THE ACTION OF THE MODULAR GROUP ON CHARACTERS OF $\mathrm{SL}(2,q)$-REPRESENTATIONS OF F_2

DARRYL MCCULLOUGH AND MARCUS WANDERLEY

Abstract. We present a general conjectural picture of Nielsen equivalence classes and T-systems of generating pairs for the groups $\mathrm{SL}(2,q)$ and $\mathrm{PSL}(2,q)$, and the closely related Markoff equivalence classes of triples of elements of the coefficient field $\mathbb{F}_q$, viewed as characters of the F_2-representations corresponding to the generating pairs. The trace of the commutator of the elements of a generating pair is an invariant of its Nielsen class, and the orbit of this trace under the action of $\mathrm{Aut}(\mathbb{F}_q)$ on $\mathbb{F}_q$ is an invariant of the T-system of the pair. We show that as long as $q \geq 13$, these invariants assume all field values except for 2, which is never the trace of a generating pair. This and other results verify parts of the general conjectural picture. We also describe computer calculations that have verified the conjectures for all $q \leq 101$. Finally, we prove that all the conjectures hold for one highly restricted but possibly infinite class of values of q.

Introduction

In this paper we will study Nielsen equivalence and T-equivalence of generating pairs of $\mathrm{SL}(2,q)$ (and $\mathrm{PSL}(2,q)$), and the closely related action of the extended modular group $\mathrm{PGL}(2,\mathbb{Z})$ on the characters of $\mathrm{SL}(2,q)$-representations of F_2, the free group on two generators. In some respects, the latter is analogous to W. Goldman's [10] study of the corresponding action on characters of $\mathrm{SL}(2,\mathbb{R})$-representations of F_2. We will give a general conjectural picture of these related structures. Besides obtaining general information about the conjectures, and proving portions of them for various cases, we will prove the full set of conjectures for a very restricted (but conceivably infinite) class of examples. Also, we will describe our computer calculations that have verified the conjectures for all $q \leq 101$.

Part of our work extends a result of R. Guralnick and I. Pak [11], which states that as primes $p \to \infty$, the number of T-systems of generating pairs of $\mathrm{SL}(2,p)$ goes to ∞. We will show that for $q = p^s \geq 13$, the number of

Date: February 26, 2006.

1991 *Mathematics Subject Classification.* Primary 20G40; Secondary 20H10 57M60.

Key words and phrases. free, group, special linear, Nielsen equivalence, T-system, Markoff, Higman invariant, modular group, representation, character, GAP.

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

T-systems of $\mathrm{SL}(2,q)$ is bounded below by $\Psi_q - 1$, where

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

This is just the number of orbits of the $\mathrm{Aut}(\mathbb{F}_q)$-action on the field of q elements $\mathbb{F}_q$. Among our conjectures is that for $q \geq 13$, the number of T-systems of generating pairs of $\mathrm{SL}(2,q)$ is exactly $\Psi_q - 1$.

Information about Nielsen equivalence classes and T-systems has applications in algebra and topology. For example, they classify two kinds of group actions on low-dimensional manifolds, which happen to be our original motivation. The first kind are free actions on 3-dimensional handlebodies, for which we refer the reader to [18] and [4]. The second kind we call *almost free* actions on 2-manifolds. These are actions that are not free, but are close to being free in the sense that the group acts transitively on the union of the fixed-point sets of its nontrivial elements. Since we are not aware of any treatment of this class of actions in the literature, we include an exposition of their basic theory, which shows how they can be classified by means of Nielsen equivalence and T-systems.

We have already mentioned that our work is somewhat analogous to W. Goldman's work on characters of F_2-representations into $\mathrm{SL}(2,\mathbb{R})$ [10]. We should also note a connection to some work B. Bowditch [2, 3], who obtained striking results on Markoff equivalence for real and complex triples. Our study of Markoff equivalence in $\mathbb{F}_q^3$ is a rudimentary extension of this viewpoint into the finite field setting.

The authors gratefully acknowledge the support of many sources in the course of this work. Travel of the first author was supported by National Science Foundation grant DMS-0102463, by the University of Oklahoma College of Arts and Sciences, and by the University of Oklahoma Office of Research Administration. The work of both authors was supported by the Universidade Federal de Pernambuco Summer in Recife program. We also owe a special thanks to Gareth Jones, who provided us with an elegant proof of the formula for Ψ_q.

1. A conjectural picture of F_2-representations in $\mathrm{SL}(2,q)$

In this section we first present the conjectural picture of Nielsen equivalence in $\mathrm{SL}(2,q)$ and of the action of the modular group on the (essential) characters. The unsupported assertions in the discussion will be verified or referenced in later parts of the paper. As we present the conjectures, we mention several of our results and explain which parts of the overall picture they verify or support. Finally, we outline the sections of the paper.

Fixing a basis of F_2, we identify the set $\mathrm{Hom}(F_2, G)$ of G-representations of F_2 with the set of pairs of elements of the group G. We call such a representation *essential* when it is surjective, that is, when the corresponding pair of elements generates G; thus we may identify the set of essential representations with the set $\mathcal{G}_2(G)$ of generating pairs of G. The group $\mathrm{Aut}(F_2)$

acts on the left on the set of essential representations, by $\phi \cdot \rho = \rho \circ \phi^{-1}$. Its orbits are called *Nielsen equivalence classes,* or just *Nielsen classes.* We denote by $\mathcal{N}$ the set of Nielsen classes of a group under discussion.

For an element $(A, B) \in \mathcal{G}_2(G)$, the union of the conjugacy classes of the commutator $[A, B] = ABA^{-1}B^{-1}$ and its inverse is a well-known invariant of the Nielsen class of (A, B), called the *Higman invariant.* Since $[A, B]$ and $[B, A]$ are usually conjugate for elements of $\mathrm{SL}(2, q)$, our Higman invariants will usually consist of a single conjugacy class.

We now specialize to the case when G is $\mathrm{SL}(2, q)$. The field with q elements will be denoted by $\mathbb{F}_q$, and throughout this paper p will denote the prime such that $q = p^s$. As will be discussed in section 14 below, each of the conjectures that we will state here implies a corresponding assertion for the case of $\mathrm{PSL}(2, q)$, while the corresponding assertion implies a weak form of the conjecture.

Five conjectures called A, B, B′, C, and W will be stated here. For easy reference, we collect here the logical relationships that either are immediate or are ensured by the results in this paper:

$$A \Leftrightarrow B$$
$$B \Leftrightarrow (B' \wedge C)$$
$$B \Rightarrow W$$

We will see below that Conjecture B′ is known in characteristic 2, so in that case Conjectures A, B, and C are equivalent, and imply Conjecture W.

Our first conjecture says that the Higman invariant is a complete invariant of Nielsen equivalence:

Conjecture A (Higman invariant classifies Nielsen classes)**.** *Two generating pairs (A, B) and (A', B') of $\mathrm{SL}(2, q)$ are Nielsen equivalent if and only if $[A, B]$ is conjugate to $[A', B']$ or to $[B', A']$.*

The conjugacy classes of $\mathrm{SL}(2, q)$ are very well-known, and corollary 6.2 below shows that for $q \geq 13$, every conjugacy class except $\{-I\}$ and those of trace 2 occurs as a Higman invariant. Thus Conjecture A would give a complete classification of the Nielsen classes.

The Higman invariant shows that the trace of the commutator $[A, B]$ is a well-defined invariant of the Nielsen class of (A, B), which we call the *trace invariant.* The conjectural picture is that the trace invariant $\mathrm{tr}\colon \mathcal{N} \to \mathbb{F}_q$ is very close to a bijection. Indeed, theorem 5.1 below implies that tr has image $\mathbb{F}_q - \{2\}$ for all $q \geq 13$ (it is an easy fact that $\mathrm{tr}([A, B])$ can never equal 2 for a generating pair), and our second conjecture completes the picture from this viewpoint:

Conjecture B (Trace invariant is nearly injective)**.** *The trace invariant $\mathrm{tr}\colon \mathcal{N} \to \mathbb{F}_q - \{2\}$ is injective except that it is two-to-one on the preimage of -2 when $q \equiv 1 \bmod 4$.*

As we will detail in section 6, well-known facts about conjugacy in $\mathrm{SL}(2, q)$ show that Conjectures A and B are equivalent.

We turn now to characters. Regarding an $\mathrm{SL}(2,q)$-representation of F_2 as a pair of elements (A,B), we define the *character* of the representation to be the element $(\mathrm{tr}(A),\mathrm{tr}(B),\mathrm{tr}(AB))$ of $\mathbb{F}_q^3$. For notational simplicity, we denote $(\mathrm{tr}(A),\mathrm{tr}(B),\mathrm{tr}(AB))$ by $\mathrm{Tr}(A,B)$ and call it the *Fricke trace,* or often just the *trace,* of the pair (A,B). A character is called *essential* when it is the character of an essential representation, that is, when it is the trace of a generating pair.

The action of $\mathrm{Aut}(F_2)$ on the essential representations induces an action of the extended modular group $\mathrm{PGL}(2,\mathbb{Z})$ on the essential characters, whose orbits we call *Markoff classes.* Explicitly, Markoff equivalence is generated by permutations of the three coordinates, together with the relation that $(\alpha,\beta,\gamma)\sim(\alpha,\beta,\alpha\beta-\gamma)$. Since this action is induced from the action on representations, there is a well-defined Fricke trace function Tr from $\mathcal{N}$ to the set $\mathcal{M}$ of Markoff classes of essential characters.

The Fricke polynomial $Q(\alpha,\beta,\gamma)=\alpha^2+\beta^2+\gamma^2-\alpha\beta\gamma-2$ has the property that $Q(\mathrm{Tr}(A,B))=\mathrm{tr}([A,B])$, and consequently the value of Q is an invariant of Markoff equivalence. For $q\geq 13$, our result that $\mathrm{tr}\colon\mathcal{N}\to\mathbb{F}_q-\{2\}$ is surjective implies that $Q\colon\mathcal{M}\to\mathbb{F}_q-\{2\}$ is surjective. This brings us to our main conjecture on essential characters:

Conjecture C (Q classifies Markoff classes)**.** *Two essential characters are Markoff equivalent if and only if they have the same Q-value.*

In summary, Conjectures C and B say respectively that for $q\geq 13$, in the diagram

$$\begin{tikzcd}
\mathcal{N} \arrow[r, "\mathrm{Tr}"] \arrow[d, "\mathrm{tr}"'] & \mathcal{M} \arrow[d, "Q"] \\
\mathbb{F}_q-\{2\} \arrow[r, equal] & \mathbb{F}_q-\{2\}
\end{tikzcd},$$

Q is a bijection, and tr is a bijection except that it is two-to-one on the preimage of -2 when $q\equiv 1 \bmod 4$.

At the end of section 8, we will see that if Conjecture B holds, then in the case when $q\equiv 1 \bmod 4$, the Fricke traces of the two elements in $\mathrm{tr}^{-1}(-2)$ are in the same Markoff class in $\mathcal{M}$. It follows that Conjecture B implies Conjecture C, as well as the following assertion:

Conjecture B′ (Fricke trace is nearly injective)**.** *The Fricke trace map* $\mathrm{Tr}\colon\mathcal{N}\to\mathcal{M}$ *is injective except that it is two-to-one on the preimage of* $Q^{-1}(-2)$ *when* $q\equiv 1 \bmod 4$.

On the other hand, Conjectures B′ and C together immediately imply Conjecture B.

Results of A. M. Macbeath, that we will present in section 4, show that if M is a Markoff class then $\mathrm{Tr}^{-1}(M)$ consists of at most two Nielsen classes, and moreover that Tr is a bijection from $\mathcal{N}$ to $\mathcal{M}$ when q is even. This leads to the following:

1) When q is even, Conjecture B$'$ holds, so Conjecture C is equivalent to Conjectures A and B.
2) When q is odd, Macbeath's results and proposition 10.5 show that if $2-\ell$ is not a square in $\mathbb{F}_q$ (where $\ell \neq 2$), then Tr is injective on the preimage of $Q^{-1}(\ell)$. That is, Conjecture B$'$ always holds over at least half of the level surfaces of Q.

For many applications, Nielsen equivalence is too strong an invariant. More natural is to extend the action of $\mathrm{Aut}(F_2)$ on representations to an action of $\mathrm{Aut}(G) \times \mathrm{Aut}(F_2)$ by letting $(\alpha, \phi)\cdot\rho = \alpha\circ\rho\circ\phi^{-1}$. The resulting equivalence classes are called *T-systems.* It is known that $\mathrm{Aut}(\mathrm{SL}(2,q))$ is generated by conjugations by elements of $\mathrm{GL}(2,q)$, which do not change $\mathrm{tr}([A,B])$, together with field automorphisms of $\mathbb{F}_q$ acting on the entries of the elements of $\mathrm{SL}(2,q)$, whose effect is to apply the same field automorphism to $\mathrm{tr}([A,B])$. Thus the orbit of $\mathrm{tr}([A,B])$ in $\mathbb{F}_q$ is an invariant of the T-system of (A,B), called the *weak trace invariant.* The number of orbits of $\mathrm{Aut}(\mathbb{F}_q)$ acting on $\mathbb{F}_q$ is the number Ψ_q stated in the second paragraph of the paper, and discussed more fully in section 7 below. Theorem 5.1 shows that for $q \geq 13$, $\Psi_q - 1$ is a lower bound for the number of T-systems of $\mathrm{SL}(2,q)$, and the next conjecture implies that this is the exact number. Denote the set of T-systems by $\mathcal{T}$ and the set of orbits of $\mathrm{Aut}(\mathbb{F}_q)$ acting on $\mathbb{F}_q$ by $\mathcal{O}_q$.

Conjecture W (Weak trace invariant classifies T-systems)**.** *For $q \geq 13$ the weak trace invariant $\mathrm{tr}\colon \mathcal{T} \to \mathcal{O}_q - \{2\}$ is a bijection.*

In section 7, we will see that Conjecture B implies Conjecture W.

As we mentioned above and will discuss in section 9, all of these conjectures have been verified computationally for all $q \leq 101$, and the minor exceptions that occur when $q \leq 11$ are also completely calculated. We will also verify all conjectures for the (almost laughably restrictive, but conceivably infinite) class of all q such that $q-1$ is prime and $q+1$ has the form $3p_1$ for some prime p_1.

Here is a brief outline of the exposition. Section 2 reviews Nielsen equivalence and T-systems, section 3 introduces the Fricke polynomial Q, and section 4 recalls results from Macbeath's elegant paper [15]. In section 5 we prove theorem 5.1 that realizes elements of $\mathbb{F}_q$ as trace invariants, and we have already mentioned the work in section 6 relating the Higman and Nielsen invariants. Section 7 introduces the weak trace invariant and gives a proof of the formula for Ψ_q. Markoff equivalence in the character variety is introduced in section 8, at which point we can present the discussion of our computer calculations in section 9.

In section 10, we review the concepts of elliptic, parabolic, and hyperbolic elements of $\mathbb{F}_q$. This has a significant application in odd characteristic, proposition 10.5 that appeared above in the discussion of Conjecture B$'$.

Apart from the final two sections, which contains the aforementioned treatments of $\mathrm{PSL}(2,q)$ and of almost free actions, the last few sections of the paper are directed toward the proof of Conjecture C, and hence all the conjectures, in the highly restricted case when one of $q+1$ or $q-1$ is prime and the other is 3 times a prime. This proof of Conjecture C requires quite a bit of work, which we hope will someday be justified by further progress at least when $p=2$. The restriction to $p=2$ arises initially because of a characterization of the hyperbolic elements, for which no analogue is apparent in odd characteristic: the inverses of the hyperbolic elements, together with 0, are exactly the kernel of the trace map from $\mathbb{F}_q$ to $\mathbb{F}_2$ (lemma 10.2).

Sections 11 and 12 are presented for arbitrary q, although only used later for even q. In section 11, we examine more closely how the level surfaces of Q meet the "slices" of $\mathbb{F}_q$ with fixed values of the first coordinate α, detailing how the pattern of their Q-values is governed by whether α is elliptic, parabolic, or hyperbolic. The action of $\mathrm{Aut}(F_2)$ on these individual slices is analyzed. Concrete examples of slices are given in section 12, to illustrate and help with understanding of section 11.

Section 13 is fully specialized to even q. A number of further simplifications occur, which allow us to obtain a workable description of the action within a fixed slice. Even with all this information, it is difficult enough to understand the full action that we require the further strong assumptions on $q-1$ and $q+1$ to force the level surface of Q into one big orbit. The need for such assumptions is rather frustrating, since in numerical calculations the orbit of a point seems to enlarge very rapidly to fill up the level surface as automorphisms are applied, but we have not been able to discern a key property that enables us to control this; it just seems random.

A possible reason for caution about the conjectures arises from section 13, where a very important role is played by the "transitive" elements of $\mathbb{F}_q-\{0\}$, which are the traces of matrices of large orders $q-1$ or $q+1$. For very large q, the proportion of such elements may become arbitrarily small, much smaller than in the cases $q\leq 101$ that we have checked computationally or the cases where $q-1$ and $q+1$ satisfy the very strong primeness assumptions. So in this possibly relevant sense the cases for which the conjectures are known are not representative of the general case. While this does not suggest that the conjectures are likely to fail, it does say that the accumulated evidence for them may not be as strong as it appears.

2. Nielsen equivalence and T-equivalence

The material in this section is very well-known. For simplicity we will restrict attention to two-generator groups G, but it can be adapted without difficulty to n-generator groups (see for example [4]).

We have defined Nielsen equivalence in the set of generating pairs $\mathcal{G}_2(G)$ as equivalence under the left action of $\mathrm{Aut}(F_2)$, and T-equivalence as equivalence under the left action of $\mathrm{Aut}(G)\times\mathrm{Aut}(F_2)$. Nielsen [20] found generators for $\mathrm{Aut}(F_2)$, which show that $\mathrm{Aut}(F_2)$ is generated by the three involutions $r\colon (x_1,x_2)\mapsto (x_1^{-1},x_1x_2)$, $s\colon (x_1,x_2)\mapsto (x_2,x_1)$, and $t\colon (x_1,x_2)\mapsto$

(x_1^{-1}, x_2), These act as follows on $\mathcal{G}_2(G)$:

$$\begin{aligned} r(g_1, g_2) &= (g_1^{-1}, g_1 g_2) \\ s(g_1, g_2) &= (g_2, g_1) \\ t(g_1, g_2) &= (g_1^{-1}, g_2) \ . \end{aligned}$$

These "basic moves" or simple compositions of them allow one to perform all product replacement moves, which replace one generator by the result of pre- or post-multiplying it by another generator or its inverse, together with the moves of replacing either generator by its inverse, or interchanging the two generators. From this viewpoint, T-equivalence just adds the additional basic moves of

$$\alpha(g_1, g_2) = (\alpha(g_1), \alpha(g_2))$$

for any $\alpha \in \mathrm{Aut}(G)$. When α is inner, this resulting pair is Nielsen equivalent to (g_1, g_2). So the action of $\mathrm{Aut}(G)$ on $\mathcal{G}_2(G)$ induces an action of $\mathrm{Out}(G)$ on the set $\mathcal{N}$ of Nielsen classes, whose quotient is the set $\mathcal{T}$ of T-systems.

3. The Fricke polynomial and the trace map on pairs

Fixing a basis of F_2, we regard representations $F_2 \to G$ for $G = \mathrm{SL}(2,q)$ or $G = \mathrm{PSL}(2,q)$ as pairs (A, B), where $A, B \in \mathrm{SL}(2,q)$. In the case $G = \mathrm{PSL}(2,q)$, we keep in mind that A and B are defined only up to sign. As in the introduction, we call a representation $F_2 \to G$ *essential* when it is surjective, that is, when (A, B) is a generating pair of G.

The *Fricke trace map* is the function $\mathrm{Tr}\colon \mathrm{Hom}(F_2, \mathrm{SL}(2,q)) \to \mathbb{F}_q^3$ defined by $\mathrm{Tr}(A,B) = (\mathrm{tr}(A), \mathrm{tr}(B), \mathrm{tr}(AB))$. The *Fricke polynomial* $Q\colon \mathbb{F}_q^3 \to \mathbb{F}_q$ is defined by

$$Q(\alpha, \beta, \gamma) = \alpha^2 + \beta^2 + \gamma^2 - \alpha\beta\gamma - 2 \ .$$

The well-known identity

$$\mathrm{tr}([A,B]) = Q(\mathrm{Tr}(A,B)) \ ,$$

can be obtained by repeatedly simplifying $\mathrm{tr}(ABA^{-1}B^{-1})$ using the fundamental identity $\mathrm{tr}(X)\,\mathrm{tr}(Y) = \mathrm{tr}(XY) + \mathrm{tr}(XY^{-1})$ for $X, Y \in \mathrm{SL}(2,F)$ (which follows from the Cayley-Hamilton identity $Y^2 - \mathrm{tr}(Y)Y + I = 0$ by multiplying on the left by XY^{-1} and taking traces). Since the expression $Q(\mathrm{Tr}(A,B))$ appears quite often in our work, we will abbreviate it to $Q(A,B)$. Since $Q(A,B) = Q(-A,B) = Q(A,-B) = Q(-A,-B)$, this Q is well-defined on $\mathrm{PSL}(2,q) \times \mathrm{PSL}(2,q)$.

4. G_0, G_1, G, and Macbeath's theorems

In this section we follow Macbeath [15]. For a fixed value of q define G_0 to be $\mathrm{SL}(2,q)$. There is a natural homomorphism $\pi\colon G_0 \to \mathrm{PSL}(2,q)$.

Let G_1 be the subgroup of $\mathrm{SL}(2,q^2)$ consisting of the matrices of the form $\begin{pmatrix} a & b \\ b^q & a^q \end{pmatrix}$. The subgroups G_0 and G_1 are conjugate in $\mathrm{SL}(2,\overline{\mathbb{F}_q})$, where $\overline{\mathbb{F}_q}$ is the algebraic closure of $\mathbb{F}_q$. From now on, when we write G_i, we mean either one of G_0 or G_1. We write G for $\mathrm{PSL}(2,q)$.

Theorem 1 of [15] says that every triple is a character:

Theorem 4.1 (Macbeath)**.** $\mathrm{Tr}\colon G_i \times G_i \to \mathbb{F}_q^3$ *is surjective.*

In section 10 below, we will give a computational proof of theorem 4.1. Also, a slight modification of Macbeath's more elegant approach is used to prove proposition 9.1.

Following [15], a subgroup of G is called *affine* if its premiage in $\mathrm{SL}(2,q)$ is conjugate, in either G_0 or G_1, into the subgroup of upper triangular matrices. Theorem 2 of [15] identifies the affine subgroups of $\mathrm{PSL}(2,q)$ in terms of Q.

Theorem 4.2 (Macbeath)**.** *A G_i-pair (A,B) generates an affine subgroup of G if and only if $Q(A,B) = 2$.*

Here, as in many places in our work, we speak of the subgroup of G generated by a G_i-pair. This means the subgroup generated by the images of A and B in G.

An $\mathbb{F}_q$-triple is called *singular* or *nonsingular* according as Q does or does not assign it the value 2. Theorem 4.2 shows that a pair with singular trace can never generate G.

Theorem 3 of [15] describes the conjugacy classes of G_i-pairs with nonsingular trace.

Theorem 4.3 (Macbeath)**.** *Let (α,β,γ) be a nonsingular $\mathbb{F}_q$-triple. If $p = 2$, then there is exactly one conjugacy class of G_i-pairs whose trace equals (α,β,γ). If $p > 2$, then there are exactly two conjugacy classes. These classes are conjugate in $\mathrm{SL}(2,\overline{\mathbb{F}_q})$, where $\overline{\mathbb{F}_q}$ is the algebraic closure of $\mathbb{F}_q$, and consequently generate isomorphic subgroups of G_i.*

5. Trace invariants

An element of $\mathbb{F}_q$ is called a *trace invariant* if it equals $\mathrm{tr}([A,B])$ (and hence equals $Q(A,B)$) for some generating pair (A,B) of $\mathrm{SL}(2,q)$. By theorem 4.2, the field element 2 is never a trace invariant. The main result of this section tells which elements of $\mathbb{F}_q$ do occur as trace invariants:

Theorem 5.1. *For $q = 2$, $q = 4$, $q = 8$, and all $q \geq 13$, the trace invariants are the elements of $\mathbb{F}_q - \{2\}$. For the remaining cases, the trace invariants are as follows:*

1) *For $q = 3$, $q = 9$, and $q = 11$, all elements except* 1 *and* 2.
2) *For $q = 5$, only* 1 *and* 3.
3) *For $q = 7$, all elements except* 0, 1, *and* 2.

For $q \leq 11$, theorem 5.1 can be proven in a few pages of elementary arguments of varying degrees of complication, and we omit these details. Also, we have checked those cases using GAP; these and other calculations are discussed in section 9 below. In the remainder of this section, we will prove theorem 5.1 for $q \geq 13$.

We first review some well-known facts about $\mathrm{PSL}(2,q)$. The subgroups of $\mathrm{PSL}(2,q)$ were determined by L. E. Dickson [5]. The following statement, in which d denotes $\gcd(2, q-1)$, is from Theorem 3(6.25) of Suzuki [24].

Theorem 5.2. *Every subgroup of* $\mathrm{PSL}(2,q)$ *is isomorphic to (at least) one of the following.*

(a) *The dihedral groups of orders* $2(q\pm 1)/d$ *and their subgroups.*
(b) *A group* H *of order* $q(q-1)/d$ *and its subgroups. A Sylow* p*-subgroup* H_0 *of* H *is elementary abelian, normal in* H*, and the factor group* H/H_0 *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 last statement in (d) is from 3(6.18) of [24]. The subgroups H_0 in (b) are also p-Sylow subgroups of $\mathrm{PSL}(2,q)$, so are conjugate to the subgroup of upper triangular elements.

We use the following terminology to refer to the subgroups described in theorem 5.2: subgroups as in (a) are called *small,* as in (c) are called *exceptional,* and as in (d) are called *linear.* The affine subgroups (defined in section 4) include the subgroups in (b) and the cyclic subgroups in (a) that are subgroups of a maximal cyclic subgroup of order $(q\pm 1)/d$. The groups in (d) coincide when $p=2$, and will be examined quite a bit more closely in section 9.

An element of $\mathrm{PSL}(2,q)$ is called *parabolic* if its trace if ± 2. Parabolic elements have order p. For nonparabolic elements, we have the following information from (2.3) and (2.4) of [9].

Lemma 5.3. *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 $x \neq 0$, 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 a straightforward calculation.

Lemma 5.4. *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* $2 - D^2 y$*.*

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

Lemma 5.5. *Assume that* $y \neq 0$*, and that* $x^4 \neq 1$ *and* $x^6 \neq 1$*. 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 $x^2 \neq 1$. Now $[H_x^{-1}, J_y] = [H_{x^{-1}}, J_y]$, so $[H_x^{-1}, J_y]$ is obtained from the expression in lemma 5.4 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 5.6. *Assume that* $q \geq 13$. *Suppose that* x *generates* $\mathbb{F}_q - \{0\}$ *and that* $y \neq 0$. *Then* H_x *and* J_y *generate* $\mathrm{SL}(2, q)$.

Proof. Since $q > 7$, we have $x^4 \neq 1$ and $x^6 \neq 1$. Let S be the image in $\mathrm{PSL}(2, q)$ of the subgroup 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 5.2. Lemma 5.5 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 linear, 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 5.3 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 5.3 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, $q = 9$. □

To see that proposition 5.6 implies theorem 5.1 in the case $q \geq 13$, let x be a generator of $\mathbb{F}_q - \{0\}$, and put $D = x - x^{-1}$. By proposition 5.6 and lemma 5.4, all traces of the form $2 - D^2y$ with $y \neq 0$ arise as trace invariants of generating pairs for $\mathrm{PSL}(2, q)$.

6. Comparison of the Higman invariant and trace invariant

In this section, we will see that each trace invariant determines a unique Higman invariant for $\mathrm{SL}(2, q)$ or $\mathrm{PSL}(2, q)$, except when the trace invariant is -2 and $q \equiv 1 \bmod 4$. In that case, there are two Higman invariants, except when $q = 9$, when there are none. Using that information, we will verify that Conjectures A and B stated in section 1 are equivalent.

The conjugacy classes in $\mathrm{SL}(2, q)$ are very well-known:

Proposition 6.1. *If $A, B \in \mathrm{SL}(2, q)$ and $\mathrm{tr}(A) \neq \pm 2$, then A is conjugate to B if and only $\mathrm{tr}(A) = \mathrm{tr}(B)$. For each of the traces 2 and -2, there are two conjugacy classes when q is even and three when q is odd.*

We will verify the second part, since we need the notation anyway. Consider an element X of $\mathrm{SL}(2, q)$ having trace 2ϵ where $\epsilon = \pm 1$. It is conjugate to a matrix of the form $\begin{pmatrix} \epsilon & \mu \\ 0 & \epsilon \end{pmatrix}$, so the nonempty set $M(X)$ of elements μ that appear in such conjugates is a complete invariant of the conjugacy class of X. 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 μ by x^2. So $M(X)$ is either 0 (when $X = \pm I$), or is the set of nonzero elements that are squares, or is the set of non-squares.

Using theorem 5.1, we can now determine the relation between Higman invariants and trace invariants.

Corollary 6.2. *Let (A, B) be a generating pair for $\mathrm{SL}(2, q)$ or $\mathrm{PSL}(2, q)$. If $q \not\equiv 1 \bmod 4$, or $\mathrm{tr}([A, B])$ is not -2, then the trace invariant determines the Higman invariant of (A, B). For $q \equiv 1 \bmod 4$ and trace -2, there are two Higman invariants when $q = 5$ or $q \geq 13$, while for $q = 9$ there are none.*

Proof. For $q \leq 11$ the theorem can be checked by direct computation, as described in section 9 below, and we assume that $q \geq 13$. Theorem 5.1 shows that 2 never occurs as a trace invariant, but that all other traces do. So proposition 6.1 shows that each trace other than -2 is the trace of a unique Higman invariant.

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

Suppose that $q \equiv 3 \bmod 4$. Then -1 is not a square in $\mathbb{F}_q$, 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 Higman invariant possible in this case.

Suppose now that $q \equiv 1 \bmod 4$. Then -1 is a square, so $M([B, A]) = -M([A, B]) = M([A, B])$, and the Higman invariant is a single conjugacy class. By theorem 5.1, at least one of the conjugacy classes of matrices of trace -2 is a Higman invariant; by conjugation we may choose a generating pair (A, B) having $[A, B] = \begin{pmatrix} -1 & t \\ 0 & -1 \end{pmatrix}$ for some nonzero $t \in \mathbb{F}_q$. Conjugating by a matrix in $\mathrm{SL}(2, q^2)$ of the form $\begin{pmatrix} \pi & 0 \\ 0 & \pi^{-1} \end{pmatrix}$, where $\pi^2 \in \mathbb{F}_q$ but $\pi \notin \mathbb{F}_q$, changes (A, B) to a generating pair (A', B') of $\mathrm{SL}(2, q)$ having commutator $\begin{pmatrix} -1 & t\pi^2 \\ 0 & -1 \end{pmatrix}$, realizing the other Higman invariant with trace -2. □

Corollary 6.2 implies that Conjecture A is equivalent to Conjecture B. For if $\mathcal{H}$ denotes the set of Higman invariants, then the trace map factors as $\mathcal{N} \to \mathcal{H} \to \mathbb{F}_q - \{2\}$, with the first function surjective. Conjecture A is that the first function is injective, as well. Corollary 6.2 shows that the first map is injective if and only if Conjecture B holds.

The pair of Nielsen classes in corollary 6.2 having trace invariant -2 but distinguished by their Higman invariants are always T-equivalent. For as seen in the proof, an automorphism $\mathrm{SL}(2, q)$ which is conjugation by a matrix $\begin{pmatrix} \pi & 0 \\ 0 & \pi^{-1} \end{pmatrix}$ in $\mathrm{SL}(2, q^2)$ with $\pi^2 \in \mathbb{F}_q$ but $\pi \notin \mathbb{F}_q$ interchanges their Higman invariants. Therefore Conjecture B implies Conjecture W.

7. Weak trace invariants

We turn now to T-equivalence. The automorphisms of $\mathrm{SL}(2, q)$ are well-understood, by the following result due to Schreier and van der Waerden [23] (see also [6] and the appendix to [12]).

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^{ϕ} *is the matrix obtained by applying an automorphism* ϕ *of* $\mathbb{F}_q$ *to each entry of* A*.*

Conjugation has no effect on the trace invariant, while applying a field automorphism to the coefficients of A and B changes the trace of $[A, B]$ by the field automorphism itself. Therefore the orbit of the trace invariant under $\mathrm{Aut}(\mathbb{F}_q)$ is an invariant of the T-system, which we call the *weak trace invariant.*

We saw at the end of section 6 that when $q \equiv 1 \bmod 4$, there are two distinct Nielsen classes which are T-equivalent. It follows that Conjecture B implies Conjecture W.

Denoting by Ψ_q the number of orbits of the action of $\mathrm{Aut}(\mathbb{F}_q)$ on $\mathbb{F}_q$, theorem 5.1 tells us immediately which orbits occur as weak trace invariants:

Corollary 7.1. *The numbers of orbits of the Frobenius automorphism that occur as weak trace invariants of generating pairs of* $\mathrm{SL}(2, q)$ *or* $\mathrm{PSL}(2, q)$ *are as follows:*

i) *If* $q = 2$*,* $q = 4$*,* $q = 8$*, or* $q \geq 13$*, 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.*

Recall that $\mathrm{Aut}(\mathbb{F}_q)$ is a cyclic group generated by the Frobenius automorphism Φ that sends each x to x^p. Since Φ has order s, the ceiling $\lceil \frac{q}{s} \rceil$ is trivially a lower bound for the number Ψ_q of orbits of $\mathrm{Aut}(\mathbb{F}_q)$, and since Φ fixes each element of the subfield $\mathbb{F}_p$, $\lceil \frac{q-p}{s} \rceil + p$ is an obvious lower bound. The exact number of orbits of $\mathrm{Aut}(\mathbb{F}_q)$ is given by the closed formula that we stated in the introduction:

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

where φ is the Euler totient function. This formula for Ψ_q is surely well-known, although we have not found an explicit statement in the literature. Experts in finite fields (we thank, in particular, H. Niederreiter) observe that the number of orbits is the same as the number of monic irreducible polynomials over $\mathbb{F}_p$ of degree dividing s, each orbit being the set of roots of one such polynomial. Consequently, the number $e(r)$ of orbits with r elements (which equals the number of monic irreducible polynomials of degree r) satisfies $q = \sum_{r|s} re(r)$, and a formula for $e(s)$ can be obtained using Möbius inversion (see for example [13, Ch. III.2]). In fact the formula for $e(s)$ is the same as our formula for Ψ_q but with φ replaced by the Möbius function μ. Summing these for r dividing s and then manipulating using the fact (also a consequence of Möbius inversion) that $\frac{\varphi(s)}{s} = \sum_{r|s} \frac{\mu(r)}{r}$ gives the formula for Ψ_q. Rather than writing out the details of that, or worse yet, leaving them to the reader, we will present here an elegant proof shown to us by Gareth Jones, that deduces the formula for Ψ_q in a few lines using Burnside's Lemma and a few of the most elementary properties of $\mathbb{F}_q$.

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 the group:

Lemma 7.2 (Burnside's Lemma)**.** *If a finite group G acts on a finite set Ω, 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.

Burnside's Lemma can be proven by elementary counting arguments (see for example [1]), and a better name for it is the Burnside-Cauchy-Frobenius formula (see [19]).

To obtain the formula for Ψ_q, we will apply Burnside's Lemma with $\Omega = \mathbb{F}_{p^s}$ and $G = \mathrm{Aut}(\mathbb{F}_{p^s})$. Recall that $\mathbb{F}_{p^r}$ occurs as a subfield of $\mathbb{F}_{p^s}$ if and only if $r|s$, and that it is the unique subfield of this order. Each element Φ^m of G has order s/r, where $r = \gcd(m,s)$, and there are $\varphi(s/r)$ elements of this order for each divisor r of s. Such an element has the same fixed points as Φ^r, since each is a power of the other, and the fixed points of Φ^r are the roots of the polynomial $x^{p^r} - x$. These roots form the subfield F_{p^r}, so $\pi(\Phi^m) = p^r$. Burnside's Lemma now yields the formula for Ψ_q.

The same argument, with q in the role of p, shows 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 $\displaystyle\frac{1}{s}\sum_{r|s} \varphi(s/r)\, q^r$.

The obvious lower bound $\lceil \frac{q-p}{s} \rceil + p$ gives the exact count whenever s is prime (or $s = 1$), since then all orbits contain s elements except those in the subfield $\mathbb{F}_p$. But even for composite s this bound is very accurate, apart from a few small values of q, because the vast majority of elements of $\mathbb{F}_q$ do not lie in any proper subfield and consequently almost all orbits have s elements. For example, using GAP [7] we find that for $\Psi_{2^{30}} = 35,792,568$, the bound of $35,791,397$ is approximately 99.9967% of the exact value, while the bound of $29,484,565,267,122,446$ is approximately 99.99999984% of $\Psi_{29^{16}} = 29,484,565,316,813,125$.

8. Markoff equivalence in the character variety

We saw in section 2 that for any group H, the action of $\mathrm{Aut}(F_2)$ on $\mathcal{G}_2(H)$ is generated by the action of three involutions r, s and t. It will be convenient to use a fourth element of $\mathrm{Aut}(F_2)$:

4. $m = tr$, which acts on $\mathcal{G}_2(H)$ by $m(A, B) = (A, AB)$.

As an automorphism of the set $H \times H$, the order of m is the least common multiple of the orders of the elements of H.

For any field F, the elements r, s, t, and m act on the set of F-triples as follows:

1) $r(\alpha, \beta, \gamma) = (\alpha, \gamma, \beta)$
2) $s(\alpha, \beta, \gamma) = (\beta, \alpha, \gamma)$
3) $t(\alpha, \beta, \gamma) = (\alpha, \beta, \alpha\beta - \gamma)$
4) $m(\alpha, \beta, \gamma) = (\alpha, \gamma, \alpha\gamma - \beta)$

Specializing to $H = \mathrm{SL}(2, F)$ for some field F, this action is induced from the action on $H \times H$ via Tr, that is, if (A, B) is a G_i-pair with $\mathrm{Tr}(A, B) = (\alpha, \beta, \gamma)$, then $r \circ \mathrm{Tr} = \mathrm{Tr} \circ r$, $s \circ \mathrm{Tr} = \mathrm{Tr} \circ s$, and $t \circ \mathrm{Tr} = \mathrm{Tr} \circ t$. For r and s this is obvious, and for t it is simply the identity $\mathrm{tr}(A^{-1}B) = \mathrm{tr}(A)\,\mathrm{tr}(B) - \mathrm{tr}(AB)$. We call the equivalence relation on F^3 generated by r, s, and t *Markoff equivalence.*

Since $\mathrm{Tr} \circ \mu = \mathrm{Tr}$ for any inner automorphism μ of F_2, the $\mathrm{Aut}(F_2)$-action on F^3 induces an action of the extended modular group $\mathrm{Aut}(F_2)/\mathrm{Inn}(F_2) = \mathrm{GL}(2, \mathbb{Z})$ on F^3. Since the element $-I$ of $\mathrm{GL}(2, \mathbb{Z})$ is represented by the automorphism that sends x_i to x_i^{-1} for both basis elements of F_2, it also acts trivially on F^3 and there is an induced action of $\mathrm{PGL}(2, \mathbb{Z})$. Thus Markoff equivalence in F^3 coincides with the orbits of this action of the extended modular group $\mathrm{PGL}(2, \mathbb{Z})$ that was used for $F = \mathbb{R}$ in [10].

Theorem 4.3 shows immediately that $\mathrm{Tr}\colon \mathcal{N} \to \mathcal{M}$ is a bijection when $p = 2$ and is (≤ 2)-to-1 when $p > 2$, where the latter terminology means that the preimage of each element of $\mathcal{M}$ contains at most two elements of $\mathcal{N}$. Moreover, we will prove in proposition 10.5 below that the Markoff class of (α, β, γ) has only one preimage Nielsen class whenever $2 - Q(\alpha, \beta, \gamma)$ is not a square in $\mathbb{F}_q$.

For the case $q \equiv 1 \bmod 4$, we saw that the two Nielsen-inequivalent generating pairs found at the end of the argument in section 6 differ by conjugation, so their Fricke traces are identical. Under the Fricke trace map $\mathcal{N} \to \mathcal{M}$, their Nielsen classes go to the same Markoff class, and consequently Conjecture B implies Conjecture C.

9. Numerical calculations

Using GAP [7], we have verified the conjectures for all $q \leq 101$. The source files for that work are available at [17]. In this section we will describe the methodology used.

From section 1, Conjecture C implies the other conjectures when q is even, and when q is odd, Conjectures C and B$'$ together imply the other conjectures.

The first step is to identify the essential characters. Denote by $P(A, B)$ the subgroup of $\mathrm{PSL}(2, q)$ generated by $\{A, B\}$. Clearly (A, B) generates a proper subgroup of $\mathrm{SL}(2, q)$ if and only if $P(A, B)$ is a proper subgroup. The cases for which this occurs were described in theorem 5.2, and can be identified from $\mathrm{Tr}(A, B) = (\alpha, \beta, \gamma)$ as follows.

1) By theorem 4.2, $P(A, B)$ is affine (which includes all cases when it is cyclic) if and only if $Q(\alpha, \beta, \gamma) = 2$.
2) Since a matrix has order 2 in $\mathrm{PSL}(2, q)$ if and only if its trace is 0, $P(A, B)$ is dihedral if and only if at least two of α, β, and γ are 0.
3) The cases when $P(A, B)$ is one of the exceptional subgroups A_4, S_4, or A_5 can be characterized by conditions on $\mathrm{Tr}(A, B) = (\alpha, \beta, \gamma)$ which are stated and verified in [16]. It is A_4 exactly when $\alpha, \beta, \gamma \in \{0, \pm 1\}$ and $Q(\alpha, \beta, \gamma) = 0$, and is S_4 exactly when $\alpha, \beta, \gamma \in \{0, \pm 1, \pm\sqrt{2}\}$, where $\sqrt{2}$ denotes a root of $x^2 - 2$, and $Q(\alpha, \beta, \gamma) = 1$. For A_5 the conditions are quite a bit more complicated, so we do not detail them here.
4) When all three of α, β, and γ lie in a proper subfield $\mathbb{F}_{p^r}$, either $P(A, B)$ is affine or $P(A, B)$ is isomorphic to a subgroup of $\mathrm{PSL}(2, p^r)$. For by theorem 4.1, there must be a pair (A', B') of elements of $\mathrm{PSL}(2, p^r)$ such that $\mathrm{Tr}(A', B') = (\alpha, \beta, \gamma)$. If $Q(\alpha, \beta, \gamma) = 2$, then $P(A, B)$ is affine, and if not, then by theorem 4.3, the subgroups $P(A, B)$ and $P(A', B')$ are isomorphic.

The remaining proper subgroups of $\mathrm{SL}(2, q)$ will be a bit of a nuisance. They are included in case (d) in theorem 5.2, and are described explicitly in (6.18) of [24] (also on p. 28 of [15]) as follows. Assume that q is odd, and observe that there is an element of $\mathbb{F}_q - \mathbb{F}_{p^r}$ whose square lies in $\mathbb{F}_{p^r}$ if and only if $\mathbb{F}_{p^{2r}} \subseteq \mathbb{F}_q$, in which case all such elements lie in $\mathbb{F}_{p^{2r}}$. If π is such an element, and $x \in \mathbb{F}_q - \mathbb{F}_{p^r}$, then $x^2 \in \mathbb{F}_{p^r}$ if and only if $x = \pi\delta$ for some some $\delta \in \mathbb{F}_{p^r}$. Write d_π for the matrix $\begin{pmatrix} \pi & 0 \\ 0 & \pi^{-1} \end{pmatrix}$. Then $\mathrm{SL}(2,p^r)$ has index 2 in the subgroup $\widehat{\mathrm{SL}}(2,p^r) = \langle \mathrm{SL}(2,p^r), d_\pi \rangle$. All that is going on is that $d_\pi^2 \in \mathrm{SL}(2,p^r)$ and d_π normalizes $\mathrm{SL}(2,p^r)$, so any element in $\widehat{\mathrm{SL}}(2,p^r) = \langle \mathrm{SL}(2,p^r), d_\pi \rangle$ can be written as $d_\pi^\epsilon A$ where $A \in \mathrm{SL}(2,p^r)$ and ϵ is 0 or 1. Note that the trace of $d_\pi^\epsilon A$ is of the form $\pi^\epsilon \delta$ for some $\delta \in \mathbb{F}_{p^r}$. Consequently, if (A,B) is a pair of elements of $\widehat{\mathrm{SL}}(2,p^r)$, then $\mathrm{Tr}(A,B) = (\pi^{\epsilon_1}\alpha, \pi^{\epsilon_2}\beta, \pi^{\epsilon_3}\gamma)$ with $\alpha, \beta, \gamma \in \mathbb{F}_{p^r}$ and $\epsilon_1 + \epsilon_2 + \epsilon_3 \equiv 0 \bmod 2$. Equivalently, $\mathrm{Tr}^2(A)$, $\mathrm{Tr}^2(B)$, $\mathrm{Tr}^2(AB)$, and $\mathrm{Tr}(A)\,\mathrm{Tr}(B)\,\mathrm{Tr}(AB)$ all lie in $\mathbb{F}_{p^r}$. We are prepared for the following variation on theorem 4.1:

Proposition 9.1. *Let $(\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma})$ be a triple of elements of $\mathbb{F}_q$ such that $\widehat{\alpha}^{\,2}$, $\widehat{\beta}^{\,2}$, $\widehat{\gamma}^{\,2}$, and $\widehat{\alpha}\,\widehat{\beta}\,\widehat{\gamma}$ all lie in a proper subfield $\mathbb{F}_{p^r}$ of $\mathbb{F}_q$. Then either $\widehat{\alpha}$, $\widehat{\beta}$, and $\widehat{\gamma}$ all lie in $\mathbb{F}_{p^r}$, or $\mathbb{F}_{p^{2r}} \subseteq \mathbb{F}_q$ and there exists a pair $(A,B) \in \widehat{\mathrm{SL}}(2,p^r)$ such that $\mathrm{Tr}(A,B) = (\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma})$.*

Proof. Our argument is a very slight modification of the proof of theorem 4.1, given as Theorem 1 in [15]. If all three coordinates lie in $\mathbb{F}_{p^r}$, in particular if q is even, then there is nothing to prove. Of the remaining possibilities, we need only consider the case when the triple is of the form $(\pi\alpha, \pi\beta, \gamma)$, with $\alpha, \beta, \gamma \in \mathbb{F}_{p^r}$, since the other cases can then be achieved by applying Nielsen moves to a pair (A,B) that we will obtain for this case.

We will seek a pair of the form

$$(A,B) = \left(\begin{pmatrix} 0 & \pi \\ -\pi^{-1} & \pi^{-1}(\pi^2\alpha) \end{pmatrix}, \begin{pmatrix} \pi x & \pi y \\ \pi^{-1}z & \pi^{-1}w \end{pmatrix} \right)$$

for which

$$\begin{aligned} \pi x + \pi^{-1}w &= \mathrm{tr}(B) = \pi\beta \\ z - y + \alpha w &= \mathrm{tr}(AB) = \gamma \\ xw - yz &= \det(B) = 1 \ . \end{aligned}$$

Eliminating x and then y from these equations reduces them to the condition

$$1 + z^2 + \pi^{-2}w^2 + \alpha zw + \beta(-1)w + \gamma(-1)z = 0 \ .$$

If (X,Y,Z) is a solution of

$$C(X,Y,Z) = X^2 + Y^2 + \pi^{-2}Z^2 + \alpha YZ + \beta XZ + \gamma XY = 0$$

with $X \neq 0$, then putting $(-1,z,w) = (-1,-Y/X,-Z/X)$ satisfies the condition. Since every quadratic form over a finite field has nonzero solutions, we need only consider the case of a nonzero solution of the form $(0,Y_0,Z_0)$. A line through this point and not tangent to the conic $C(X,Y,Z) = 0$ will intersect the conic in a point with $X \neq 0$, giving the desired solution. Such a line exists unless $(0, Y_0, Z_0)$ is a singular point of the conic. But the singularity condition is $0 = \gamma Y_0 + \beta Z_0$, $0 = 2Y_0 + \alpha Z_0$, and $0 = \alpha Y_0 + 2\pi^{-2} Z_0$, and the latter two equations imply that $\pi^2 = (2\alpha^{-1})^2$, in contradiction to the fact that $\pi \notin \mathbb{F}_{p^r}$. □

Equipped with proposition 9.1, we now consider the remaining kind of proper subgroup.

5) If p is odd and α^2, β^2, γ^2, and $\alpha\beta\gamma$ lie in a proper subfield $\mathbb{F}_{p^r}$ of $\mathbb{F}_q$, but at least one of α, β, and γ does not lie in $\mathbb{F}_{p^r}$, then either $P(A, B)$ is affine or $P(A, B)$ is isomorphic to a subgroup of the image of $\widehat{\mathrm{SL}}(2, p^r)$ in $\mathrm{PSL}(2, q)$. For by proposition 9.1, there exists a pair (A', B') of elements of $\widehat{\mathrm{SL}}(2, p^r)$ such that $\mathrm{Tr}(A', B') = (\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma})$. If $Q(\alpha, \beta, \gamma) = 2$, then $P(A, B)$ is affine, and if not, then by theorem 4.3 the subgroups $P(A, B)$ and $P(A', B')$ are isomorphic.

Conditions 1) through 5) make it easy to check when a triple in $\mathbb{F}_q^3$ is essential. In particular, a computational proof of theorem 5.1 for the cases when $q \leq 11$ can be carried out simply by removing from $\mathbb{F}_q^3$ all the triples satisfying one of the five conditions, then finding the values that Q assumes on the remaining subset.

Here is how our program verifies Conjecture C for $q \leq 101$. For a fixed q and a fixed value ℓ other than 2, it finds the essential triples of $\mathbb{F}_q^3$ with Q-value equal to ℓ, and forms singleton lists each containing one of these triples. Then it combines any two of these lists whenever an element of one is equivalent to an element of the other under one of r, s, or t. In all cases it finishes with only a single list for each of the possible trace invariants listed in theorem 5.1.

To verify Conjecture B$'$ by carrying out the same procedure at the level of generating pairs requires a great deal more computing capacity, and on typical desktop machines such as our is only feasible for $q \leq 9$. To overcome this, we use a different approach, which assumes that Conjecture C has already been verified for the value of q. We may assume that q is odd, since all conjectures follow from Conjecture C when q is even. Assuming that q satisfies Conjecture C, we exploit the fact (theorem 4.3) that each nonsingular triple is the trace of only two conjugacy classes of pairs. For the Q-value -2, there is nothing to prove since Conjecture C together with corollary 6.2 show that there are exactly two Nielsen classes with trace -2. For the other values of Q, we seek two pairs that have the same Fricke trace— a triple with this particular Q-value— that are Nielsen equivalent but not conjugate. Finding such pairs shows that there is only one Nielsen class mapping to the Markoff class of that triple. Since Conjecture C tells us there is only one Markoff class with the given Q-value, this establishes Conjecture B$'$ for that Q-value.

Here is the actual algorithm. Fixing a q with $13 \leq q$ and an $\ell \in \mathbb{F}_q - \{2, -2\}$, consider the graph with vertices the set of generating pairs with Q-value ℓ, with edges labeled by r running from each (A, B) to $r(A, B)$, and similar edges labeled s and t. The program chooses a pair (A_0, B_0) for which $Q(A_0, B_0) = \ell$, then takes random walks in the graph, starting from (A_0, B_0). Traveling along an edge corresponds to applying a basic Nielsen move. If some walk returns to a pair with Fricke trace $\operatorname{Tr}(A_0, B_0)$, but not conjugate to (A_0, B_0), then the two conjugacy classes of pairs with that Fricke trace are Nielsen equivalent. Corollary 6.2 shows that such a walk cannot exist when $q \equiv 1 \bmod 4$ and $\operatorname{tr}([A, B]) = -2$, but in all other cases we found many such walks. In the cases when $2-\ell$ was not a square, lemma 10.6 below shows that the Nielsen equivalent pairs (A_0, B_0) and (A_0^{-1}, B_0^{-1}) are nonconjugate, and the program immediately found very short walks taking (A_0, B_0) to a conjugate of (A_0^{-1}, B_0^{-1}). When $2 - \ell$ was not a square, the shortest walks varied considerably in length, with a few cases as short as one step, and most cases in the range of 10 to 50 steps. The longest walks needed were around 200 steps, with repeated program runs sometimes giving other walks of length in the low 100's for those cases, but with these "difficult" cases always among the longest walks needed.

10. Parabolic, elliptic, and hyperbolic elements

In our remaining work, we will need more information about the fields $\mathbb{F}_q$ and their elements. To set notation, we denote by u a generator of $C_{q-1} = \mathbb{F}_q - \{0\}$. In the software for the computer-assisted calculations that we discussed in section 9, we used for u the primitive element denoted by $Z(q)$ in the GAP computer algebra system [7]. For the unique quadratic extension $\mathbb{F}_{q^2}$ of $\mathbb{F}_q$, the group of nonzero elements is $C_{q^2-1} = \mathbb{F}_{q^2} - \{0\}$ and is generated by $Z(q^2)$. The element u is $Z(q^2)^{q+1}$, and we denote by v the element $Z(q^2)^{q-1}$. The latter generates a subgroup $C_{q+1} \subset C_{q^2-1}$, and its powers are exactly the elements of $\mathbb{F}_{q^2}$ that satisfy $x^{q+1} = 1$, that is, the elements of norm 1. For p odd, $C_{q-1} \cap C_{q+1} = C_2$, generated by $u^{(q-1)/2} = v^{(q+1)/2} = -1$. For $p = 2$, $C_{q-1} \cap C_{q+1} = \{1\}$. The subgroup of C_{q^2-1} generated by $C_{q-1} \cup C_{q+1}$ is the set of squares, so is all of C_{q^2-1} when $p = 2$ and has index 2 when $p > 2$.

An element α of $\mathbb{F}_q$ is called *elliptic, parabolic* or *hyperbolic* according as the equation $\lambda^2 - \alpha\lambda + 1 = 0$ has zero, one, or two distinct roots in $\mathbb{F}_q$. An element of $\mathbb{F}_q$ is hyperbolic if and only if it can be written as $u^i + u^{-i}$ with $u^i \neq \pm 1$. The elements $v^j + v^{-j}$ lie in $\mathbb{F}_q$, and for $v^j \neq \pm 1$ are exactly the elliptic elements. We denote the sets of elliptic and hyperbolic elements of a field under discussion by E and H respectively. If q is even, then E contains $\frac{1}{2}q$ elements, H contains $\frac{1}{2}(q-2)$ hyperbolic elements, and 0 is the unique parabolic element. If q is odd, then E contains $\frac{1}{2}(q-1)$ elliptic elements, H contains $\frac{1}{2}(q-3)$ hyperbolic elements, and 2 and -2 are the parabolic elements.

The following lemma gives a simple criterion to identify the type of an element in fields of odd characteristic. By $\mathbb{F}_q^2$ we denote the set of elements of $\mathbb{F}_q$ that are squares.

Lemma 10.1. *Let $\alpha \in \mathbb{F}_q$, with $p > 2$.*

1) *α is parabolic if and only if $\alpha^2 - 4 = 0$.*
2) *α is hyperbolic if and only if $\alpha^2 - 4 \in \mathbb{F}_q^2 - \{0\}$.*
3) *α is elliptic if and only if $\alpha^2 - 4 \notin \mathbb{F}_q^2$.*

Proof. It suffices to prove the only if direction of all three statements. The parabolic case is immediate. If α is hyperbolic, then writing $\alpha = u^i + u^{-i}$ gives $\alpha^2 - 4 = (u^i - u^{-i})^2 \in \mathbb{F}_q^2 - \{0\}$. If α is elliptic, then $\alpha = v^j + v^{-j}$ produces $\alpha^2 - 4 = (v^j - v^{-j})^2$. If this is the square of an element in $\mathbb{F}_q$, then $v^j - v^{-j} \in \mathbb{F}_q$, so $2v^j \in \mathbb{F}_q$. Since $p > 2$, this implies that $v^j \in \mathbb{F}_q$, a contradiction. □

Recall that if $r = p^m$ is a prime power and $n > 1$, the trace map $\mathrm{Tr}\colon \mathbb{F}_{r^n} \to \mathbb{F}_r$ is the $\mathbb{F}_r$-linear transformation defined by $\mathrm{Tr}(x) = x + x^{p^m} + x^{p^{2m}} + \cdots + x^{p^{(n-1)m}}$. The trace map has many well-known properties which can be found in any text on finite fields, such as [13]. In particular, $\mathrm{Tr}(xy)$ defines a nondegenerate symmetric $\mathbb{F}_r$-valued bilinear form on $\mathbb{F}_{r^n} \times \mathbb{F}_{r^n}$, making $\mathbb{F}_{r^n}$ an inner product space over $\mathbb{F}_r$. Consequently, every subspace W determines an orthogonal subspace $W^\perp$ of complementary dimension. Since the characteristic is nonzero, W and $W^\perp$ may have nontrivial intersection. The kernel of Tr is exactly $\mathbb{F}_r^\perp$.

In characteristic 2, the trace map leads to an elegant and useful description of the hyperbolic and elliptic elements. Let q be even, and for $S \subseteq \mathbb{F}_q - \{0\}$, denote by S^{-1} the set consisting of the inverses of the elements of S. Since 0 is the unique parabolic element when q is even, both H^{-1} and E^{-1} are defined.

Lemma 10.2. *Assume that $\mathbb{F}_q$ has characteristic 2, and let $\mathrm{Tr}\colon \mathbb{F}_q \to \mathbb{F}_2$ be the trace map to $\mathbb{F}_2$. Then $\mathrm{Tr}^{-1}(0) = H^{-1} \cup \{0\}$ and $\mathrm{Tr}^{-1}(1) = E^{-1}$.*

Proof. Let $\alpha \in H$. Then $x^2 + \alpha x + 1 = 0$ for some $x \in \mathbb{F}_q$, so $(x\alpha^{-1})^2 + (x\alpha^{-1}) + \alpha^{-2} = 0$. Since the traces of conjugate elements are equal, applying Tr to this equation gives $\mathrm{Tr}(\alpha^{-2}) = 0$ and hence $\mathrm{Tr}(\alpha^{-1}) = 0$. Conversely, if $\alpha \neq 0$ and $\mathrm{Tr}(\alpha^{-1})$ and hence $\mathrm{Tr}(\alpha^{-2})$ are 0, then we can write $\alpha^{-2} = \beta^2 + \beta$ for some element β of $\mathbb{F}_q$ (see for example Theorem 2.25 of [13]), and multiplying by α^2 shows that $x^2 + \alpha x + 1 = 0$ has a solution. We conclude that $H^{-1} = \mathrm{Tr}^{-1}(0) - \{0\}$. Since $\mathbb{F}_q = \{0\} \cup H \cup E$ and E has $q/2$ elements, it follows that $E^{-1} = \mathrm{Tr}^{-1}(1)$. □

For odd characteristic, we do not know a result analogous to lemma 10.2.

Lemma 10.2 has the following consequences.

Corollary 10.3. *Let q be even and let W denote the subspace of $\mathbb{F}_q$ spanned by the subset $\{\kappa_1, \ldots, \kappa_r\}$ of $\mathbb{F}_q$. Then $(\kappa_1 H \cap \cdots \cap \kappa_r H)^{-1} \cup \{0\} = W^\perp$.*

Proof. Using lemma 10.2, $x \in (\kappa H)^{-1} \cup \{0\} = \kappa^{-1}(H^{-1} \cup \{0\})$ if and only if $\mathrm{Tr}(\kappa x) = 0$, that is, $x \in (\kappa \mathbb{F}_2)^\perp$. So $(\cap \kappa_i H)^{-1} \cup \{0\} = \cap(\kappa_i \mathbb{F}_2)^\perp = W^\perp$. □

Corollary 10.4. *Let q be even and suppose that the elements $\kappa_1, \ldots, \kappa_n$ of $\mathbb{F}_q$ are linearly independent over $\mathbb{F}_2$. Then*

(a) *$\cap_{j=1}^n \kappa_j E$ contains $q/2^n$ elements.*

(b) *For $1 \leq m \leq n$, $(\cap_{i=1}^m \kappa_i H) \cap (\cap_{j=m+1}^n \kappa_j E)$ contains $q/2^n - 1$ elements.*

Proof. Corollary 10.3 shows that $\cap_{i=1}^n \kappa_i H$ has $q/2^n - 1$ elements, which is part (b) for $m = n$. Now, let $K_i = \kappa_i H \cup \{0\}$, so that $\cap_{i=1}^n K_i = \{0\} \cup (\cap_{i=1}^n \kappa_i H)$ has $1/2^n$ elements. Write L_i for $\mathbb{F}_q - K_i$, which is $\kappa_i E$. Inducting on $n - m$, we have $1 + |(\cap_{i=1}^m \kappa_i H) \cap (\cap_{j=m+1}^n \kappa_j E)| = |(\cap_{i=1}^m K_i) \cap (\cap_{j=m+1}^n L_j)| = |(\cap_{i=1}^m K_i) \cap (\cap_{j=m+2}^n L_j)| - |(\cap_{i=1}^{m+1} K_i) \cap (\cap_{j=m+2}^n L_j)| = q/2^{n-1} - q/2^n = q/2^n$, establishing the rest of (b). Part (a) follows from induction and part (b) since $|(\cap_{j=1}^n L_j)| = |(\cap_{j=2}^n L_j)| - |K_1 \cap (\cap_{j=2}^n L_i)| = q/2^{n-1} - q/2^n$. □

The type of an element of G_i— that is, parabolic, elliptic, or hyperbolic— is defined to be the type of its trace.

Consider an $\mathbb{F}_q^3$-triple (α, β, γ) with α either hyperbolic or elliptic. Write $\alpha = x + x^{-1}$, with x of the form u^i or v^j according as α is hyperbolic or elliptic. Since $x \neq x^{-1}$, there is a unique pair (a, d) of elements of $\mathbb{F}_{q^2}$ such that $a + d = \beta$ and $ax + dx^{-1} = \gamma$. Explicitly, $a = (\gamma - \beta x^{-1})/(x - x^{-1})$ and $d = (\beta x - \gamma)/(x - x^{-1})$. In the hyperbolic case, a and d lie in $\mathbb{F}_q$, and we may select any b and c in $\mathbb{F}_q$ with $bc = ad - 1$. In the elliptic case, we compute that $d = a^q$. Since $ad - 1 \in \mathbb{F}_q$, we may select any $b \in \mathbb{F}_{q^2}$ with $b^{q+1} = ad - 1$, and we may choose $c = b^q$ for this b. Thus (α, β, γ) is the trace of a G_i-pair— a G_0-pair when A is hyperbolic, and a G_1-pair when it is elliptic— of the form

$$(A, B) = \left(\begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix}, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) ,$$

with a and d uniquely determined by $\mathrm{Tr}(A, B)$ and the choice of which of x and x^{-1} is considered to be x. We say that such pairs are *in normal form.* We note that this gives a computational proof of theorem 4.1, but more importantly, for a pair in normal form, it is straightforward to compute the following formula, which we call the *Fundamental Equation:*

$$Q(\alpha, \beta, \gamma) = 2 - bc(\alpha^2 - 4)$$

The Fundamental Equation helps establish the following proposition that was mentioned in section 1.

Proposition 10.5. *Let (α, β, γ) be an essential character. If $2 - Q(\alpha, \beta, \gamma)$ is not a square in $\mathbb{F}_q$, then the Markoff class of (α, β, γ) is the Fricke trace of a unique Nielsen class.*

Since (A, B) is always Nielsen equivalent to (A^{-1}, B^{-1}), proposition 10.5 follows directly from theorem 4.3 and the following computational lemma.

Lemma 10.6. *Let (A, B) be a G_i-pair with $Q(A, B) \neq 2$. Then (A, B) is conjugate to (A^{-1}, B^{-1}) if and only if $2 - Q(A, B)$ is a square in $\mathbb{F}_q$.*

It seems interesting to compare lemma 10.6 with Lemma 3.4.5 of [10], which says that in $\mathrm{SL}(2,\mathbb{R})$, A and B are hyperbolic elements whose axes cross if and only if $\mathrm{tr}([A,B]) < 2$. Since $\mathrm{tr}([A,B]) = Q(A,B)$, this condition is equivalent (when $Q(A,B) \neq 0$) to $2 - Q(A,B)$ being a square in $\mathbb{R}$. Of course, the isometry of $\mathbb{H}^2$ that rotates through an angle of π fixing the intersection point of the axes conjugates (A,B) to (A^{-1},B^{-1}).

Proof of lemma 10.6. When q is even, $2 - \mathrm{Tr}(A,B)$ is always a square. On the other hand, $\mathrm{Tr}(A,B) = \mathrm{Tr}(A^{-1},B^{-1})$, so theorem 4.3 shows that (A,B) is conjugate to (A^{-1},B^{-1}), establishing the lemma. So we will assume that q is odd.

Assume first that at least one of A, B, or AB is not parabolic. By Nielsen moves we may assume that A is not parabolic. Conjugate to put (A,B) into normal form.

Suppose that A is hyperbolic. We have $\beta = a + d$ and $\gamma = \mathrm{tr}(AB) = ax + dx^{-1}$. If (A,B) is conjugate to (A^{-1},B^{-1}), then the conjugating matrix X must be of the form $\begin{pmatrix} 0 & s \\ s^{-1} & 0 \end{pmatrix}$, and $XBX^{-1} = B^{-1}$ forces $b = cs^2$. Conversely, the condition $b = cs^2$, which allows the conjugation to be carried out, is equivalent to $bc = (cs)^2$, that is, bc is a square in $\mathbb{F}_q$. By the Fundamental Equation, $bc = (2 - Q(A,B))/(\alpha^2 - 4)$. Since $\alpha^2 - 4 = (x - x^{-1})^2$, bc is a square in $\mathbb{F}_q$ if and only if $2 - Q(A,B)$ is a square.

Suppose now that A is elliptic, and $B = \begin{pmatrix} a & b \\ b^q & a^q \end{pmatrix}$. By lemma 10.1, $\alpha^2 - 4$ is not a square in $\mathbb{F}_q$, and the Fundamental Equation shows that $2 - Q(A,B) = b^{q+1}(\alpha^2 - 4)$, so we must show that (A,B) is conjugate to (A^{-1},B^{-1}) if and only if b^{q+1} is not the square of an element of $\mathbb{F}_q$.

The condition that $XAX^{-1} = A^{-1}$ in G_1 is equivalent to X being of the form $\begin{pmatrix} 0 & s \\ s^q & 0 \end{pmatrix}$ where $s^{q+1} = -1$, and then $XBX^{-1} = B^{-1}$ exactly when $s^{-2} = b^{q-1}$. So we must show that these two equations hold for some $s \in \mathbb{F}_{q^2}$ if and only if b^{q+1} is not a square in $\mathbb{F}_q$.

If they hold, then raising both sides of the second equation to the power $-(q+1)/2$ gives $s^{q+1} = (b^{q+1})^{(1-q)/2}$. If b^{q+1} were a square in $\mathbb{F}_q$, say $b^{q+1} = c^2$, then we would have $-1 = (c^2)^{(1-q)/2} = c^{1-q} = 1$.

Suppose that b^{q+1} is not a square in $\mathbb{F}_q$. Let $s = b^{(1-q)/2}$. Then $s^{q+1} = (s^2)^{(q+1)/2} = (b^{1-q})^{(q+1)/2} = (b^{q+1})^{(1-q)/2} = -1$.

We may now assume that A and B (and AB, although that is not needed here) are parabolic. We work in G_0, and by conjugating we may assume that A is of the form $\begin{pmatrix} \epsilon & x \\ 0 & \epsilon \end{pmatrix}$, with $\epsilon = \pm 1$ and $x \neq 0$. It is straightforward to check that A is conjugate to A^{-1} if and only if -1 is a square in $\mathbb{F}_q$, say, $r^2 = -1$. In this case, any matrix of the form $X = \begin{pmatrix} r & s \\ 0 & -r \end{pmatrix}$ conjugates A to

A^{-1}. Writing B as $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the condition that $XBX^{-1} = B^{-1}$ is equivalent to $r(d-a) = sc$. If $c \neq 0$, this can be satisfied, while if $c = 0$, then since B is parabolic we have $a = d$ and again it can be satisfied. So (A,B) is conjugate to (A^{-1}, B^{-1}) exactly when -1 is a square in $\mathbb{F}_q$. On the other hand, for $\gamma = \mathrm{tr}(AB)$ we have $Q(A,B) = 4+4+\gamma^2 \pm 4\gamma - 2 = 2 + (\gamma \pm 2)^2$, so $2 - Q(A,B)$ is a square if and only if -1 is a square. Again, the criterion holds. □

11. The geometry of the Q-levels

For each $\ell \in \mathbb{F}_q$, we denote $Q^{-1}(\ell)$ by Q_ℓ, and call it a *level surface* of Q, or a *Q-level.* We will examine intersections of the Q-levels with the "slices" of $\mathbb{F}_q^3$ having a fixed value for one of the coordinates. Since there are Nielsen automorphisms that interchange the coordinates, it is sufficient to focus initially on a single coordinate. The discussion is necessarily rather notation-intensive, so it may be helpful to read this section in tandem with the concrete examples presented in section 12.

For $\alpha \in \mathbb{F}_q$, define U_α to be the set of $\mathbb{F}_q$-triples whose first entry is equal to α, and for $\ell \in \mathbb{F}_q$, define $U_{\alpha,\ell} = Q_\ell \cap U_\alpha$. Denote by $\mathbb{D}$ the subgroup of $\mathrm{Aut}(F_2)$ generated by $\{t, m\}$ (where t and m were defined in sections 2 and 8 respectively). It is an infinite dihedral group, whose action on $\mathbb{F}_q^3$ preserves each $U_{\alpha,\ell}$. Any action of $\mathbb{D}$ on a finite set induces an effective action of a finite quotient of $\mathbb{D}$ that is a dihedral group (allowing the possibilities $D_2 = C_2 \times C_2$, $D_1 = C_2$, and $D_0 = \{1\}$).

Finally, we define another quantity that will play an important role in our work. For $\alpha, \ell \in \mathbb{F}_q$ with $\alpha \neq \pm 2$, define

$$k(\alpha,\ell) = 1 - (\ell - 2)(\alpha^2 - 4)^{-1} \ .$$

When the values of α and ℓ are fixed, as in the next proposition, we often just write k for $k(\alpha,\ell)$.

Proposition 11.1. *Let $\alpha \in \mathbb{F}_q$ be hyperbolic, and write $\alpha = x + x^{-1}$ for some $x \in \mathbb{F}_q$.*

1) *When $k \neq 0$, that is, when $\ell \neq \alpha^2 - 2$, $U_{\alpha,\ell}$ is a "hyperbola" with $q-1$ points. Explicitly, if $C_{\alpha,\ell}$ is the set of pairs $(a, k/a)$ with $a \in \mathbb{F}_q - \{0\}$, then sending $(a, k/a)$ to $(\alpha, a + k/a, ax + k/(ax))$ is a bijection from $C_{\alpha,\ell}$ to $U_{\alpha,\ell}$. In these $C_{\alpha,\ell}$-coordinates, the action of $\mathbb{D}$ on $U_{\alpha,\ell}$ becomes $m(a, k/a) = (ax, k/(ax))$ and $t(a, k/a) = (k/a, a)$.*
2) *When $k = 0$, that is, when $\ell = \alpha^2 - 2$, $U_{\alpha,\ell}$ is a "degenerate hyperbola" with $2q - 1$ points, consisting of the two straight lines $\gamma = x\beta$ and $\gamma = x^{-1}\beta$. The action of $\mathbb{D}$ fixes their intersection point $(\alpha, 0, 0)$, and on the other points it acts by $m(\alpha, \beta, x\beta) = (\alpha, x\beta, x^2\beta)$, $m(\alpha, \beta, x^{-1}\beta) = (\alpha, x^{-1}\beta, x^{-2}\beta)$, and $t(\alpha, \beta, x\beta) = (\alpha, \beta, x^{-1}\beta)$.*

Proof. Assume first that $k \neq 0$. Using the normal form and Fundamental Equation from section 10, we find that $U_{\alpha,\ell}$ consists of the $(\alpha, \beta, \gamma) =$

$(x+x^{-1}, a+d, ax+dx^{-1})$ such that $ad=k$ (with the ordered pair (a,d) uniquely determined by (α,β,γ) and the ordered pair (x,x^{-1})). The map $\phi\colon C_{\alpha,\ell}\to U_{\alpha,\ell}$ defined by $\phi(a,k/a)=(\alpha,a+k/a,ax+k/(ax))$ is easily seen to be a bijection. Using the formulas $t(\alpha,\beta,\gamma)=(\alpha,\beta,\alpha\beta-\gamma)$ and $m(\alpha,\beta,\gamma)=(\alpha,\gamma,\alpha\gamma-\beta)$ given near the beginning of section 8, one verifies that $t\,\phi(a,k/a)=\phi(k/a,a)$ and $m\,\phi(a,k/a)=\phi(ax,k/(ax))$. Changing the choice of which member of the pair $\{x,x^{-1}\}$ is considered to be x interchanges a and d, and the resulting formulas for the actions of m and t define the same effect.

When $k=0$, the equation $\ell=Q(\alpha,\beta,\gamma)$ works out to $0=\beta^2-\alpha\beta\gamma+\gamma^2=(\beta-x\gamma)(\beta-x^{-1}\gamma)$, giving the two intersecting straight lines for $U_{\alpha,\ell}$. Since $\alpha=x+x^{-1}$, the action of m on the line $\gamma=x\beta$ is

$$m(\alpha,\beta,x\beta)=(\alpha,x\beta,(\alpha x-1)\beta)=(\alpha,x\beta,x^2\beta)$$

and similarly for the line $\gamma=x^{-1}\beta$. The action of t is

$$t(\alpha,\beta,x\beta)=(\alpha,\beta,(\alpha-x)\beta)=(\alpha,\beta,x^{-1}\beta)\ .$$

□

Proposition 11.2. *Let $\alpha\in\mathbb{F}_q$ be elliptic, and write $\alpha=x+x^q$ with $x\in\mathbb{F}_{q^2}-\mathbb{F}_q$ and $x^{q+1}=1$.*

1) *When $k\neq 0$, that is, when $\ell\neq\alpha^2-2$, $U_{\alpha,\ell}$ is an "ellipse" with $q+1$ points. Explicitly, if $C_{\alpha,\ell}$ is the set of pairs (a,a^q) with $a\in\mathbb{F}_{q^2}$ and $a^{q+1}=k$, then sending (a,a^q) to $(\alpha,a+k/a,ax+k/(ax))$ is a bijection from $C_{\alpha,\ell}$ to $U_{\alpha,\ell}$. In these $C_{\alpha,\ell}$-coordinates, the action of $\mathbb{D}$ on $U_{\alpha,\ell}$ becomes $m(a,k/a)=(ax,k/(ax))$ and $t(a,k/a)=(k/a,a)$.*
2) *When $k=0$, that is, when $\ell=\alpha^2-2$, $U_{\alpha,\ell}$ is a "degenerate ellipse" consisting only of $(\alpha,0,0)$.*

Proof. For $k\neq 0$, calculating as in proposition 11.1 shows that $U_{\alpha,\ell}$ consists of the $(\alpha,\beta,\gamma)=(x+x^q,a+a^q,ax+(ax)^q)$ such that $a^{q+1}=k$. In $\mathbb{F}_{q^2}$ there are $q+1$ choices for a, and the map from $C_{\alpha,\ell}$ to $U_{\alpha,\ell}$ is again seen to be bijective, with the action as described. When $k=0$, the factorization $0=(\beta-x\gamma)(\beta-x^{-1}\gamma)$ has the unique solution $(\beta,\gamma)=(0,0)$ in $\mathbb{F}_q\times\mathbb{F}_q$. □

Proposition 11.3. *Let $\alpha\in\mathbb{F}_q$ be parabolic.*

1) *If $p>2$ and $\alpha=2\epsilon$, then $Q_{\alpha,\ell}$ is empty if $\ell-2$ is not a square, while if $\ell-2=s^2$, $Q_{\alpha,\ell}$ is the set of triples of the form $(2\epsilon,\beta,\epsilon\beta\pm s)$, which is a pair of disjoint lines if $\ell\neq 2$ and a single line if $\ell=2$. The action of m is $m(2\epsilon,\beta,\epsilon\beta\pm s)=(2\epsilon,\epsilon\beta\pm s,\epsilon(\epsilon\beta\pm s)\pm\epsilon s)$, so for $\ell\neq 2$, m preserves each line if $\epsilon=1$ and interchanges them if $\epsilon=-1$. The action of t is $t(2\epsilon,\beta,\epsilon\beta\pm s)=(2\epsilon,\beta,\epsilon\beta\mp s)$, so t interchanges the two lines. For $\ell=2$, t acts trivially, and m acts as an involution when $\epsilon=-1$ and trivially when $\epsilon=1$.*
2) *For $p=2$ and $\ell=s^2$, $Q_{\alpha,\ell}$ is the line consisting of the points of the form $(0,\beta,\beta+s)$. The action of t is trivial and the action of m is*

1	u	0	0
u	u	0	0
0	1	u	u
$\alpha=0$	0	u	1

1	0	0	u
u	0	u	0
0	u	0	0
$\alpha=u$	0	u	1

1	0	u	0
u	0	0	u
0	u	0	0
$\alpha=1$	0	u	1

FIGURE 1. Slices for $q = 3$.

$m(0, \beta, \beta + s) = (0, \beta + s, \beta)$, so m is an involution if $\ell \neq 0$ and acts trivially if $\ell = 0$.

Proof. For $p > 2$ and $\alpha = 2\epsilon$, the equation $\ell = Q(\alpha, \beta, \gamma)$ is $\ell - 2 = (\gamma - \epsilon\beta)^2$, so $Q_{\alpha,\ell}$ is empty when $\ell - 2$ is not square. When $\ell - 2 = s^2 \neq 0$, $Q_{\alpha,\ell}$ consists of the two disjoint lines $\gamma = \epsilon\beta \pm s$. We have $m(2\epsilon, \beta, \epsilon\beta \pm s) = (2\epsilon, \epsilon\beta \pm s, \epsilon(\epsilon\beta \pm s) \pm \epsilon s)$, so m preserves each line if $\epsilon = 1$ and interchanges them if $\epsilon = -1$. For t, we have $t(2\epsilon, \beta, \epsilon\beta \pm s) = (2\epsilon, \beta, \epsilon\beta \mp s)$, so t interchanges the two lines. When $\ell - 2 = 0$, we have $\gamma = \epsilon\beta$, and the remarks about the action are easily checked.

For $p = 2$ any ℓ can be written uniquely as s^2. We find $s^2 = Q(\alpha, \beta, \gamma) = (\gamma + \beta)^2$, which says that $\gamma = \beta + s$ and $Q_{\alpha,\ell}$ is a line, and the action works out as stated. □

12. EXAMPLES OF SLICES

Figure 1 shows the slices for $\mathbb{F}_3 = \{0, u, 1\}$. For each fixed value of α, the horizontal coordinate is β, the vertical coordinate is γ, and the (β, γ)-entry is $Q(\alpha, \beta, \gamma)$. The sixteen triples with $Q(\alpha, \beta, \gamma) = 0$ are the single $\mathrm{Aut}(F_2)$-orbit of traces of generators. The element 0 is the unique elliptic element of $\mathbb{F}_3$, indeed $0 = Z(9)^2 + Z(9)^{-2}$, where $Z(9)$ is the multiplicative generator of $\mathbb{F}_9 - \{0\}$ provided by GAP.

The slice for $\alpha = 0$ is as described in proposition 11.2. We have $k = 1 - (\ell - 2)(0^2 - 4) = \ell - 1$, so $k = 0$ occurs for $U_{0,1}$, giving the degenerate ellipse $(0, 0, 0)$, while $U_{0,0}$ and $U_{0,u}$ are ellipses each containing four points. The elements $\alpha = u$ and $\alpha = 1$ are parabolic, and their slices are as described in proposition 11.3.

Figure 2 shows the slices for $\mathbb{F}_4 = \{0, u, u^2, 1\}$. The elements u and u^2 are elliptic. There is one parabolic element, 0, and one hyperbolic element 1. In the slice for $\alpha = 1$, the degenerate hyperbola is $U_{1,1}$. The level surface U_1 consists of the three inessential characters $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$, and one $\mathrm{Aut}(F_2)$-orbit of length 18.

Finally, figures 3 and 4 show the slices for $\mathbb{F}_8$ for $\alpha = u$ and $\alpha = u^3$. Since u is elliptic, U_u contains one degenerate ellipse $(u, 0, 0)$ and seven nondegenerate ellipses each containing nine points. The element u^3 is hyperbolic, and U_{u^3} contains one degenerate hyperbola of fifteen points and seven nondegenerate hyperbolas of seven points each. In each figure we have indicated the degenerate conics in boldface.

1	1	u	u^2	0
u^2	u	1	0	u^2
u	u^2	0	1	u
0	0	u^2	u	1
$\alpha=0$	0	u	u^2	1

1	u	u	1	1
u^2	1	0	0	1
u	0	u	0	u
0	u^2	0	1	u
$\alpha=u$	0	u	u^2	1

1	u^2	1	u^2	1
u^2	0	0	u^2	u^2
u	1	0	0	1
0	u	1	0	u^2
$\alpha=u^2$	0	u	u^2	1

1	0	1	1	0
u^2	u^2	1	u^2	1
u	u	u	1	1
0	1	u	u^2	0
$\alpha=1$	0	u	u^2	1

FIGURE 2. Slices for $q = 4$.

1	u^6	u^6	0	u^4	0	u^3	u^3	u^4
u^6	u^3	u^6	1	u^6	u^5	u^5	1	u^3
u^5	u^5	u	u^3	u^4	u^4	u	u^5	u^3
u^4	u^4	u^5	1	1	0	u^4	u^5	0
u^3	1	u	u	u^6	1	u^4	u^6	u^4
u^2	u	0	u^3	u	1	u^3	1	0
u	0	u^5	0	u	u^5	u	u^6	u^6
0	$\mathbf{u^2}$	0	u	1	u^4	u^5	u^3	u^6
$\alpha=u$	0	u	u^2	u^3	u^4	u^5	u^6	1

FIGURE 3. Slice for $q = 8$ and $\alpha = u$.

1	u^2	u^4	$\mathbf{u^6}$	u^2	u^5	$\mathbf{u^6}$	u^5	u^4
u^6	u	$\mathbf{u^6}$	u	0	$\mathbf{u^6}$	0	u^5	u^5
u^5	u^4	u^4	u^3	$\mathbf{u^6}$	u^3	0	0	$\mathbf{u^6}$
u^4	u^5	1	$\mathbf{u^6}$	1	u^3	u^3	$\mathbf{u^6}$	u^5
u^3	0	$\mathbf{u^6}$	u^2	1	1	$\mathbf{u^6}$	0	u^2
u^2	u^3	u	u^2	u^2	$\mathbf{u^6}$	u^3	u	$\mathbf{u^6}$
u	1	u	u	$\mathbf{u^6}$	1	u^4	$\mathbf{u^6}$	u^4
0	$\mathbf{u^6}$	1	u^3	0	u^5	u^4	u	u^2
$\alpha=u^3$	0	u	u^2	u^3	u^4	u^5	u^6	1

FIGURE 4. Slice for $q = 8$ and $\alpha = u^3$.

13. THE CASE OF EVEN CHARACTERISTIC

Throughout this section we assume that $p = 2$, so that $q = 2^s$ for some s. For x in $\mathbb{F}_q$, we denote the unique square root $x^{q/2}$ of x by $\sqrt{x}$. The expression $k = 1-(\ell-2)(\alpha^2-4)^{-1}$ from section 11 becomes $k = 1+\ell\alpha^{-2}$, so is 0 exactly when $\alpha = \sqrt{\ell}$. We introduce the notation

$$\kappa = 1+\sqrt{\ell}\alpha^{-1}$$

for the square root of k.

In characteristic 2, the set of values of the second coordinate β dthat appear in triples in $U_{\alpha,\ell}$ can be described in a somewhat simpler way. Denote this set by $\beta(U_{\alpha,\ell})$. Suppose first that $\alpha = u^d+u^{-d}$ is hyperbolic, so that proposition 11.1 applies. When $\kappa = 0$, it shows that $\beta(U_{\alpha,\ell}) = \mathbb{F}_q$, while for $\kappa \neq 0$ it gives

$$\beta(U_{\alpha,\ell}) = \{a+\kappa^2/a \mid a \in \mathbb{F}_q^*\} = \{\kappa(a/\kappa+1/(a/\kappa)) \mid a \in \mathbb{F}_q^*\} = \kappa\,(H\cup\{0\})\,,$$

where $\mathbb{F}_q^*$ denotes $\mathbb{F}_q - \{0\}$ and as usual H denotes the set of hyperbolic elements of $\mathbb{F}_q$. Similarly, applying proposition 11.2 when $\alpha = v^d+v^{-d}$ is elliptic shows that for $\kappa = 0$, $\beta(U_{\alpha,\ell}) = \{0\}$ and for $\kappa \neq 0$

$$\beta(U_{\alpha,\ell}) = \kappa\,(E\cup\{0\})$$

although one detail should be mentioned. Since $\kappa \in \mathbb{F}_q$, we have $(a/\kappa)^{q+1} = k/\kappa^{q+1} = \kappa^2/\kappa^2 = 1$, so $a/\kappa+1/(a/\kappa)$ is indeed elliptic.

We will need to understand the β-coordinates of the $\mathbb{D}$-orbits of $U_{\alpha,\ell}$, where $\mathbb{D}$ is the subgroup of $\mathrm{Aut}(F_2)$ generated by $\{t,m\}$, as discussed near the beginning of section 11. To simplify notation, write $h(i)$ for the element u^i+u^{-i}, noting that $h(i) = h(-i) = h(i\pm(q-1))$ and $h(0) = 0$. Now, fix a hyperbolic element $\alpha = h(d)$. Denote $\gcd(d,q-1)$ by d_0, so that the order of u^d in $\mathbb{F}_q^*$ is $(q-1)/d_0$, which we denote by d_1.

For $n \in \mathbb{Z}$, put

$$H_{d_0}(n) = \{h(d_0 i+n) \mid i \in \mathbb{Z}\} = \{h(di+n) \mid i \in \mathbb{Z}\}\ .$$

We have $H_{d_0}(n) = H_{d_0}(-n)$. For odd j, $H_{d_0}(\frac{d_0-j}{2}) = H_{d_0}(\frac{d_0+j}{2})$, since $h(d_0 i+\frac{d_0-j}{2}) = h(d_0(-i)-\frac{d_0-j}{2}) = h(d_0(-i-1)+\frac{d_0+j}{2})$. Also, $H_d(0)$ contains $(d_1+1)/2$ points, while for $1 \leq n \leq (d_0-1)/2$, $H_d(n)$ contains d_1 points, and $\{H_{d_0}(0), H_{d_0}(1), \ldots, H_{d_0}(\frac{d_0-1}{2})\}$ is a partition of $H\cup\{0\}$.

Let $\kappa h(i) \in \beta(U_{\alpha,\ell})$. From the description of the action of m given in proposition 11.1, the β-coordinate of the image under m of a point with β-coordinate $\kappa\,h(i) = \kappa u^i+\kappa^2(1/(\kappa u^i))$ is $\kappa u^{d+i}+\kappa^2(1/(\kappa u^{d+i})) = \kappa\,h(d+i)$. Since t fixes the β-coordinate of each point, the set of β-coordinates of an m-orbit is the same set as the β-coordinates of the $\mathbb{D}$-orbit that contains it. Thus we have obtained the following description.

Proposition 13.1. *Assume that $p = 2$ and let $\alpha = h(d)$ be a hyperbolic element of $\mathbb{F}_q$. Put $d_0 = \gcd(d,q-1)$ and $d_1 = (q-1)/d_0$. For the $q-1$ values of ℓ with $\kappa \neq 0$, $\beta(U_{\alpha,\ell}) = \kappa(H\cup\{0\})$. Moreover,*

1) *The sets of β-coordinates of the $\mathbb{D}$-orbits of $U_{\alpha,\ell}$ are $\kappa H_{d_0}(0)$, $\kappa H_{d_0}(1)$, $\ldots$, $\kappa H_{d_0}((d_0-1)/2)$.*

2) *The set $\kappa H_{d_0}(0)$ contains $(d_1+1)/2$ points, and each $\kappa H_{d_0}(n)$ for $1 \le n \le (d_0-1)/2$ contains d_1 points.*

By a very similar argument, putting $e(d) = v^d + v^{-d}$ and

$$E_{d_0}(n) = \{e(d_0 i + n) \mid i \in \mathbb{Z}\}$$

we have

Proposition 13.2. *Assume that $p = 2$ and let $\alpha = e(d)$ be an elliptic element of $\mathbb{F}_q$. Put $d_0 = \gcd(d, q+1)$ and $d_1 = (q+1)/d_0$. For the $q-1$ values of ℓ with $\kappa \neq 0$, $\beta(U_{\alpha,\ell}) = \kappa(E \cup \{0\})$. Moreover,*

1) *The sets of β-coordinates of the $\mathbb{D}$-orbits of $U_{\alpha,\ell}$ are $\kappa E_{d_0}(0)$, $\kappa E_{d_0}(1)$, $\ldots$, $\kappa E_{d_0}((d_0-1)/2)$.*
2) *The set $\kappa E_{d_0}(0)$ contains $(d_1+1)/2$ points, and each $\kappa E_{d_0}(j)$ for $1 \le j \le (d_0-1)/2$ contains d_1 points.*

A hyperbolic (respectively, elliptic) element α is called *transitive* exactly when $\alpha = h(d)$ (respectively, $\alpha = e(d)$) with $\gcd(d, q-1) = 1$ (respectively, $\gcd(d, q+1) = 1$). We remark that a matrix $A \in \mathrm{SL}(2,q)$ has order $q-1$ or $q+1$ in $\mathrm{PSL}(2,q)$ if and only if $\mathrm{tr}(A)$ is transitive, but we will not need this fact.

Write $V_{\alpha,\ell}$ for $U_{\alpha,\ell} - \{(\alpha,0,0)\}$. We have $V_{\alpha,\ell} = U_{\alpha,\ell}$ except when $\kappa = 0$, and propositions 13.1 and 13.2 show that for $\kappa = 0$, $\beta(V_{\alpha,\ell})$ is $\mathbb{F}_q^*$ when α is hyperbolic and is empty when α is elliptic. Propositions 13.1 and 13.2 show that for all transitive α, $\mathbb{D}$ acts transitively on $V_{\alpha,\ell}$.

For two subsets $X, Y \subset \mathbb{F}_q^3$, we write $X \sim_M Y$ when every element of X is Markoff equivalent to every element of Y.

Proposition 13.3. *Suppose that $q-1$ or $q+1$ is prime. Then for each $\ell \neq 0$, there is a Markoff equivalence class that contains $V_{\alpha,\ell}$ for every transitive element α of $\mathbb{F}_q$.*

Proof. Fix ℓ. For any transitive element α, the elements of $V_{\alpha,\ell}$ lie in a single $\mathbb{D}$-orbit, so are Markoff equivalent.

Suppose for now that $q-1$ is prime, so that every hyperbolic element is transitive. We fix a hyperbolic element α_1, and will prove that $V_{\alpha_1,\ell} \sim_M V_{\alpha_2,\ell}$ for any transitive element α_2.

Write κ_i for $1 + \sqrt{\ell}\,\alpha_i$. If α_2 is elliptic, then we may assume that $\kappa_2 \neq 0$, since if $\kappa_2 = 0$ then $V_{\alpha_2,\ell}$ is empty.

It is sufficient to prove that $\beta(V_{\alpha_1,\ell}) \cap \beta(V_{\alpha_2,\ell})$ contains a transitive element α_3. For then, each $V_{\alpha_i,\ell}$ contains a point of the form $(\alpha_i, \alpha_3, \gamma_i)$ in $V_{\alpha_i,\ell}$. The points $(\alpha_3, \alpha_i, \gamma_i)$ lie in $V_{\alpha_3,\ell}$, so are equivalent. Since $(\alpha_i, \alpha_3, \gamma_i) \sim_M (\alpha_3, \alpha_i, \gamma_i)$, $V_{\alpha_1,\ell} \sim_M V_{\alpha_3,\ell} \sim_M V_{\alpha_2,\ell}$.

If one of the κ_i is 0, say κ_1, then $\beta(V_{\alpha_1,\ell}) = \mathbb{F}_q$ and $V_{\alpha_1,\ell}$ contains a point of the form $(\alpha_1, \alpha_2, \gamma)$, hence $\alpha_1 \in \beta(V_{\alpha_1,\ell}) \cap \beta(V_{\alpha_2,\ell})$. So we may assume that both κ_i are nonzero. If α_2 is hyperbolic, then corollary 10.3 shows that

$\kappa_1 H \cap \kappa_2 H \cap H$ is nonempty, and if α_2 is elliptic, it shows that $\kappa_1 H \cap \kappa_2 E \cap H$ is nonempty, giving the desired transitive element in $\beta(V_{\alpha_1,\ell}) \cap \beta(V_{\alpha_2,\ell})$.

If $q+1$ is prime, then all elliptics are transitive, and the argument is similar, with α_1 selected to be elliptic with $\kappa_1 \neq 0$. □

The proof of theorem 13.5 will use the following easy observation about cyclic groups of prime order:

Lemma 13.4. *Let $S \subseteq C_P$ where P is prime, and let $x \neq 1$. If $xS = S$, then S is either empty or $S = C_P$. Suppose that S is a nonempty, proper subset and $xS \subset S \cup \{y\}$. Then S is of the form $\{x^{-1}y, x^{-2}y, \ldots, x^{-n}y\}$ for some n.*

Proof. The first statement is immediate since x must be a generator of C_P. For the second statement, consider the subset $S' = \{x^{-1}y, x^{-2}y, \ldots, x^{-n}y\}$, where $n+1$ is the minimal value for which $x^{-(n+1)}y \notin S$. Then $x(S - S') = S - S'$ so $S' = S$. □

We can now prove Conjecture C for a restricted set of q.

Theorem 13.5. *Let $q = 2^s$ and suppose that one of $q+1$ or $q-1$ is prime and the other is 3 times a prime. Then there are exactly $q-1$ Markoff classes of essential characters, classified by their Q-values.*

The only case of theorem 13.5 for $q+1$ prime is $s = 4$. For numbers of the form $2^{2k+1}+1$ are always divisible by 3, while if $s = 2k$ then $q-1$ factors as $(2^k-1)(2^k+1)$ and is of the form $3p_1$ only when $k = 2$. (Actually, the proof of theorem 13.5 adapts, in a rather degenerate manner, to $q = 4$.) For $q-1$ prime, theorem 13.5 applies when $s \in \{3, 5, 7, 13, 17, 19, 31, 61, 127\}$, and perhaps for other values as well. It might apply to infinitely many cases, of course it is a well-known open problem even to determine whether there are infinitely many primes among the Mersenne numbers $2^s - 1$.

Proof of theorem 13.5. From theorem 5.1 we know that the Q-values of the Markoff classes are all of $\mathbb{F}_q - \{0\}$.

Fix $\ell \in \mathbb{F}_q - \{0\}$ and suppose first that $q-1$ is prime and $q+1 = 3p_1$ with p_1 prime. Since $q-1$ is prime, all hyperbolic elements are transitive, and proposition 13.3 shows that there is a Markoff class $M(\ell)$ containing $V_{\alpha,\ell}$ for all transitive elements α. For transitive α, the only time a triple in any of the slices $U_{\alpha,\ell}$ is not contained in $V_{\alpha,\ell}$ is when $\kappa = 0$ and the triple is $(\sqrt{\ell}, 0, 0)$. This triple is the trace of pairs which generate dihedral subgroups and hence is not essential. So it remains to show that $M(\ell)$ contains the essential characters in $U_{\alpha,\ell}$ for all nontransitive α with $\kappa \neq 0$. Since every essential triple has at least two nonzero coordinates, we may further assume that α is nonzero.

Using the notation of proposition 13.2, the set of nontransitive elements is $E_3(0) \cup E_{p_1}(0)$. Since $1 = v^{p_1} + v^{-p_1}$, $E_{p_1}(0) = \{0, 1\}$. Therefore the set of nontransitive elements is $E_3(0) \cup \{1\}$. It contains $(p_1+3)/2$ elements, unless $q = 8$, in which case $E_3(0) = \{0, 1\}$ contains $(p_1+1)/2$ elements.

Fix an essential character $(\alpha, \beta, \gamma) \in U_{\alpha,\ell}$ with α nontransitive. Necessarily its corresponding $\kappa \neq 0$, since as we have seen $U_{\sqrt{\ell},\ell}$ contains no essential character.

Since (α, β, γ) is essential, at least two of its coordinates are nonzero, and by Markoff equivalence we may assume that they include α and β. Also, since the coordinates do not all lie in a proper subfield, we may assume that $\alpha \neq 1$. That is, $\alpha \in E_3(0) - E_{p_1}(0)$ and $\beta \neq 0$.

By proposition 13.2, the set of β-coordinates for the $\mathbb{D}$-orbit of (α, β, γ) is $\kappa E_3(0) \cup \kappa E_3(1)$, with each of these two sets being the set of β-coordinates for some $\mathbb{D}$-orbit of $U_{\alpha,\ell}$. If $\beta \in \kappa E_3(1)$, then since $\kappa E_3(1)$ contains p_1 values, and there are only $(p_1 + 3)/2$ nontransitive elements (or $(p_1 + 1)/2$, if $q = 8$), (α, β, γ) is Markoff equivalent to some $(\alpha, \beta', \gamma')$ with β' transitive, so $(\alpha, \beta, \gamma) \in M(\ell)$. So we may assume that $\beta \in \kappa E_3(0) \subset E_3(0) \cup \{1\}$ and $\kappa E_3(0) \subseteq E_3(0) \cup \{1\}$. If $q = 8$ this is impossible since $\kappa \neq 1$. Otherwise, lemma 13.4 shows that $E_3(0)$ is of the form $\{0, \kappa^{-1}, \kappa^{-2}, \ldots, \kappa^{-(p_1-1)/2}\}$. But this is a contradiction, since $E_3(0)$ is closed under the Frobenius automorphism.

Suppose now that $q + 1$ is prime and $q - 1 = 3p_1$. We proceed as before, interchanging the roles of elliptic and hyperbolic. The set of nontransitive elements is $H_3(0) \cup \{1\}$, and contains $(p_1 + 1)/2$ elements, so we may examine an essential triple (α, β, γ) with α a nontransitive hyperbolic other than 1, and $\beta \neq 0$. If $\alpha = \sqrt{\ell}$ so that $\kappa = 0$, then according to proposition 11.1 the β-sets of the $\mathbb{D}$-orbits in $U_{\alpha,\ell}$ are of the form $\{\beta, u^3\beta, u^6\beta, \ldots, u^{q-4}\beta\}$, so contain p_1 elements. At least one of these must be transitive, so we may assume that $\kappa \neq 0$. The remainder of the proof is then analogous to the case when $q - 1$ was prime. □

As explained in section 1, Conjecture C implies all of the conjectures for these values of q, so we the following is a consequence of theorem 13.5:

Corollary 13.6. *Suppose that one of $q - 1$ or $q + 1$ is prime and the other is 3 times a prime. Then there are exactly $q - 1$ Nielsen classes of generating pairs, classified by their trace invariants.*

As far as extending these results to more values of q, we have little idea how to make headway for odd q, as the even case relied on several major simplifications that do not seem to have analogues in the odd case. For the even case, a roadblock to extending our method is our inability to make some usable statement about the effect of the element m on β-coordinates, specifically about the values of i and j that are obtained when elements of the form $\kappa(a+\kappa/a)$ are rewritten in the form u^i+u^{-i} or v^j+v^{-j}. For the first main step, proposition 13.3, one can weaken the hypotheses a bit by more careful use of corollary 10.3 in the proof. Since this gives some additional information about Markoff equivalence and its proof is fairly short, we give one such result here. It applies to many values of q such as 2^{11}, for which $2^{11} - 1 = 23 \cdot 89$.

Proposition 13.7. *Suppose that no more than* $3q/16-2$ *hyperbolic elements are nontransitive. Then for each* $\ell \neq 0$*, there is a Markoff equivalence class* $M(\ell)$ *such that for every transitive element* α *of* $\mathbb{F}_q$*,* $U_{\alpha,\ell} \subset M(\ell)$*.*

Proof. Arguing as in the proof of proposition 13.3, it suffices to reach a contradiction assuming that $\beta(V_{\alpha_1,\ell})\cap\beta(V_{\alpha_2,\ell})$ contains no transitive hyperbolic element. Let α_3 be a transitive hyperbolic element such that κ_3 is not in the $\mathbb{F}_2$-subspace of $\mathbb{F}_q$ spanned by κ_1 and κ_2.

Assume first that α_2 is hyperbolic. By corollary 10.4, $\kappa_1 H \cap \kappa_2 H \cap H$ has exactly $q/8-1$ elements, which are assumed nontransitive, each $\kappa_i H \cap \kappa_3 H \cap H$ has $q/8-1$ elements, and $\kappa_1 H \cap \kappa_2 H \cap \kappa_3 H \cap H$ has $q/16-1$ elements. For each i, the $q/16$ elements of $\kappa_i H \cap \kappa_3 H \cap H$ that are not in $\kappa_1 H \cap \kappa_2 H \cap \kappa_3 H \cap H$ cannot all be nontransitive, since then the set of nontransitive hyperbolic elements would contain the $q/8-1$ elements of $\kappa_1 H \cap \kappa_2 H \cap H$, plus these additional $q/16$.

When α_2 is elliptic, the argument is exactly the same except that $\kappa_2 E$ is used in place of $\kappa_2 H$. □

Of course, there is an analogous result assuming that there is a sufficient density of transitive elliptic elements.

Finally, as remarked in section 1, the key role played by transitive elements, while possibly an artifact of our approach to theorem 13.5, does give some reason for caution. The cases in which the conjectures are known to hold, that is, $q \leq 101$ and the cases of theorem 13.5, all have rather high densities of transitive elements, so at least in this sense they are not representative of the general case. For very large q, this density can be arbitrarily close to 0.

14. The case of $\mathrm{PSL}(2,q)$

In this section, we will adapt the conjectures to the case of $\mathrm{PSL}(2,q)$. Since $\mathrm{PSL}(2,q) = \mathrm{SL}(2,q)$ when q is even, we will assume throughout this section that q is odd.

For now, we consider coefficients in an arbitrary field F. We have already noted that even when (A,B) represents a pair in $\mathrm{PSL}(2,F)$, so that A and B are only defined up to sign, the trace of $[A,B]$ is well-defined. To make the Fricke trace $\mathrm{Tr}(A,B)$ well-defined on Nielsen classes in $\mathcal{G}_2(\mathrm{PSL}(2,F))$, it is sufficient to extend Markoff equivalence by adding the additional involutions $(\alpha,\beta,\gamma) \to (-\alpha,\beta,-\gamma)$ and $(\alpha,\beta,\gamma) \to (\alpha,-\beta,-\gamma)$. This extends the $\mathrm{PGL}(2,\mathbb{Z})$-action on F^3 to an action of $(C_2 \times C_2) \circ \mathrm{PGL}(2,\mathbb{Z})$ on F^3. We call the resulting relation *weak Markoff equivalence.*

We will use $\overline{\mathcal{N}}$ and $\overline{\mathcal{M}}$ to denote the Nielsen classes for $\mathrm{PSL}(2,F)$ and the set of weak Markoff classes in F^3, so that there are natural surjections $\mathcal{N} \to \overline{\mathcal{N}}$ and $\mathcal{M} \to \overline{\mathcal{M}}$. For finite F, at least, the following conjecture seems reasonable:

Conjecture P (No essential difference between SL and PSL)**.** *$\mathcal{N} \to \overline{\mathcal{N}}$ and $\mathcal{M} \to \overline{\mathcal{M}}$ are bijections.*

Specializing to the case $F = \mathbb{F}_q$, Conjecture A implies that $\mathcal{N} \to \overline{\mathcal{N}}$ is bijective, since we have a composition $\mathcal{N} \to \overline{\mathcal{N}} \to \{\text{Higman invariants}\}$ with the first map surjective. Similarly, Conjecture C implies that $\mathcal{M} \to \overline{\mathcal{M}}$ is bijective. Thus Conjecture A implies Conjecture P. Similarly, Conjecture W implies that $\mathcal{T} \to \overline{\mathcal{T}}$ is a bijection, where $\overline{\mathcal{T}}$ denotes the T-systems of $\mathrm{PSL}(2,q)$.

On the other hand, versions of the conjectures for $\mathrm{PSL}(2,q)$ imply weak forms of the conjectures, by means of the following observation:

Proposition 14.1. *The natural maps $\mathcal{N} \to \overline{\mathcal{N}}$ and $\mathcal{M} \to \overline{\mathcal{M}}$ are (≤ 2)-to-1.*

Proof. Suppose that (A,B) and (A',B') in $\mathcal{G}_2(\mathrm{SL}(2,q))$ are Nielsen equivalent as elements of $\mathcal{G}_2(\mathrm{PSL}(2,q))$. A sequence of Nielsen moves changing (A',B') to (A,B) up to signs changes (A',B') to one of (A,B), $(-A,B)$, $(A,-B)$, or $(-A,-B)$.

If A has odd order k, then $(-A)^k = -I$, so $(-A,B) \sim (-A,-B)$, where $\sim$ indicates Nielsen equivalence. On the other hand, if A has even order $2k$, then $A^k = -I$ so $(A,B) \sim (A,-B)$. By the same reasoning applied to B, either $(A,-B) \sim (-A,-B)$ or $(A,B) \sim (-A,B)$. Each of the four possible combinations lead to (at least) three of (A,B), $(-A,B)$, $(A,-B)$, or $(-A,-B)$ being Nielsen equivalent, showing that $\mathcal{N} \to \overline{\mathcal{N}}$ is (≤ 2)-to-1.

Since the four equivalent $\mathbb{F}_q$-triples (α,β,γ), $(-\alpha,\beta,-\gamma)$, $(\alpha,-\beta,-\gamma)$, and $(-\alpha,-\beta,\gamma)$ are the Fricke traces of (A,B), $(-A,B)$, $(A,-B)$, and $(-A,-B)$, the previous argument shows that they lie in at most two Markoff classes. It follows that $\mathcal{M} \to \overline{\mathcal{M}}$ is (≤ 2)-to-1. □

Consider, for example, the following conjecture:

Conjecture PA (Higman invariant classifies Nielsen classes)**.** *Two generating pairs (A,B) and (A',B') of $\mathrm{PSL}(2,q)$ are Nielsen equivalent if and only if $[A,B]$ is conjugate to $[A',B']$ or to $[B',A']$ in $\mathrm{SL}(2,q)$.*

Conjecture PA is that the latter map in $\mathcal{N} \to \overline{\mathcal{N}} \to \{\text{Higman invariants}\}$ is a bijection, and proposition 14.1 then implies that $\mathcal{N} \to \{\text{Higman invariants}\}$ is (≤ 2)-to-1. The patterns are similar for Conjectures C and W.

The argument in section 6 showing the equivalence of Conjectures A and B applies to the analogous projective versions.

As for Conjecture B′, the Fricke trace $\mathrm{Tr}\colon \mathcal{G}_2(\mathrm{SL}(2,q)) \to \mathbb{F}_q^3$ carries each Nielsen equivalence class onto an entire Markoff class, since the $\mathrm{Aut}(F_2)$-action on $\mathbb{F}_q^3$ that produces Markoff equivalence is induced by the $\mathrm{Aut}(F_2)$-action on $\mathcal{G}_2(\mathrm{SL}(2,q))$ that produces Nielsen equivalence. So two different elements of $\mathcal{M}$ have the same image in $\overline{\mathcal{M}}$ if and only if two different elements of $\mathcal{N}$ have the same image in $\overline{\mathcal{N}}$. This shows that Conjecture B′ implies the analogous statement for $\mathrm{Tr}\colon \overline{\mathcal{N}} \to \overline{\mathcal{M}}$.

15. Almost free actions

We call an action of a (nontrivial) finite group G on a compact surface F *fixed-point-transitive* if G acts transitively on $\cup_{1\neq g\in G}\operatorname{Fix}(g)$. If in addition, the fixed-point set is nonempty, we say that the action is *almost free.* These may be regarded as the actions that are as close as possible to being free. For closed F, they correspond exactly to the free actions on bounded surfaces which act transitively with nontrivial stabilizers on the set of boundary components.

From now on, we assume implicitly that all surfaces are closed and orientable, and that all actions preserve orientation. Fix an almost free action of G on a surface F of genus g. The stabilizers of the fixed points are conjugate cyclic subgroups of G of some order $n \geq 2$. The quotient orbifold $\mathcal{O}$ has one order-n cone point, and underlying manifold $|\mathcal{O}|$ an orientable surface of some genus g_0. The orbifold fundamental group $\pi_1^{\mathrm{orb}}(\mathcal{O})$, that is, the set of lifts of elements of G to the universal cover of F, is an extension $1 \to \pi_1(F) \to \pi_1^{\mathrm{orb}}(\mathcal{O}) \to G \to 1$; explicitly, $\pi_1^{\mathrm{orb}}(\mathcal{O})$ can be given by a presentation of the form $\langle a_1, b_1, \ldots, a_{g_0}, b_{g_0} \mid \Big(\prod_{i=1}^{g_0}[a_i,b_i]\Big)^n = 1\rangle$.

Since $\pi_1(F)$ is torsionfree, the image of $\prod_{i=1}^{g_0}[a_i,b_i]$ in G must have order n. Therefore G is nonabelian.

Using orbifold Euler characteristic, we have $\chi_{\mathrm{orb}}(\mathcal{O}) = \chi(|\mathcal{O}|) + \frac{1}{n} - 1$ and $\chi(F) = |G|\,\chi_{\mathrm{orb}}(\mathcal{O})$, giving $g_0 = \frac{g-1}{|G|} + \frac{n+1}{2n}$. This yields immediately:

1) $g \geq 2$.
2) If $g \leq |G|$, then $g_0 = 1$ and $g = 1 + \frac{n-1}{2n}|G|$.

When $g \leq |G|$ we call the almost free action *large.* Our main result shows that large almost free actions are classified by Nielsen equivalence classes. As usual, $\mathcal{G}_2(G)$ will denote the set of generating pairs of G. Recall that two actions $\varphi_i \colon G \to F_i$, $i = 1, 2$, are *weakly equivalent* if there are a diffeomorphism $h \colon F_1 \to F_2$ and an automorphism $\alpha \colon G \to G$ so that $\varphi_1(g) = h^{-1}\varphi_2(\alpha(g))\,h$ for all $g \in G$, and are called *equivalent* if this can be achieved with α the identity automorphism.

Theorem 15.1. *Let G be a nonabelian group. Then the equivalence classes (respectively, weak equivalence classes) of large almost free actions of G on closed orientable surfaces correspond bijectively to the Nielsen classes (respectively, T-systems) of $\mathcal{G}_2(G)$. For the action on a surface F corresponding to the equivalence class of a generating pair $\{A, B\}$, the order of the stabilizers of the fixed points equals the order n of $[A, B]$ in G. Moreover, $n = \frac{|G|}{|G|+\chi(F)}$, or equivalently the genus of F is $1 + \frac{n-1}{2n}\,|G|$.*

In particular, when $\mathcal{G}_2(G)$ is empty, the theorem implies that G does not act almost freely on any F.

Proof of theorem 15.1. Denote by $\mathcal{O}_n$ the orbifold with underlying manifold the torus, and one cone point of order n, so that $\pi_1^{\mathrm{orb}}(\mathcal{O}_n) = \langle A, B \mid [A, B]^n =$

$1\rangle$. We have observed that the quotient orbifold $\mathcal{O}$ of any large almost free action of G is homeomorphic to some $\mathcal{O}_n$. The image of the standard generating pair (A,B) of $\pi_1^{\mathrm{orb}}(\mathcal{O}_n)$ in G is an element of $\mathcal{G}_2(G)$. The Nielsen class of this element is independent of the choice of identification of $\mathcal{O}$ with $\mathcal{O}_n$. To see this, we note first that any other identification differs from this previous one by an orbifold homeomorphism, which induces an automorphism of $\pi_1^{\mathrm{orb}}(\mathcal{O}_n)$. It is known that all automorphisms of $\pi_1^{\mathrm{orb}}(\mathcal{O}_n)$ are induced by automorphisms of the free group on $\{A,B\}$, so different choices produce Nielsen equivalent pairs for (A,B) in $\pi_1^{\mathrm{orb}}(\mathcal{O}_n)$, and hence produce Nielsen equivalent elements of $\mathcal{G}_2(G)$.

There are two fine points here. The first concerns the fact that automorphisms of $\pi_1^{\mathrm{orb}}(\mathcal{O}_n)$ are induced by automorphisms of F_2, or equivalently that any two generating pairs of $\pi_1^{\mathrm{orb}}(\mathcal{O}_n)$ differ by Nielsen moves. The general problem of classifying generating pairs for two-generator Fuchsian groups is quite difficult. As far as we know, the case of the groups $\langle A,B \mid [A,B]^n=1\rangle$, which we use here, appears first as Lemma 3 of [21], but much of that paper is incorrect. The later [22] gives a full algebraic solution of the classification and a discussion of the troubled history of the problem, in particular the case we need appears as Theorem 1(2.1) there. See also [8].

The second fine point is that the automorphism of $\pi_1^{orb}(\mathcal{O}_n)$ preserves $\pi_1(F)$ and induces an inner automorphism on $G=\pi_1^{orb}(\mathcal{O}_n)/\pi_1(F)$, since it is induced by a homeomorphism that is induced by an equivalence of the actions. When it is only a weak equivalence, an outer automorphism of G may be induced, so the generating pairs of $G=\pi_1^{orb}(\mathcal{O}_n)/\pi_1(F)$ may map to elements of $\mathcal{G}_2(G)$ that differ by an automorphism of G, and the pairs in $\mathcal{G}_2(G)$ need only be T-equivalent. Thus the assignment is indeed well-defined as a function from equivalence classes (respectively, weak equivalence classes) of actions to Nielsen classes (T-systems) in $\mathcal{G}_2(G)$.

To see that this assignment is injective, suppose that two actions on surfaces F_1 and F_2 are assigned Nielsen equivalent elements $(\overline{A_1},\overline{B_1})$ and $(\overline{A_2},\overline{B_2})$ of $\mathcal{G}_2(G)$. As we have seen, we may assume that they are actions by orbifold covering transformations of covers of the same quotient orbifold $\mathcal{O}_n$. A sequence of Nielsen moves in G taking $(\overline{A_1},\overline{B_1})$ to $(\overline{A_2},\overline{B_2})$ lifts to a sequence of Nielsen moves in $\mathcal{G}_2(\pi_1^{\mathrm{orb}}(\mathcal{O}_n))$ taking (A,B) to some (A',B') which also projects to $(\overline{A_2},\overline{B_2})$. These Nielsen moves are induced by orbifold homeomorphisms of $\mathcal{O}_n$ (the product replacements by Dehn twists, the inversions and the interchange of generators by orientation-reversing involutions). So there is an orbifold homeomorphism of $\mathcal{O}_n$ that induces an isomorphism from the extension $1\to\pi_1(F_1)\to\pi_1^{\mathrm{orb}}(\mathcal{O}_n)\to G\to 1$ to the extension $1\to\pi_1(F_2)\to\pi_1^{\mathrm{orb}}(\mathcal{O}_n)\to G\to 1$. This isomorphism of extensions is the identity on G, so the lifted homeomorphism is an equivalence of the actions of G as covering transformations. For weak equivalence, the argument is the same except that the isomorphism of extensions may induce a nontrivial isomorphism of the quotient groups G, so the lifted homeomorphism may be only a weak equivalence of the actions of G as covering transformations.

Finally, a generating pair $\{\overline{A}, \overline{B}\}$ of G for which $[\overline{A}, \overline{B}]$ has order n defines a surjective homomorphism $\pi_1(\mathcal{O}_n) \to G$. In a one-relator group, every torsion element is conjugate to a power of the relator (see for example Theorem IV.5.2 of [14]). Therefore the kernel of the homomorphism is torsionfree, so the covering orbifold of $\mathcal{O}_n$ corresponding to the kernel is a manifold on which G acts almost freely. The Euler characteristic calculations shown above give its genus to be $1 + \frac{n-1}{2n}\,|G|$. □

Using these results, we can easily classify the orientation-preserving large fixed-point transitive actions on the surface of genus 3. The formula $g = 1 + \frac{n-1}{2n}|G|$ becomes $|G| = \frac{4n}{n-1}$, and only the values $n = 2$, 3, and 5 yield integers for $|G|$, giving respectively 8, 6, and 5. The latter cannot occur since G is nonabelian, so G must be one of D_3, D_4, or Q_8. For D_3 there is a generating pair with $[A, B]$ of order 3, and for D_4 and Q_8 there are generating pairs with $[A, B]$ of order 2, so fixed-point transitive actions of these three groups exist. In all three cases, it is not difficult to check that $\mathcal{G}_2(G)$ has only one Nielsen class (the dihedral cases appear as theorem 14 of [18]), so these actions are unique up to equivalence.

Another interesting example is $G = A_5$. Regard it as $\mathrm{PSL}(2, 4)$. Theorem 15.1 shows that the orientation-preserving large fixed-point-transitive actions of A_5 are classified by the Nielsen equivalence classes of generating pairs of $\mathrm{PSL}(2, 4)$. These are are well-known; in the framework of our theory we know they are classified by their trace invariants 1, u, and u^2 in $\mathbb{F}_4$. The element $1 = u + u^{-1}$ is hyperbolic, and is the trace of matrices of order $n = 3$, the order of u, so the Nielsen class with trace invariant 1 corresponds to a fixed-point transitive action of A_5 on the surface of genus $g = 1+\frac{3-1}{2\cdot 3}\cdot 60 = 21$. The elements $u = v^2+v^{-2}$ and $u^2 = v+v^{-1}$ are elliptic (where $v = Z(8)^3$ as in section 10) and are the traces of matrices of order $n = 5$, the orders of v and v^2. These correspond to two equivalence classes of fixed-point transitive actions on the surface of genus $g = 1+\frac{5-1}{2\cdot 5}\cdot 60 = 25$, which are weakly equivalent.

References

1. Kenneth P. Bogart, An obvious proof of Burnside's lemma, *Amer. Math. Monthly* 98 (1991), 927–928.
2. B. H. Bowditch, A proof of McShane's identity via Markoff triples, *Bull. London Math. Soc.* 28 (1996), 73–78.
3. B. H. Bowditch, Markoff triples and quasi-Fuchsian groups, *Proc. London Math. Soc.* (3) 77 (1998), 697–736.
4. A. Costa and D. McCullough, Orientation-reversing free actions on handlebodies, *J. Pure Appl. Alg.* 204 (2006) 155-169.
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. GAP: Groups, Algorithms, and Programming, available at the St. Andrews GAP website `http://turnbull.mcs.st-and.ac.uk/~gap/` .
8. J. Gilman, Two-generator discrete subgroups of $\mathrm{PSL}(2,R)$, *Mem. Amer. Math. Soc.* 117 (1995), no. 561, x+204 pp.
9. H. Glover and D. Sjerve, The genus of $\mathrm{PSL}_2(q)$, *J. Reine Angew. Math.* 380 (1987), 59–86.
10. W. Goldman, The modular group action on real SL(2)-characters of a one-holed torus, *Geom. Topol.* 7 (2003), 443–486.
11. R. Guralnick and I. Pak, On a question of B. H. Neumann, *Proc. Amer. Math. Soc.* 131 (2003), 2021–2025.
12. Loo-Keng Hua, On the automorphisms of the symplectic group over any field, *Ann. of Math.* (2) 49, (1948), 739–759.
13. R. Lidl and H. Niederreiter, *Introduction to finite fields and their applications,* Cambridge University Press, Cambridge, (1986).
14. R. C. Lyndon and P. E. Schupp, *Combinatorial group theory,* Springer-Verlag, 1977.
15. 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.
16. D. McCullough, Exceptional subgroups of $\mathrm{SL}(2,F)$, preprint available at `www.math.ou.edu/~dmccullough/`
17. D. McCullough, GAP programs for *The action of the modular group on characters of* $\mathrm{SL}(2,q)$*-representations of* F_2*,* available at `www.math.ou.edu/~dmccullough/`
18. D. McCullough and M. Wanderley, Free actions on handlebodies, *J. Pure Appl. Algebra.* 181 (2003), 85–104.
19. P. Neumann, A lemma which is not Burnside's, *Math. Sci.* 4 (1979), 133–141
20. J. Nielsen, Die Isomorphismengruppe der allgemeinen unendlichen Gruppe mit zwei Erzeugenden, *Math. Ann.* 78 (1918), 385–397.
21. N. Purzitsky and G. Rosenberger, Two generator Fuchsian groups of genus one, *Math. Z.* 128 (1972), 245-251, with correction *Math. Z.* 132 (1973), 261–262.
22. G. Rosenberger, All generating pairs of all two-generator Fuchsian groups, *Arch. Math.* 46 (1986), 198–204.
23. O. Schrier and B. L. van der Waerden, Die Automorphismen der projektiven Gruppen, *Abh. Math. Sem. Univ. Hamburg* 6 (1928), 303–322.
24. M. Suzuki, *Group Theory, vol. I,* Springer-Verlag, 1982.

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

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`