\mathrm{SL}(2,q)$ *or of* $\mathrm{PSL}(2,q)$ *has the form* $A \mapsto PA^{\phi}P^{-1}$*, where* $P$ *is an element of* $\mathrm{GL}(2,q)$*, and* $A^{\phi}$ *is the matrix obtained by applying an automorphism* $\phi$ *of* $\mathbb{F}_q$ *to each entry of* $A$*.*

We define the *weak trace invariant* of a generating pair of $\mathrm{PSL}(2,q)$ to be the set of elements of $\mathbb{F}_q$ equivalent to their trace invariant under automorphisms of $\mathbb{F}_q$. The Schreier-van der Waerden theorem shows that the weak trace invariant is an invariant of the weak Nielsen equivalence class of the generating pair.

It is well-known that for $q = p^s$, the automorphism group of $\mathbb{F}_q$ is cyclic of order $s$, generated by the Frobenius automorphism $\sigma$ that sends $x$ to $x^p$. We denote by $\Psi_q$ the number of orbits of the action of $\sigma$ on $\mathbb{F}_q$. Our Main Corollary follows directly from the Main Theorem:

**Main Corollary.** *The number of orbits of the Frobenius automorphism that occur as weak trace invariants of generating pairs of* $\mathrm{PSL}(2,q)$ *are as follows:*

i) *If* $q = 2$*,* $q = 4$*,* $q = 8$*, or* $q > 11$*, then* $\Psi_q - 1$ *orbits occur.*
ii) *If* $q = 3$*,* $q = 9$*, or* $q = 11$*, then* $\Psi_q - 2$ *orbits occur.*
iii) *If* $q = 5$ *or* $q = 7$*, then* $\Psi_q - 3$ *orbits occur.*

Since $\sigma$ has order $s$, the number $\dfrac{q}{s}$ is an obvious lower bound for $\Psi_q$. In section 5 we see that

$$\Psi_q = \frac{1}{s}\sum_{r|s}\varphi(s/r)\ p^r\ ,$$

where $\varphi(n)$ is the Euler function defined by $\varphi(1) = 1$ and $\varphi(n)$ is the number of positive integers less than $n$ and relatively prime to $n$ when $n > 1$. As we will note at the end of section 5, $\dfrac{q}{s}$ is a very accurate estimate of $\Psi_q$, except for some small values of $q$.

Whenever $\{A, B\}$ is a generating pair of $\mathrm{PSL}(2,q)$ and $\widetilde{A}$ and $\widetilde{B}$ are elements of $\mathrm{SL}(2,q)$ that project to $A$ and $B$, $\{\widetilde{A}, \widetilde{B}\}$ is also a generating pair of $\mathrm{SL}(2,q)$. Consequently the Main Theorem and Main Corollary immediately imply the same statements with $\mathrm{PSL}(2,q)$ replaced by $\mathrm{SL}(2,q)$.

The Main Corollary gives a lower bound for the number of $\mathrm{PSL}(2,q)$-defining subgroups of the free group on two generators, up to automorphism (see the discussion of $G$-defining subgroups in section 1). In section 8, we give an application which was the original motivation for our work. As we will explain there, it is known that the weak Nielsen equivalence classes of generating vectors of minimal cardinality of a finite group $G$ correspond to the distinct (orientation-preserving) free actions of $G$ on the lowest genus

of 3-dimensional orientable handlebody admitting a free $G$-action (for two-generator groups, this genus is $1 + |G|$). Thus the numbers in the Main Corollary furnish lower bounds for the number of minimal-genus free actions of the groups $\mathrm{PSL}(2, q)$. In particular, we see in section 9 that these groups have many more distinct free actions, relative to their orders, than previously known examples.

Section 1 contains a review of Nielsen equivalence and the Nielsen invariant, and gives some basic properties of the trace invariant. Sections 2 through 4 comprise the proof of the Main Theorem. As mentioned above, section 5 develops the formula for $\Psi_q$. The relation between the trace invariant and the Nielsen invariant is examined in section 6, where we obtain a very restricted class of examples of distinct Nielsen equivalence classes with the same trace invariant. In section 7 we discuss Nielsen equivalence in $A_5$, in particular noting that all equivalence classes are distinguished by the Nielsen invariant, and by the trace invariant when this group is regarded as $\mathrm{PSL}(2, 4)$, but not when it is regarded as $\mathrm{PSL}(2, 5)$. Section 8 contains our application to group actions on 3-dimensional handlebodies, and the final section is a discussion of some open questions.

We thank the referee of [27] for suggesting the investigation of Nielsen equivalence in $\mathrm{PSL}(2, q)$. We also thank Marston Conder, Gareth Jones, Andy Magid, Harald Niederreiter, and Günter Pilz for valuable consultations.

Throughout this paper, $p$ denotes the prime for which $q = p^s$ for some $s \geq 1$. We write $C_{q-1}$ for the cyclic group of nonzero elements of $\mathbb{F}_q$ under multiplication. By $d$ we denote $\gcd(2, p-1)$, which is the order of the center of $\mathrm{SL}(2, q)$. The order of $\mathrm{PSL}(2, q)$ is $q(q^2 - 1)/d$.

## 1. Nielsen equivalence, the Nielsen invariant, and the trace invariants

A *generating vector* for a group is a tuple $S = (s_1, \ldots, s_n)$ such that $\{s_1, \ldots, s_n\}$ is a generating set. A generating vector $T = (t_1, \ldots, t_n)$ is said to be obtained from $S$ by a *Nielsen move* if there is a $j$ so that $s_i = t_i$ for all $i \neq j$, and for some $k \neq j$, $t_j$ equals $s_j s_k$. Also, the replacement of an entry by its inverse is a Nielsen move. Sequences of these two basic moves allow one to replace $s_j$ by $s_j s_k^{-1}$ or by $s_k^{\pm 1} s_j$, and consequently allow the interchange of any two elements, as in the sequence $(a, b, c) \to (a, ba, c) \to (a^{-1}b^{-1}a, ba, c) \to (a^{-1}b^{-1}a, a, c) \to (b^{-1}, a, c) \to (b, a, c)$. The generating vectors $S$ and $T$ are called *(Nielsen) equivalent* if there is a sequence of Nielsen moves that changes one to the other. They are called *weakly (Nielsen) equivalent* if there is an automorphism $\alpha$ of $G$ such that $(\alpha(s_1), \ldots, \alpha(s_n))$ is equivalent to $T$. Note that equivalent or weakly equivalent generating vectors must have the same number of elements.

For a group $G$, a normal subgroup $S$ of the free group $W_n$ of rank $n$ is called a *$G$-defining subgroup* if $W_n/S \cong G$. It was noted by B. H. Neumann

and H. Neumann [30] that the number of weak equivalence classes of generating $n$-vectors of $G$ equals the number of $G$-defining subgroups of $W_n$, up to automorphism of $W_n$ (see the discussion in [11, p. 542]).

We will now define a slight generalization of the usual Nielsen invariant. Let $G$ be a group, and $\widetilde{G} \to G$ a quotient map whose kernel is central in $\widetilde{G}$. For $g \in G$, denote by $\widetilde{g}$ some element of $\widetilde{G}$ that maps to $G$. For $g, h \in G$, the centrality of the kernel shows that the commutator $[\widetilde{g}, \widetilde{h}]$ is independent of the choices of $\widetilde{g}$ and $\widetilde{h}$.

**Lemma 1.1.** *If $(s_1, s_2)$ and $(t_1, t_2)$ are Nielsen equivalent generating 2-vectors of $G$, then $[\widetilde{s_1}, \widetilde{s_2}]$ is conjugate either to $[\widetilde{t_1}, \widetilde{t_2}]$ or to $[\widetilde{t_2}, \widetilde{t_1}]$.*

*Proof.* It suffices to check the result for the two basic Nielsen moves. We have, for example, $[\widetilde{s_1 s_2}, \widetilde{s_2}] = [\widetilde{s_1}\widetilde{s_2}, \widetilde{s_2}] = [\widetilde{s_1}, \widetilde{s_2}]$, and $[\widetilde{s_1^{-1}}, \widetilde{s_2}] = [\widetilde{s_1}^{-1}, \widetilde{s_2}] = \widetilde{s_1}^{-1}[\widetilde{s_2}, \widetilde{s_1}]\widetilde{s_1}$. The other cases are similar. □

Define the *Nielsen invariant* with *values in* $\widetilde{G}$ of a generating 2-vector $(t_1, t_2)$ to be the set of (possibly equal) conjugacy classes of $[\widetilde{t_1}, \widetilde{t_2}]$ and $[\widetilde{t_2}, \widetilde{t_1}]$. Lemma 1.1 shows that this is an invariant of Nielsen equivalence.

The particular case we will use is $G = \mathrm{PSL}(2, q)$ and $\widetilde{G} = \mathrm{SL}(2, q)$. We abuse notation by writing elements of $\mathrm{PSL}(2, q)$ as matrices, in which case the Nielsen invariant of a generating pair $\{A, B\}$ of $\mathrm{PSL}(2, q)$ is the pair of conjugacy classes in $\mathrm{SL}(2, q)$ of $[A, B]$ and $[B, A]$. Since these matrices and any of their conjugates have the same trace, we may define the *trace invariant* of a generating pair $\{A, B\}$ of $\mathrm{PSL}(2, q)$ to be the trace of $[A, B]$. Since the trace invariant is determined by the Nielsen invariant, it is an invariant of the Nielsen equivalence class of a generating pair of $\mathrm{PSL}(2, q)$.

It turns out that the trace invariant for $\mathrm{PSL}(2, q)$ is very nearly a complete invariant of the Nielsen invariant. The relation between them is examined in section 6 below.

In the remainder of this section, we develop some basic properties of trace invariants. In the proofs, we use some of the well-known properties of $\mathrm{PSL}(2, q)$ that are given at the beginning of section 2 below.

There is one element of $\mathbb{F}_q$ that can never be a trace invariant.

**Proposition 1.2.** *If $A$ and $B$ generate $\mathrm{PSL}(2, q)$, then the trace of $[A, B]$ is not equal to 2.*

*Proof.* Suppose that $\mathrm{tr}([A, B]) = 2$. Then $[A, B]$ is conjugate in $\mathrm{PSL}(2, q)$ to some matrix $T_\lambda = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$. If $\lambda = 0$, then $A$ and $B$ commute and cannot generate. Otherwise, assume by conjugation of $A$ and $B$ that $[A, B] = T_\lambda$, and write $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. We have $ABA^{-1} = T_\lambda B = \begin{pmatrix} a + c\lambda & b + d\lambda \\ c & d \end{pmatrix}$, and equating the traces shows that $c = 0$. Since $[B, A] = T_{-\lambda}$, $A$ is also upper triangular, so that the subgroup generated by $A$ and $B$ is contained in the subgroup of upper triangular matrices of $\mathrm{PSL}(2, q)$. □

The following proposition gives some trace invariants when $q$ is prime.

**Proposition 1.3.** *If $x \in C_{p-1}$, then the matrices $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}$ generate $\mathrm{PSL}(2,p)$, and their commutator has trace $2+x^2$.*

*Proof.* By a theorem of Dickson (see 3(6.21) of [39]), these matrices generate $\mathrm{PSL}(2,p)$. The trace calculation is straightforward. □

An element $A \in \mathrm{PSL}(2,q)$ is called *parabolic* when $\mathrm{tr}(A)$ is $\pm 2$. For any $q$, generating pairs containing a parabolic matrix can never give traces other than those in proposition 1.3.

**Proposition 1.4.** *If $A, B$ generate $\mathrm{PSL}(2,q)$, and $A$ is parabolic, then the trace of $[A,B]$ is of the form $2+x^2$ for some $x \in C_{q-1}$.*

*Proof.* We may assume that the trace of $A$ is 2, and after conjugation in $\mathrm{PSL}(2,q)$ that $A = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$. Writing $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ gives $[A,B] = \begin{pmatrix} 1+ac\lambda+c^2\lambda^2 & (1-a^2)\lambda - ac\lambda^2 \\ c^2\lambda & 1-ac\lambda \end{pmatrix}$, so the trace of $[A,B]$ is $2+(c\lambda)^2$. If $c\lambda = 0$, then $[A,B]$ has trace 2, and proposition 1.2 shows that $A$ and $B$ could not generate. □

## 2. The case of $q > 11$

In this section, we will find generating pairs that realize many elements of $\mathbb{F}_q$ as values of the trace invariant, enough elements to prove the Main Theorem for all $q > 11$. We first list some well-known facts about $\mathrm{PSL}(2,q)$.

From [12, Proposition 2.1], we have

**Proposition 2.1.** *Let $A \in \mathrm{SL}(2,q)$ have trace $\alpha$ and eigenvalues $\lambda_1$ and $\lambda_2$.*

1) *If $\lambda_1 = \lambda_2$, then $A$ is conjugate to $\begin{pmatrix} \lambda_1 & \mu \\ 0 & \lambda_1 \end{pmatrix}$.*
2) *If $\lambda_1 \neq \lambda_2$, then $A$ is conjugate to $\begin{pmatrix} \alpha & 1 \\ -1 & 0 \end{pmatrix}$.*

The eigenvalues of $A$ are the roots of $\lambda^2 - \mathrm{tr}(A)\lambda + 1 = 0$, so are equal only when they are both 1 or both $-1$. In particular, all matrices in $\mathrm{SL}(2,q)$ of the same trace are conjugate in $\mathrm{SL}(2,q)$ unless the trace is $\pm 2$. Part 1 shows that all elements of $\mathrm{SL}(2,q)$ of trace $\pm 2$ have order $p$ or $2p$, and have order $p$ in $\mathrm{PSL}(2,q)$. Thus the order of any element in $\mathrm{PSL}(2,q)$ is determined by its trace.

Write $T$ for the subgroup of $\mathrm{PSL}(2,q)$ consisting of all matrices of the form $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$. Since $T$ is isomorphic to the additive group of $\mathbb{F}_q$, it is an elementary abelian $p$-group of order $q$, and is a $p$-Sylow subgroup of $\mathrm{PSL}(2,q)$.

The subgroups of $\mathrm{PSL}(2,q)$ were determined by L. E. Dickson [5]. The following statement is from Theorem 3(6.25) of Suzuki [39].

**Theorem 2.2.** *Every subgroup of* $\mathrm{PSL}(2,q)$ *is isomorphic to (at least) one of the following.*

(a) *The dihedral groups of orders* $2(q\pm1)/d$ *and their subgroups.*
(b) *A group* $H$ *of order* $q(q-1)/d$ *and its subgroups. A Sylow* $p$*-subgroup* $Q$ *of* $H$ *is elementary abelian, normal in* $H$*, and the factor group* $H/Q$ *is a cyclic group of order* $(q-1)/d$*.*
(c) $A_4$*,* $S_4$*, or* $A_5$*.*
(d) $\mathrm{PSL}(2,p^r)$ *or* $\mathrm{PGL}(2,p^r)$ *where* $r$ *divides* $s$*. The latter subgroup can occur only when* $p>2$*.*

The last statement in (d) is from 3(6.18) of [39]. Note that the subgroups $Q$ in (b) are $p$-Sylow subgroups of $\mathrm{PSL}(2,q)$, so are conjugate to $T$.

We use the following terminology to refer to the subgroups described in theorem 2.2: subgroups as in (a) are called *small,* as in (b) are called *affine,* as in (c) are called *exceptional,* and as in (d) are called *projective.*

From (2.3) and (2.4) of [12], we have the following information.

**Lemma 2.3.** *The orders of nonparabolic elements of* $\mathrm{PSL}(2,q)$ *are exactly the divisors of* $(q+1)/d$ *and* $(q-1)/d$*. In particular, the maximum order of a nonparabolic element of* $\mathrm{PSL}(2,q)$ *is* $(q+1)/d$*.*

For $x,y\in\mathbb{F}_q$ with $x\neq0$, put $H_x=\begin{pmatrix}x&0\\0&x^{-1}\end{pmatrix}$ and $J_y=\begin{pmatrix}y+1&1\\y&1\end{pmatrix}$. The next lemma is straightforward.

**Lemma 2.4.** *Put* $D=x-x^{-1}$*. Then* $[H_x,J_y]=\begin{pmatrix}1-Dxy&Dx(y+1)\\-Dx^{-1}y&1+Dx^{-1}y\end{pmatrix}$*. Consequently, the trace of* $[H_x,J_y]$ *is* $2-D^2y$*.*

The next lemma will ensure that $H_x$ and $J_y$ do not generate a small or affine subgroup.

**Lemma 2.5.** *Assume that* $y\neq0$*, and that* $x^4\neq1$ *and* $x^6\neq1$*. Then* $[H_x,J_y]$ *and* $[H_x^{-1},J_y]$ *do not commute in* $\mathrm{PSL}(2,q)$*.*

*Proof.* Again write $D=x-x^{-1}$, which is nonzero since $x^2\neq1$. Now $[H_x^{-1},J_y]=[H_{x^{-1}},J_y]$, so $[H_x^{-1},J_y]$ is obtained from the expression in lemma 2.4 by replacing each appearance of $x$ with $x^{-1}$ (hence each $D$ with $-D$). One then calculates

$$[H_x,J_y]\,[H_x^{-1},J_y]=\begin{pmatrix}1+D^3xy(y+1)&D^2(y+1)-D^3xy(y+1)\\D^2y+D^3x^{-1}y^2&1-D^3x^{-1}y(y+1)\end{pmatrix}.$$

Again, by replacing each $x$ by $x^{-1}$, we obtain

$$[H_x^{-1},J_y]\,[H_x,J_y]=\begin{pmatrix}1-D^3x^{-1}y(y+1)&D^2(y+1)+D^3x^{-1}y(y+1)\\D^2y-D^3xy^2&1+D^3xy(y+1)\end{pmatrix}.$$

If these matrices are equal, their $(2,1)$ entries show that $x = -x^{-1}$, in contradiction to the assumption that $x^4 \neq 1$. So assume that $p \neq 2$ and the matrices differ by multiplication by $-I$. From the $(2,1)$ entries, we have $y - Dxy^2 = -y - Dx^{-1}y^2$, or $D^2y = 2$. From the $(1,1)$ entries, we find that $1 - D^3x^{-1}y(y+1) = -1 - D^3xy(y+1)$, which implies that $D^4y(y+1) = -2$, and using $D^2y = 2$ this leads to $D^2 = -3$. But the equation $D^2 = -3$ says that $x^2 - 2 + x^{-2} = -3$, that is, $x^4 + x^2 + 1 = 0$. Multiplying by $x^2 - 1$ shows that $x^6 = 1$, in contradiction to the hypothesis. □

**Proposition 2.6.** *Assume that $q > 11$. Suppose that $x$ generates $C_{q-1}$ and that $y \neq 0$. Then $H_x$ and $J_y$ generate* $\mathrm{PSL}(2, q)$.

*Proof.* Since $q > 7$, we have $x^4 \neq 1$ and $x^6 \neq 1$. Let $S$ be the subgroup of $\mathrm{PSL}(2, q)$ generated by $\{H_x, J_y\}$. Note that the order of $H_x$ is $(q-1)/d$. We assume that $S \neq \mathrm{PSL}(2, q)$, and consider the four possibilities given in theorem 2.2. Lemma 2.5 shows that $S$ is not small or affine. Since $(q-1)/2$ is at least 6, $H_x$ has order more than 5, so $S$ cannot be exceptional.

Assume that $S$ is projective, and consider first the case that $S$ is isomorphic to $\mathrm{PSL}(2, p^r)$, where $r$ is a proper divisor of $s$. By lemma 2.3 the order of $H_x$ is no more than $(p^r + 1)/d$. Since $r < s$, this is less than $(p^s - 1)/d$, the known order of $H_x$.

The remaining possibility is that $p > 2$ and $S$ is isomorphic to $\mathrm{PGL}(2, p^r)$. Since $H_x^2$ is contained in a subgroup isomorphic to $\mathrm{PSL}(2, p^r)$, lemma 2.3 shows that $(p^s - 1)/2$, the order of $H_x$, is no more than $p^r + 1$. This can hold only when $p = 3$ and $s = 2$, that is, $q = 9$. □

Proposition 2.6 implies the Main Theorem in the case $q > 11$. For let $x$ be a generator of $C_{q-1}$, and put $D = x - x^{-1}$. By proposition 2.6 and lemma 2.4, all traces of the form $2 - D^2y$ with $y \neq 0$ arise as trace invariants of generating pairs for $\mathrm{PSL}(2, q)$. As we will discuss below, the proof of proposition 2.6 also applies with only slight modifications to the case of $q = 8$.

## 3. The cases of $q < 11$ but $q \neq 7, 11$

The cases when $q$ is small could be checked by finite calculation, but we prefer to use algebraic arguments that we find more illuminating.

For $q = 2$, the Main Theorem is immediate from proposition 1.2.

For $q = 3$, $\mathrm{PSL}(2, 3)$ is $A_4$. Since the commutator subgroup consists of involutions, the trace invariant of any generating pair is 0. In fact, there is only one Nielsen equivalence class of generating 2-vector. We do not need this fact to prove the Main Theorem, but it will be relevant in section 6. So we will now verify it, by checking that any generating 2-vector $(x, y)$ is Nielsen equivalent to $((1\ 3\ 4), (2\ 3\ 4))$.

We may assume that $x$ has order 3, and replacing $y$ by $xy$ if necessary that $y$ also has order 3. Write $x = (a\ b\ c)$ and $y = (d\ e\ f)$. Replacing $x$ by $x^{-1}$, if necessary, we can and always will assume that $a < b < c$, and

similarly that $d < e < f$. If neither $x$ or $y$ contains both 1 and 2, then they must have the desired forms (1 3 4) and (2 3 4). So we may assume that $x = (1\ 2\ c)$.

Suppose first that $c = 3$. Necessarily, $y$ contains 4, so conjugating $y$ by a power of $x$, we may assume that $y = (2\ 3\ 4)$. Then, conjugating $x$ by $y$ makes $x = (1\ 3\ 4)$. The case when $c = 4$ is similar.

**Proposition 3.1.** *All elements of* $\mathbb{F}_4$ *other than* 0 *occur as trace invariants.*

*Proof.* For $q = 4$, $\mathrm{PSL}(2,4) \cong A_5$. Proposition 1.2 show that 0 cannot occur as a trace invariant. Write $\mathbb{F}_4$ as $\{0, 1, x, x+1\}$ where $x^2 = x + 1$. Put $A = \begin{pmatrix} x+1 & 0 \\ 0 & x \end{pmatrix}$, a matrix of order 3, and for $\alpha \in \mathbb{F}_4$ and $\beta \in \{x, x+1\}$ put $B_{\alpha,\beta} = \begin{pmatrix} \beta+\alpha & \beta \\ \beta^{-1}(\alpha^2+\alpha\beta+1) & \alpha \end{pmatrix}$. Since the trace of $B_{\alpha,\beta}$ is $\beta$, it has order 5, so $A$ and $B_{\alpha,\beta}$ generate $\mathrm{PSL}(2,4)$. One computes that $\mathrm{tr}([A,B]) = \alpha^2 + \alpha\beta + 1$, and taking $(\alpha,\beta)$ to be $(0,x)$, $(1,x)$, and $(1,x+1)$ gives all nonzero elements of $\mathbb{F}_4$. □

**Proposition 3.2.** *The elements of* $\mathbb{F}_5$ *that occur as trace invariants are* 1 *and* 3*.*

*Proof.* The orders of elements of $\mathrm{PSL}(2,5)$ are 2 for elements of trace 0, 3 for elements of trace $\pm 1$, and 5 for nonidentity elements of trace $\pm 2$.

Let $A, B \in \mathrm{PSL}(2,5)$. Suppose for contradiction that none of $A$, $B$, or $AB$ has trace $\pm 2$, so that $A^r = B^s = (AB)^t$ with $r, s, t \in \{2, 3\}$. We must have $r = s = t = 3$, since if one of them is 2, then $A$ and $B$ generate either a dihedral group or $A_4$. Regard $\mathrm{PSL}(2,5)$ as $A_5$. We will show that, after possibly changing $\{A, B\}$ by Nielsen equivalence, $AB$ has order 5. By inner automorphism, we may assume that $A = (123)$. Since $A$ and $B$ generate, $B$ must move 4 and 5, so (possibly replacing $B$ by $B^{-1}$) we may write $B = (a45)$. Conjugating by $A$ some number of times makes $B = (345)$, but then $AB = (12453)$ has order 5 and trace $\pm 2$. We conclude that for a generating pair, at least one of $A$, $B$, or $A$ has trace $\pm 2$. Proposition 1.4 then shows that the trace invariant must be 1 or 3. By proposition 1.3, both of these are achieved. □

**Proposition 3.3.** *All elements of* $\mathbb{F}_8$ *other than* 0 *occur as trace invariants.*

*Proof.* The proof of proposition 2.6 applies when $q = 8$. For the order of the element $H_x$ is 7, so $S$ is not exceptional, nor can it be one of the projective subgroups $\mathrm{PSL}(2,2)$, $\mathrm{PGL}(2,2)$, $\mathrm{PSL}(2,4)$, or $\mathrm{PGL}(2,4)$. Proposition 3.3 then follows just as the Main Theorem follows from proposition 2.6. □

**Proposition 3.4.** *If* $q = 9$*, then all elements except* 1 *and* 2 *occur as trace invariants.*

*Proof.* We write $\mathbb{F}_9$ as $\{a + bx \mid a, b \in \mathbb{F}_3,\ x^2 = x + 1\}$. The Frobenius automorphism exchanges $x$ and $2x + 1$. Matrices of trace 2 have order 3, of

traces $x+1$ and $2x+2$ have order 4, and of traces $x$, $2x+1$, $x+2$, and $2x$ have order 5.

We first use the matrices defined in section 2. In particular, the matrix $H_x = \begin{pmatrix} x & 0 \\ 0 & x+2 \end{pmatrix}$ has trace $2x+2$ so has order 4. By lemma 2.4, the trace of $[H_x, J_y]$ is $2-y$. Theorem 2.2 shows quickly that any element of order 4 together with any element of order 5 forms a generating pair. In particular, $H_x$ together with each of the matrices $J_{x+1}$ and $J_{2x+1}$ generates, giving trace invariants $2x+1$ and $x+1$ respectively.

The matrix $J_{x+2}$ has trace $x+1$ so has order 4, and the trace of $[H_x, J_{x+2}]$ is $2x$ so $[H_x, J_{x+2}]$ has order 5. Theorem 2.2 then shows that $H_x$ and $J_{x+2}$ are a generating pair, giving $2x$ as a trace invariant.

The elements $2x+1$, $x+1$, and $2x$ lie in three different orbits of $\mathbb{F}_9 - \mathbb{F}_3$ under the Frobenius automorphism. Applying the Frobenius automorphism to the entries of each of these three pairs of generators yields a pair whose trace invariant is the other element of each orbit. Thus we have all elements of $\mathbb{F}_9 - \mathbb{F}_3$ as trace invariants.

The elements $\begin{pmatrix} x+1 & 1 \\ -1 & 0 \end{pmatrix}$ and $\begin{pmatrix} x & 1 \\ -1 & 0 \end{pmatrix}$ have orders 4 and 5 respectively, so they generate, giving 0 as a trace invariant.

In principle, arguments as in propositions 4.5 and 4.6 below ought to show these are the only possible trace invariants, but we have an easy way out. D. Stork [38] showed that there were exactly four weak equivalence classes of PSL$(2,9)$-generating pairs (in [38] this group is regarded as $A_6$). The elements we have so far obtained form four orbits of the Frobenius automorphism of $\mathbb{F}_9$, so a realization of any other element would imply the existence of a fifth weak equivalence class of generating pair. □

## 4. The cases of $q = 7$ and $q = 11$

We will adapt some of the ideas developed by Macbeath in [25]. Write $G_0$ for SL$(2,q)$, $G$ for PSL$(2,q)$, and $k$ for $\mathbb{F}_q$. A *$G_0$-pair* is a pair $(A,B)$ of elements of $G_0$. A *$k$-triple* is an element of $k^3$. Associated to any $G_0$-pair $(A,B)$ is the $k$-triple $(\mathrm{tr}(A), \mathrm{tr}(B), \mathrm{tr}(AB))$, which we denote by $T(A,B)$. Theorem 1 of [25] shows the following.

**Proposition 4.1.** *Every $k$-triple is of the form $T(A,B)$ for some $G_0$-pair.*

The expression $\alpha^2+\beta^2+\gamma^2-\alpha\beta\gamma-4$ plays a key role in [25]. We will denote it by $Q(\alpha,\beta,\gamma)$. The relation between $k$-triples and trace invariants is based on the following calculation.

**Proposition 4.2.** *If $T(A,B) = (\alpha,\beta,\gamma)$, then* $\mathrm{tr}([A,B]) = 2 + Q(\alpha,\beta,\gamma)$.

*Proof.* First suppose that $A = \begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix}$ and $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, so that $AB = \begin{pmatrix} ax & bx \\ cx^{-1} & dx^{-1} \end{pmatrix}$. Solving the equations $a+d=\beta$ and $ax+dx^{-1}=\gamma$, we

obtain $a = (\gamma - \beta x^{-1})/D$ and $d = (\beta x - \gamma)/D$, where $D = x - x^{-1}$. So $ad - 1 = -Q(\alpha, \beta, \gamma)/D^2$. Since $ad - bc = 1$, we have $-bc = Q(\alpha, \beta, \gamma)/D^2$, and by further calculation,

$$\operatorname{tr}([A, B]) = 2ad - bc(x^2 + x^{-2}) = 2 + Q(\alpha, \beta, \gamma) \ .$$

Now suppose that $A = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$. Then we find that $\gamma = \beta + \lambda c$ and that

$$\operatorname{tr}([A, B]) = 2 + (\lambda c)^2 = 2 + Q(2, \beta, \gamma) \ .$$

□

We will use the following Nielsen moves:

1) $\widetilde{a}(A, B) = (A^{-1}, B)$,
2) $\widetilde{b}(A, B) = (B, A)$,
3) $\widetilde{c}(A, B) = (A^{-1}, AB)$.

Note that $\widetilde{a}$, $\widetilde{b}$, and $\widetilde{c}$ generate all Nielsen moves, since $\widetilde{a}\,\widetilde{c}\,(A, B) = (A, AB)$. By $a$, $b$, and $c$ we denote the corresponding permutations induced on the set of $k$-triples via $T$:

1) $a(\alpha, \beta, \gamma) = (\alpha, \beta, \alpha\beta - \gamma)$;
2) $b(\alpha, \beta, \gamma) = (\beta, \alpha, \gamma)$;
3) $c(\alpha, \beta, \gamma) = (\alpha, \gamma, \beta)$;

That $\widetilde{b}$ and $\widetilde{c}$ are well-defined is immediate, and for $\widetilde{a}$ it is simply the identity that $\operatorname{tr}(A^{-1}B) = \operatorname{tr}(A)\operatorname{tr}(B) - \operatorname{tr}(AB)$.

Since $\widetilde{a}\,\widetilde{c}\,(A, B) = (A, AB)$, the third entry of $(ac)^{k-1}(\alpha, \beta, \gamma)$ expresses the trace of $A^k B$ in terms of the traces of $A$, $B$, and $AB$. This will be useful in the proof of proposition 4.6 below, so we state it as a lemma.

**Lemma 4.3.** *The third entry of $(ac)^{k-1}T(A, B)$ is the trace of $A^k B$.*

We say that the $k$-triples $(\alpha, \beta, \gamma)$ and $(\alpha', \beta', \gamma')$ are equivalent, and write $(\alpha, \beta, \gamma) \sim (\alpha', \beta', \gamma')$, if they lie in the same orbit of $k^3$ under the action of the permutations $a$, $b$, and $c$. In particular, permuting the entries of a $k$-triple gives an equivalent triple. Since $\widetilde{a}$, $\widetilde{b}$, and $\widetilde{c}$ generate all Nielsen moves, the following proposition is immediate.

**Proposition 4.4.** *Let $(\alpha, \beta, \gamma)$ and $(\alpha', \beta', \gamma')$ be $k$-triples, and suppose that $T(A, B) = (\alpha, \beta, \gamma)$. Then $(\alpha, \beta, \gamma) \sim (\alpha', \beta', \gamma')$ if and only if there is a pair $(A', B')$, Nielsen equivalent to $(A, B)$, with $T(A', B') = (\alpha', \beta', \gamma')$.*

We remind the reader of the following presentations for the dihedral and exceptional subgroups of $\mathrm{PSL}(2, q)$ (see for example [4]): $D_n = \langle A, B \mid A^2 = B^2 = (AB)^n = 1\rangle$, $A_4 = \langle A, B \mid A^2 = B^3 = (AB)^3 = 1\rangle$, $S_4 = \langle A, B \mid A^2 = B^3 = (AB)^4 = 1\rangle$, $A_5 = \langle A, B \mid A^2 = B^3 = (AB)^5 = 1\rangle$. Manipulation of these presentations shows that for any of them, permuting the exponents of the three relations yields a presentation of an isomorphic group.

**Proposition 4.5.** *The elements of* $\mathbb{F}_7$ *that occur as trace invariants are the elements other than* 0, 1, *and* 2.

There is an easy way to deduce proposition 4.5. One can check quickly, as we do in the first paragraph of the proof below, that the elements of $\mathbb{F}_7$ other than 0, 1, and 2 occur as trace invariants. D. Stork [38] proved that there are exactly four weak equivalence classes of generating pairs for PSL$(2, 7)$, so there can be at most four trace invariants (the Frobenius automorphism acts trivially, so weak trace invariants and trace invariants are the same). But we give here an independent proof, since it is short and provides a template for the case of PSL$(2, 11)$ in proposition 4.6.

*Proof.* In any PSL$(2, q)$, matrices of trace 0 have order 2 and matrices of trace $\pm 1$ have order 3. In PSL$(2, 7)$, matrices of trace $\pm 3$ have order 4. Since the squares of nonzero elements of $\mathbb{F}_7$ are 1, 2, and 4, proposition 1.3 shows that 3, 4, and 6 occur as trace invariants. By proposition 4.1, there is a pair $(A, B)$ for which $T(A, B) = (1, 1, 4)$. Proposition 4.2 shows that $\operatorname{tr}([A, B]) = 5$. Consequently, $[A, B]$ has order 7 and the subgroup of PSL$(2, 7)$ generated by $A$ and $B$ is not exceptional. Since $AB$ has order 4, the subgroup is also not affine or small, so $A$ and $B$ generate. Therefore all elements of $\mathbb{F}_7$ other than 0, 1, and 2 do appear as trace invariants.

For elements $A$ and $B$ in SL$(2, q)$, we denote by $S(A, B)$ the subgroup of PSL$(2, q)$ that they generate. Proposition 1.2 shows that if $\operatorname{tr}([A, B]) = 2$ then $S(A, B) \neq \mathrm{PSL}(2, 7)$.

To complete the proof, we consider a pair $(A, B)$ with $T(A, B) = (\alpha, \beta, \gamma)$, and must show that either $S(A, B) \neq \mathrm{PSL}(2, q)$ or $\operatorname{tr}([A, B])$ is not 0 or 1. Proposition 1.4 shows that $\operatorname{tr}([A, B])$ is not 0 or 1 whenever an entry of $T(A, B)$ is $\pm 2$, so from now on we assume that this is not the case.

Changing $(A, B)$ by Nielsen equivalence (or replacing $A$ or $B$ by its product with $-I$) changes neither $S(A, B)$ nor $\operatorname{tr}([A, B])$. Consequently, by proposition 4.4, we are free to change $T(A, B)$ by equivalence. In particular, by permuting the entries of $(\alpha, \beta, \gamma)$, we may always assume that $\alpha \leq \beta \leq \gamma$ (regarding these as integers between 0 and 6). Since $T(-A, B) = (-\alpha, \beta, -\gamma)$ and $T(A, -B) = (\alpha, -\beta, -\gamma)$, we may further assume that $\alpha \leq \beta \leq 3$.

If $(\alpha, \beta) = (0, 0)$, then $A$ and $B$ have order 2 so $S(A, B)$ is dihedral. If $(\alpha, \beta) = (0, 1)$, then $A$ has order 2 and $B$ has order 3. Since $\gamma$ is either 1 or $\pm 3$, $AB$ has order 3 or 4 and hence $S(A, B)$ is $A_4$ or $S_4$.

Suppose that $(\alpha, \beta) = (1, 1)$. Proposition 4.2 shows that $\operatorname{tr}([A, B]) = \gamma^2 - \gamma$. Since $\gamma$ is $\pm 1$ or $\pm 3$, we obtain $\operatorname{tr}([A, B])$ equal to 0 or 1 only when $\gamma = 1$. But $(1, 1, 1) \sim (0, 1, 1)$, so $S(A, B) \cong A_4$.

Suppose that $(\alpha, \beta) = (1, 3)$. Then $\operatorname{tr}([A, B]) = \gamma^2 - 3\gamma + 1$, and since $\gamma = \pm 3$ this is 0 or 1 only when $\gamma = 3$. But $(1, 3, 3) \sim (0, 1, 3)$, so $S(A, B) \cong S_4$.

Finally, suppose that $(\alpha, \beta) = (3, 3)$. Then $\operatorname{tr}([A, B]) = \gamma^2 - 2\gamma + 2$, and since $\gamma = \pm 3$ this is not 0 or 1. □

The next proposition will complete the proof of the Main Theorem.

**Proposition 4.6.** *The elements of* $\mathbb{F}_{11}$ *that occur as trace invariants are the elements other than* 1 *and* 2.

*Proof.* Our approach is analogous to the case of PSL(2, 7) in proposition 4.5. In PSL(2, 11), matrices of traces $\pm 3$ and $\pm 4$ have order 5, and of trace $\pm 5$ have order 6.

Again, the generating pairs containing parabolics give all $2 + x^2$ as trace invariants, which are the values 0, 3, 5, 6, and 7. Theorem 2.2 shows quickly that any element of order 6 together with any element of order 5 give a generating pair. Thus, for example, the cases $T(A, B, AB) = (3, 1, 6)$, $(4, 6, 6)$, $(3, 3, 6)$, and $(3, 0, 5)$ give respectively the values 4, 8, 9, and 10 as trace invariants.

We use the notation of the proof of proposition 4.5, the only differences at the outset being that $\alpha \leq \beta \leq 5$, and that the only problematical case is when $\mathrm{tr}([A, B]) = 1$.

If $(\alpha, \beta) = (0, 0)$, then $A$ and $B$ have order 2 and $S(A, B)$ is dihedral. If $(\alpha, \beta) = (0, 1)$, then $A$ has order 2, $B$ has order 3, and $S(A, B)$ is exceptional unless $\gamma = \pm 5$. But in the latter case, $\mathrm{tr}([A, B]) = 2$.

For $(\alpha, \beta) = (1, 1)$, $\mathrm{tr}([A, B]) = \gamma^2 - \gamma$, which is 1 only when $\gamma = 4$ or $\gamma = -3$. Suppose that $\gamma = 4$. Changing $(A, B)$ by Nielsen equivalence, we may assume that $T(A, B, AB) = (4, 1, 1)$. Since $(ac)^2(4, 1, 1) = (4, 3, 0)$, lemma 4.3 shows that $\mathrm{tr}(A^3B) = 0$. So we have $(A^3)^5 = B^3 = (A^3B)^2 = 0$, showing that $S(A, B) = A_5$. Suppose that $\gamma = -3$. Since $a(1, 1, -3) = (1, 1, 4)$, we again have $S(A, B) = A_5$.

Suppose now that $(\alpha, \beta) = (1, 3)$. Then $\mathrm{tr}([A, B]) = \gamma^2 - 3\gamma - 3$, which equals 1 only when $\gamma = 4$ and $\gamma = -1$. Since $(1, 3, 4) \sim (1, 1, 4)$ and $(1, 3, -1) \sim (1, -1, 3)$, with the latter changed to $(1, 1, -3)$ upon replacing $B$ by $-B$, these both give $S(A, B) = A_5$.

The cases when $(\alpha, \beta)$ is $(3, 3)$, $(3, 4)$, $(3, 5)$, $(4, 4)$, $(4, 5)$, or $(5, 5)$ never produce $\mathrm{tr}([A, B]) = 1$. □

## 5. The orbits of the Frobenius automorphism

There are several ways to obtain the formula

$$\Psi_{p^s} = \frac{1}{s} \sum_{r|s} \varphi(s/r)\, p^r$$

stated in the introduction. The best we have seen is the following one, shown to us by Gareth Jones. It is based on the Burnside-Cauchy-Frobenius formula, commonly called Burnside's Lemma (see [31] for its history). Burnside's Lemma says that the number of orbits of a finite group acting on a finite set equals the average number of fixed points of the elements of $G$. Short proofs are readily available in the literature (for example, [1]).

**Lemma 5.1** (Burnside's Lemma)**.** *If a finite group $G$ acts on a finite set $\Omega$, then the number of orbits is given by*

$$\frac{1}{|G|}\sum_{g\in G}\pi(g)$$

*where $\pi(g)$ is the number of points fixed by $G$.*

In the remainder of this section, we use $\mathbb{F}(r)$ to denote $\mathbb{F}_{p^r}$. To obtain the formula for $\Psi_q$, we will apply Burnside's Lemma with $\Omega = \mathbb{F}(s)$ and $G = \mathrm{Aut}(\mathbb{F}(s))$. We will use the following elementary facts about finite fields:

1) $\mathbb{F}(r)$ occurs as a subfield of $\mathbb{F}(s)$ if and only if $r|s$.
2) When $\mathbb{F}(r)$ occurs as a subfield of $\mathbb{F}(s)$, it is the unique subfield of this order.
3) $\mathrm{Aut}(\mathbb{F}(s))$ is cyclic of order $s$, generated by the Frobenius automorphism $\sigma$ that sends each $x$ to $x^p$. The elements fixed by $\sigma^m$ are exactly the subfield $\mathbb{F}(r)$, where $r$ is the greatest common divisor $\gcd(m, s)$.

The last fact implies that if $s/r = \gcd(m, s)$, then $\pi(\sigma^m) = p^{s/r}$. Since each element in cyclic group of order $s$ has order $r$ dividing $s$, and there are exactly $\varphi(r)$ such elements, Burnside's Lemma gives the number of orbits to be

$$\frac{1}{s}\sum_{r|s}\varphi(r)\,p^{s/r} = \frac{1}{s}\sum_{r|s}\varphi(s/r)\,p^{r}\ .$$

We remark that this proof extends in a straightforward manner to show that for any prime power $q$, the number of orbits of the action of the Galois group $\mathrm{Aut}_{\mathbb{F}_q}\mathbb{F}_{q^s}$ on $\mathbb{F}_{q^s}$ is $\dfrac{1}{s}\displaystyle\sum_{r|s}\varphi(s/r)\,q^r$.

The crude lower bound of $\frac{q}{s}$ for $\Psi_q$ is accurate because the vast majority of elements of $\mathbb{F}_q$ do not lie in a proper subfield, so practically all orbits have $s$ elements. For example, we compute that $\Psi_{3^{12}} = 44,368$, while $3^{12}/12$ gives $44,287$, or approximately 99.8% of the exact value, and $\Psi_{2^{30}} = 35,792,568$, with the crude lower bound of $35,791,395$ approximately 99.9967% of the exact value.

## 6. Comparison of the Nielsen invariant and trace invariant

In this section, we will see that the trace invariant coincides with the Nielsen invariant (with values in $\mathrm{SL}(2, q)$) of $\mathrm{PSL}(2, q)$, except when the trace invariant is $-2$ and $q \equiv 1 \mod 4$. In this case, there are at most two Nielsen invariants. For some values of $q$, both Nielsen invariants occur, giving examples of Nielsen inequivalent generating pairs with the same trace invariant.

Consider an element $A$ of $\mathrm{SL}(2, q)$ with trace $2\epsilon$ where $\epsilon = \pm 1$. Define $M(A)$ to be the set of $\mu$ for which $A$ is conjugate to $\begin{pmatrix}\epsilon & \mu\\ 0 & \epsilon\end{pmatrix}$. Of course,

$M(A)$ is a complete invariant of the conjugacy class of $A$. Conjugation by an element $P$ of $\mathrm{SL}(2,q)$ takes $\begin{pmatrix}\epsilon & \mu\\ 0 & \epsilon\end{pmatrix}$ to $\begin{pmatrix}\epsilon & \mu'\\ 0 & \epsilon\end{pmatrix}$ if and only if $P$ is upper triangular. In this case, writing $P = \begin{pmatrix}x & b\\ 0 & x^{-1}\end{pmatrix}$, the effect of conjugation by $P$ is to multiply $\mu$ by $x^2$. So $M(A)$ is either 0 (when $A = \pm I$), or is the set of elements of $C_{q-1}$ that are squares, or is the set of non-squares. We have verified the following addendum to proposition 2.1, given in [12, p. 64].

**Proposition 6.1.** *For $\epsilon = \pm 1$, the conjugacy classes of elements of $\mathrm{SL}(2,q)$ of trace $2\epsilon$ are determined by the invariant $M(A)$. When $p = 2$, there is one conjugacy class besides that of $\epsilon I$. When $p$ is odd, there are two conjugacy classes besides that of $\epsilon I$.*

We can now give the main result of this section.

**Theorem 6.2.** *Let $\{A, B\}$ generate $\mathrm{PSL}(2,q)$. If $q \not\equiv 1 \mod 4$, or the trace invariant of $\{A, B\}$ is not $-2$, then the trace invariant determines the Nielsen invariant of $\{A, B\}$. When $q \equiv 1 \mod 4$ and the trace invariant is $-2$, there are at most two Nielsen invariants, distiguishable by the invariant $M([A, B])$.*

*Proof.* The trace determines the conjugacy class of an element of $\mathrm{SL}(2,q)$ for traces other than $\pm 2$, and proposition 1.2 shows that 2 never occurs as a trace invariant. So for traces other than $-2$, the trace determines the Nielsen invariant.

Suppose from now on that the trace invariant of the generating pair $\{A, B\}$ equals $-2$. In particular, by proposition 1.2, $p \neq 2$. Since $A$ and $B$ cannot commute in $\mathrm{PSL}(2,q)$, $[A, B] \neq -I$.

Suppose that $q \equiv 3 \mod 4$. Then $-1$ is not a square in $C_{q-1}$, so $M([B, A]) = -M([A, B]) \neq M([A, B])$. Thus the conjugacy classes of $[A, B]$ and $[B, A]$ are distinct and are the two conjugacy classes of matrices of trace $-2$ other than $-I$, so there is only one Nielsen invariant possible in this case.

Suppose now that $q \equiv 1 \mod 4$. Then $-1$ is a square in $C_{q-1}$, so $M([B, A]) = -M([A, B]) = M([A, B])$, and the Nielsen invariant is a single conjugacy class. Since the two conjugacy classes other than $-I$ are distinguished by the invariant $M([A, B])$, the last sentence of the theorem is established. □

We remark that the proof of theorem 6.2 shows that for $A \in \mathrm{SL}(2,q)$, $A$ is conjugate to $A^{-1}$ except when $\mathrm{tr}(A) = -2$ and $q \equiv 3 \mod 4$. Consequently, the Nielsen invariant of $\{A, B\}$ is simply the conjugacy class of $[A, B]$, apart from this exceptional case.

We now examine the realization of the Nielsen invariant in the case when it is distinct from the trace invariant. For $\mu \in C_{q-1}$, define $S(\mu)$ to be the set of elements of $C_{q-1}$ of the form $\lambda^2\mu$. Recall the elements $H_x$ and $J_y$ from section 2.

**Theorem 6.3.** *Suppose that $q > 11$ and $q \equiv 1 \mod 4$. Let $x$ generate $C_{q-1}$, and let $D = x - x^{-1}$. Then $\{H_x, J_{4D^{-2}}\}$ is a generating pair with trace invariant $-2$, and $M([H_x, J_{4D^{-2}}]) = S(x^2 - 1)$.*

*Proof.* Since $q > 11$, proposition 2.6 ensures that this is a generating pair, and lemma 2.4 shows that its trace invariant is $-2$.

Suppose that $P \begin{pmatrix} -1 & \mu \\ 0 & -1 \end{pmatrix} P^{-1} = [H_x, J_{4D^{-2}}]$. Writing $P$ as $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and using lemma 2.4, we compute that

$$\begin{pmatrix} -ac\mu - 1 & a^2\mu \\ -c^2\mu & ac\mu - 1 \end{pmatrix} = \begin{pmatrix} 1 - 4xD^{-1} & Dx(4D^{-2} + 1) \\ -4x^{-1}D^{-1} & 1 + 4x^{-1}D^{-1} \end{pmatrix},$$

so $M([H_x, J_{4D^{-2}}]) = S(\mu) = S(4x^{-1}D^{-1}) = S(Dx) = S(x^2 - 1)$. □

We do not know in general the possible values that can occur for $S(x^2-1)$ in theorem 6.3. For $q = 5$ and $q = 9$ only non-squares are obtained, and for $q = 13$ only squares, but both are obtained when $q = 17$ and $q = 29$. Marston Conder has shown us a proof using Paley graphs that when $p \equiv 1 \mod 4$, exactly $(q - 5)/4$ of the elements of the form $x^2 - 1$ in $\mathbb{F}_q$ are squares, and exactly $(q - 1)/4$ are non-squares. Computations suggest that the hypothesis of theorem 6.3 will hold for all large $q$.

We note that $S(\mu)$ is not an invariant of weak equivalence, indeed it is never invariant under all the automorphisms of $\mathrm{PSL}(2, q)$ resulting from conjugation by elements of $\mathrm{GL}(2, q)$.

## 7. The case of $A_5$

We do not have a general picture of the Nielsen equivalence classes of generating pairs of the groups $\mathrm{PSL}(2, q)$, although we hope to make further progress on this. The case of $A_5$ can be worked out by hand, and gives a pretty example which we now present.

When we regard $A_5$ as $\mathrm{PSL}(2, 4)$, proposition 3.1 gives three trace invariants realized. Of course, $x$ and $x + 1$ form an orbit under the Frobenius automorphism, so there are only two weak trace invariants. When we regard $A_5$ as $\mathrm{PSL}(2, 5)$, as in proposition 3.2, only two trace invariants occur, that is, over $\mathbb{F}_5$ there are at least two Nielsen equivalence classes with the same trace invariant.

In fact, $A_5$ has exactly the three equivalence classes forming two weak equivalence classes that are detected over $\mathbb{F}_4$ or, as we will now see, by the Nielsen invariant (with values in $A_5$ itself). First recall that $A_5$ has 24 elements of order 5, forming two conjugacy classes, 20 conjugate elements of order 3, and 15 conjugate involutions. From [30], any generating pair is Nielsen equivalent to a pair consisting of a 5-cycle and an involution. There are 360 such pairs, but for each 5-cycle, five involutions pair with it to generate a dihedral subgroup, so only 240 such pairs generate. We call these 240 pairs *admissible.* Starting from the pair $\{(12345), (12)(34)\}$ and simply conjugating one generator by the other, and taking inverses of the

5-cycles, we eventually obtain all twelve of the 5-cycles conjugate to (12345) in an equivalent admissible pair, and each such admissible pair produces four others by conjugating the involution by the 5-cycle, giving a total of 60 equivalent admissible pairs. These all have Nielsen invariant equal to the conjugacy class of (12345). The nontrivial outer autmorphism of $A_5$ can be represented by conjugation by (45), which sends the equivalence class containing these sixty admissible pairs to the equivalence class of the pair $\{(12354), (12)(35)\}$, which has Nielsen invariant the conjugacy class of (12354) and consequently is a distinct Nielsen equivalence class. The generating pair $\{(12345), (13)(24)\}$ has Nielsen invariant the conjugacy class of order 3 elements, and by a similar process of conjugating one generator by another we obtain sixty admissible pairs equivalent to this one. However, we also have an equivalence

$$\{(12345), (13)(24)\} \sim \{(12345) \cdot (13)(24), (13)(24)\} =$$
$$\{(14532), (13)(24)\} \sim \{(12354), (13)(24)\} \sim \{(12354), (13)(25)\} ,$$

the penultimate step by inverting the 5-cycle, and the final one by conjugating the involution by the 5-cycle. Therefore these sixty are equivalent to their sixty images under the outer automorphism, and the third equivalence class contains the other 120 admissible pairs.

The fact that there are exactly two weak equivalence classes of generating pairs was proven in [30].

## 8. Group actions on handlebodies

Actions of finite groups on 3-dimensional handlebodies have been extensively studied. A general theory of such actions was given in [26] and [15], and the actions on very low genera were examined in [16]. Actions with the genus small relative to the order of the group were investigated in [28], while [14] treats the special case of orientation-reversing involutions.

We consider here only smooth (or PL) effective actions of a finite group $G$ preserving orientation on a 3-dimensional handlebody $V$. Two such actions $\phi_1, \phi_2 \colon G \to \mathrm{Diff}(V)$ are said to be *equivalent* if there is a diffeomorphism $h$ of $V$ such that $h\phi_1(g)h^{-1} = \phi_2(g)$ for all $g \in G$, and *weakly equivalent* if there is an automorphism $\alpha$ of $G$ such that $\phi_1$ and $\phi_2 \circ \alpha$ are equivalent.

*Free* orientation-preserving actions of finite groups on handlebodies were studied for cyclic groups by J. Przytycki [35], and in general in [27]. It is an elementary fact (see section 1 of [27]) that a finite group $G$ acts freely and preserving orientation on a 3-dimensional handlebody if and only if the genus of that handlebody has the form $1 + |G|(n-1)$ with $n$ an integer greater than or equal to $\mu(G)$, the minimal number of generators of $G$.

It is known that the equivalence classes (respectively, weak equivalence classes) of free orientation-preserving actions of $G$ on the handlebody of genus $1 + |G|(n-1)$ correspond to the Nielsen equivalence classes (respectively, weak Nielsen equivalence classes) of $n$-element generating sets of $G$.

This has been known for some time and can be deduced from earlier results in the literature, although the first explicit statement seems to be theorem 2.3 of [27]. As a consequence of this correspondence, the Main Corollary gives lower bounds for the number of weak equivalence classes of free actions of the groups $\mathrm{PSL}(2,q)$ on the handlebody of genus $1+q(q^2-1)/d$. Using our crude but effective lower bound for $\Psi_q$, we can state a simple consequence:

**Corollary 8.1.** *For $q > 11$, $\mathrm{PSL}(2,q)$ has at least $\frac{q}{s} - 1$ weak equivalence classes of orientation-preserving free actions on the handlebody of genus $1+q(q^2-1)/d$.*

In the discussion at the end of section 9, we will see that the groups $\mathrm{PSL}(2,q)$ appear to have many more actions, relative to their orders, than previously known examples.

## 9. Open questions

There are intriguing open questions about Nielsen equivalence. Let $G$ denote a finite generated group, and write $\mu(G)$ for the minimal cardinality of a generating set of $G$.

*Question 1:* For $n > \mu(G)$, are any two generating $n$-vectors Nielsen equivalent?

According to p. 92 of [24], this was first asked by F. Waldhausen. Question 1 has been answered affirmatively for a number of cases, including solvable groups [8], $\mathrm{PSL}(2,p)$ with $p$ prime [11], $\mathrm{PSL}(2,2^m)$ with $m \geq 2$ [9], $\mathrm{PSL}(2,3^p)$ with $p$ prime [27], and the Suzuki groups $\mathrm{Sz}(2^{2m-1})$ [9]. The corresponding question for infinite groups has been resolved negatively. The such first example appears to be due to G. A. Noskov [33], and general constructions are given in [10].

An affirmative answer to Question 1 for a given group implies affirmative answers to the next two questions, and conversely they together imply the affirmative for Question 1.

*Question 2:* Is every generating $n$-vector $(g_1,\ldots,g_n)$ of $G$ Nielsen equivalent to a generating $n$-vector of the form $(h_1,\ldots,h_{\mu(G)},1,\ldots,1)$?

*Question 3:* If $(g_1,\ldots,g_n)$ and $(h_1,\ldots,h_n)$ are any two generating $n$-vectors of $G$, are $(g_1,\ldots,g_n,1)$ and $(h_1,\ldots,h_n,1)$ Nielsen equivalent?

These general questions concern generating tuples of more than minimal length. We state also a general problem:

*Problem 4:* Investigate Nielsen equivalence of minimal-length generating tuples.

In this paper, we have made progress on Problem 4 for $\mathrm{PSL}(2,q)$, but we do not know how far we remain from a full classification of the equivalence classes. When $q < 11$, the weak trace invariant is a complete invariant for weak equivalence. Indeed, there is one weak equivalence class for $q = 2$ and

$q = 3$ (see section 3), there are two for $q = 4$ and $q = 5$ (see section 7), and in the cases of $q = 7$ and $q = 9$ (since $\mathrm{PSL}(2, 9) \cong A_6$), D. Stork [38] showed that there are exactly four. All of these agree with the number of weak trace invariants found in the Main Corollary, although we have no reason to expect this to continue for larger $q$. We do not know any inequivalent generating pairs with the same trace invariant other than the very restricted examples from section 6, and these are distinguished by the Nielsen invariant. So as far as we know, the Nielsen invariant is a complete invariant of equivalence for generating pairs of $\mathrm{PSL}(2, q)$, although again we have no reason to expect this to hold in general.

A specific aspect of Problem 4 is:

*Problem 5:* Find groups with a large number of weak equivalence classes of minimal-length generating tuples.

Dunwoody [7] gave examples of solvable groups with arbitrarily large numbers of weak equivalence classes of generating $n$-vectors.

**Theorem** (Dunwoody)**.** *For every pair of integers $n > 1$ and $N > 0$ and every prime $p$, there exists a $p$-group $G(n, N)$, nilpotent of length 2 and with $\mu(G(n, N)) = n$, for which there are at least $N$ weak equivalence classes of generating $n$-vectors.*

Let us examine $N$ relative to $|G|$ for the Dunwoody examples. When $n > 2$, the group $G(n, N)$ has larger order than $G(2, N)$, so $n = 2$ achieves (as far as we know) the most equivalence classes for its order. Here is the construction of the groups $G(2, N)$. Fix a prime $p$, and let $q = p^s$ with $s \geq 1$. Let $A$ be a cyclic group generated by an element $a$ of order $q^4$, and $B$ a cyclic group generated by an element $b$ of order $q^5$. Form the semidirect product $A \circ B$ in which $b^{-1}ab = a^{q^2+1}$. The elements $a^{q^3}$ and $b^{q^4}$ are central and have order $q$, so the element $a^{q^3}b^{-q^4}$ generates a central subgroup of order $q$. Define $G(q)$ to be the quotient of $A \circ B$ by this subgroup, so the order of $G(q)$ is $q^8$. In [7], it is shown that $G(q)$ has at least $\frac{1}{2}(p-1)p^{s-1}$ weak Nielsen equivalence classes of generating 2-vectors. This bound is on the order of $\frac{1}{2}$ times the eighth root of $|G(q)|$. For the groups $\mathrm{PSL}(2, q)$, our results give $\frac{q}{s} - 1$ as a lower bound, and this is on the order of $\frac{1}{s}$ times the cube root of $|\mathrm{PSL}(2, q)|$.

## References

1. Kenneth P. Bogart, An obvious proof of Burnside's lemma, *Amer. Math. Monthly* 98 (1991), 927–928.
2. A. M. Brunner, Metabelian representations and the isomorphism problem, *Math. Z.* 162 (1978), 235–240.
3. D. J. Collins, Generation and presentation of one-relator groups with centre, *Math. Z.* 157 (1977), 63–77.
4. H. S. M. Coxeter and W. O. J. Moser, Generators and relations for discrete groups (Fourth edition), Ergebnisse der Mathematik und ihrer Grenzgebiete 14, Springer-Verlag, Berlin-New York, 1980.

5. L. E. Dickson, Linear groups: with an exposition of the Galois field theory, Leipzig (1901), reprinted by Dover Publications, Inc., New York (1958).
6. J. Dieudonné, On the automorphisms of the classical groups, with a supplement by Loo-Keng Hua, *Mem. Amer. Math. Soc.* 2 (1951), 1–122.
7. M. Dunwoody, On $T$-systems of groups, *J. Australian Math. Soc.* 3 (1963), 172–179.
8. M. Dunwoody, Nielsen transformations, in *Computational Problems in Abstract Algebra* (Proc. Conf. Oxford, 1967), Pergamon, Oxford (1970), 45–46.
9. M. Evans, $T$-systems of certain finite simple groups, *Math. Proc. Cambridge Phil. Soc.* (1993), 9–22.
10. M. Evans, Presentations of groups involving more generators than are necessary, *Proc. London Math. Soc.* (3) 67 (1993), 106–126.
11. R. Gilman, Finite quotients of the automorphism group of a free group, *Can. J. Math.* 29 (1977), 541–551.
12. H. Glover and D. Sjerve, The genus of $\mathrm{PSl}_2(q)$, *J. Reine Angew. Math.* 380 (1987), 59–86.
13. Loo-Keng Hua, On the automorphisms of the symplectic group over any field, *Ann. of Math.* (2) 49, (1948), 739–759.
14. J. Kalliongis and D. McCullough, Orientation-reversing involutions of handlebodies, *Trans. Amer. Math. Soc.* 348 (1996),1739–1755.
15. J. Kalliongis and A. Miller, Equivalence and strong equivalence of actions on handlebodies, *Trans. Amer. Math. Soc.* 308 (1988), 721–745.
16. J. Kalliongis and A. Miller, The symmetries of genus one handlebodies, *Canad. J. Math.* 43 (1991), 371–404.
17. R. Lidl and H. Neiderreiter, *Introduction to finite fields and their applications* (revised edition), Cambridge University Press (1994).
18. M. Lustig, Nielsen equivalence and simple-homotopy type, *Proc. London Math. Soc.* (3) 62 (1991), 537–562.
19. M. Lustig and Y. Moriah, Nielsen equivalence in Fuchsian groups and Seifert fibered spaces, *Topology* 30 (1991), 191–204.
20. M. Lustig and Y. Moriah, Generalized Montesinos knots, tunnels and $N$-torsion, *Math. Ann.* 295 (1993), 167–189.
21. M. Lustig and Y. Moriah, Generating systems of groups and Reidemeister-Whitehead torsion, *J. Algebra* 157 (1993), 170–198.
22. M. Lustig, Y. Moriah, $N$-torsion and applications, in *Geometric group theory,* Vol. 1 (Sussex, 1991), 159–168, London Math. Soc. Lecture Note Ser., 181, Cambridge Univ. Press, Cambridge, 1993.
23. M. Lustig and Y. Moriah, On the complexity of the Heegaard structure of hyperbolic 3-manifolds, *Math. Z.* 226 (1997), 349–358.
24. R. C. Lyndon and P. E. Schupp, *Combinatorial Group Theory,* Springer-Verlag, 1977.
25. A. M. Macbeath, Generators of the linear fractional groups, *Number Theory* (Proc. Sympos. Pure Math., Vol. XII, Houston, Tex., 1967), Amer. Math. Soc., Providence, R.I. (1969) 14–32.
26. D. McCullough, A. Miller, and B. Zimmermann, Group actions on handlebodies, *Proc. London Math. Soc.* (3) 59 (1989), 373–416.
27. D. McCullough and M. Wanderley, Free actions on handlebodies, *J. Pure Appl. Alg.* 181 (2003), 85–104.
28. A. Miller and B. Zimmermann, Large groups of symmetries of handlebodies, *Proc. Amer. Math. Soc.* 106 (1989), 829–838.
29. Y. Moriah, Heegaard splittings of Seifert fibered spaces, *Invent. Math.* 91 (1988), 465–481.
30. B. H. Neumann and H. Neumann, Zwei Klassencharakteristischer Untergruppen und ihre Factorgruppen, *Math. Nachr.* 4 (1951) 106–125.
31. P. Neumann, A lemma which is not Burnside's, *Math. Sci.* 4 (1979), 133–141

32. J. Nielsen, Die Isomorphismengruppe der allgemeinen unendlichen Gruppe mit zwei Erzeugenden, *Math. Ann.* 78 (1918), 385–397.
33. G. A. Noskov, Primitive elements in a free group, *Math. Zametki* 30 (1981), 497–500.
34. S. J. Pride, The isomorphism problem for two-generator one-relator groups with torsion is solvable, *Trans. Amer. Math. Soc.* 227 (1977), 109–139.
35. J. H. Przytycki, Free actions of $\mathbb{Z}_n$ on handlebodies, *Bull. Acad. Polonaise des Sciences* XXVI (1978), 617–624.
36. G. Rosenberger, The isomorphism problem for cyclically pinched one-relator groups, *J. Pure Appl. Algebra* 95 (1994), 75–86.
37. O. Schrier and B. L. van der Waerden, Die Automorphismen der projektiven Gruppen, *Abh. Math. Sem. Univ. Hamburg* 6 (1928), 303–322.
38. D. Stork, The action of the automorphism group of $F_2$ upon the $A_6$- and PSL$(2,7)$-defining subgroups of $F_2$, *Trans. Amer. Math. Soc.* 172 (1972), 111–117.
39. M. Suzuki, *Group Theory, vol. I,* Springer-Verlag, 1982.
40. P. J. Webb, Minimal relation modules of free nilpotent groups, *Arch. Math.* (Basel) 37 (1981), 193–197.

Department of Mathematics, University of Oklahoma, Norman, Oklahoma 73019, USA
*E-mail address*: dmccullough@math.ou.edu
*URL*: www.math.ou.edu/~dmccullo/

Departmento de Matematica, Universidade Federal de Pernambuco, Av. Prof. Luiz Freire, s/n, Cid. Universitaria-Recife-PE, CEP 50.740-540, Brazil
*E-mail address*: mvw@dmat.ufpe.br