JOURNAL OF ALGEBRA **176**, 139–181 (1995)

On the Probability That a Group Element is p-Singular

I. M. Isaacs*

Department of Mathematics, University of Wisconsin, Madison, Wisconsin 53706

and

W. M. Kantor† and N. Spaltenstein

Department of Mathematics, University of Oregon, Eugene, Oregon 97403

Communicated by George Glauberman

Received July 29, 1992

1. INTRODUCTION

Suppose that G is a permutation group of degree n and let p be a prime divisor of $|G|$. In computational group theory it is a natural and important problem to find an element of G of order p. A polynomial-time (but impractical) algorithm for this is given in [Ka]. In practice, an element of the desired type is obtained by "randomly" choosing elements of G and computing their orders. After a few tries, and with some luck, a p-singular element (i.e., one of order divisible by p) frequently turns up. The purpose of this article is to make it clear just how well this procedure can be expected to work.

MAIN THEOREM. *Let G be a permutation group of degree n whose order is divisible by a prime p. The following then hold.*

(a) *The probability that an element of G has order divisible by p is at least $1/n$.*

(b) *Equality occurs above if and only if either G is sharply 2-transitive with n a power of p or G is the full symmetric group S_n with $n = p \geq 5$.*

It is easy to see in the two situations described in (b) that the probability that a random element is p-singular is exactly $1/n$. The proof of the rest

*Research partially supported by the NSF.
†Research partially supported by the NSF and NSA.

0021-8693/95 $12.00

of the theorem begins with a reduction to the case in which G is "almost simple" and is a primitive permutation group. (A group G is *almost simple* if its socle is a nonabelian simple group; in other words, if G is contained between a nonabelian simple group and its automorphism group.) The proof is then completed by an analysis of the various simple groups, using the classification theorem. The reduction occurs in Section 2, the alternating and sporadic groups are considered in Section 3, and the analysis of the groups of Lie type is presented in Sections 4–10.

The fact that equality can occur in the main theorem shows that the naive algorithm mentioned earlier cannot work well in all cases. In fact, as has been pointed out by J. J. Cannon, it works rather poorly in practice when $G = PSL(2, p)$ with $n = p + 1$; but in that situation it is straightforward to check that the probability is $2/(n - 1)$. It is easy to construct other examples where this procedure works poorly. On the other hand, the analysis in Sections 5–9 shows that the situation is better for Lie type groups of characteristic different from p. For these groups, the probability that an element is p-singular is always at least $1/p^2$, independent of the type of group; and it is also at least $(1 - p^{-1})/2h$, where h is the Coxeter number of the associated Weyl group. (These results are contained in Theorems 5.2 and 5.1, respectively.) Strong estimates are also obtained in the case of groups of Lie type in characteristic p (Theorem 10.1).

In her thesis [Ga], A. Gambini independently also found lower bounds for the probability that an element of G has order divisible by p when G is a permutation group of degree n whose order is divisible by p. While her results are not sharp, some of her methods are similar to ours. In particular, she also reduced the problem to simple groups and appealed to the classification in order to complete her proof.

2. REDUCTION

If G is any finite group, let $\mu_m(G)$ denote the probability that an element of G has order divisible by the positive integer m. Unless stated otherwise, our concern will be with the case in which $m = p$ is a fixed prime, and usually we will suppress the subscript and write $\mu(G)$ instead of $\mu_p(G)$. A simple example of this notation is the following elementary observation.

LEMMA 2.1. *If A is abelian then $\mu(A) = 1 - 1/|A|_p$.*

Here, k_p denotes the p-part of an integer k. The following is another trivial but useful observation.

LEMMA 2.2. *If $N \lhd G$, then $\mu(G) \geq \mu(G/N) + \mu(N)/|G{:}N|$.*

Proof. If a coset Nx is p-singular as an element of G/N, then every element in the coset is p-singular. This accounts for $\mu(G/N)|G/N||N| = \mu(G/N)|G|$ p-singular elements in $G - N$. The desired inequality now follows from the fact that N contains $\mu(N)|N|$ p-singular elements. ∎

We now *assume that the Main Theorem holds when G is almost simple*, and we begin work toward proving the Main Theorem in general. We suppose that G is a permutation group on a set X with $|X| = n$, that $p | |G|$ and that $\mu(G) \leq 1/n$. Working by induction, we can also *assume that the Main Theorem holds for each permutation group whose degree is less than n*. Our task is to show that G is one of the groups mentioned in part (b) of the Main Theorem (and hence $\mu(G) = 1/n$). By our assumption, we may suppose that G is not almost simple, and so our goal is to prove that G is sharply 2-transitive with n a power of p. We *assume that this is not the case* and eventually derive a contradiction.

We now proceed in several steps.

Step 1. *G is transitive on X.*

Otherwise, let $Y \subset X$ be an orbit such that the induced permutation group G^Y of G on Y has order divisible by p. Since $|Y| < n$, we have $\mu(G^Y) \geq 1/|Y| > 1/n$ by the inductive hypothesis. Also, G^Y is a homomorphic image of G, and hence $1/n \geq \mu(G) \geq \mu(G^Y)$ by Lemma 2.2. This is a contradiction.

Step 2. *G is primitive on X.*

Otherwise, let Σ be a nontrivial block system, so that $|\Sigma| < n$. If the group G^{Σ} induced by G on Σ has order divisible by p, Lemma 2.2 and the inductive hypothesis yield $1/n \geq \mu(G) \geq \mu(G^{\Sigma}) \geq 1/|\Sigma| > 1/n$, a contradiction. Thus $p \nmid |G^{\Sigma}|$.

Let M denote the kernel of the action of G on Σ. Then M is an intransitive normal subgroup of G having order divisible by p, and we may assume that Σ is the set of orbits of M. For $Y \in \Sigma$, the groups M^Y all have the same order, and hence $p | |M^Y|$ for all $Y \in \Sigma$. If H denotes the (set) stabilizer of Y in G, then $p | |H^Y|$ and $|G:H| = |\Sigma| = n/|Y|$. Counting p-singular elements, and using Lemma 2.2 and the inductive hypothesis, we have

$$|G|/n \geq \mu(G)|G| \geq \mu(H)|H| \geq \mu(H^Y)|H| \geq |H|/|Y| = |G|/n,$$

and hence equality holds throughout.

In particular, all p-singular elements of G lie in H, and hence lie in M since $M = \mathrm{core}_G(H)$, the G-core of H. But M contains elements of order p, and hence $C_G(M) \leq M$.

Also, $\mu(H) = \mu(H^Y) = 1/|Y|$, and so by induction and part (b) of the Main Theorem, either H^Y is sharply 2-transitive with $|Y|$ a power of p, or else H^Y is the full symmetric group and $|Y| = p \geq 5$. In either case,

$O_{p'}(H^Y)$ is trivial. Since $O_{p'}(M) \lhd H$, it follows that $O_{p'}(M)$ acts trivially on Y; as $Y \in \Sigma$ was arbitrary, we must have $O_{p'}(M) = 1$. Since $\mu(H) = \mu(H^Y)$, Lemma 2.2 implies that $\mu(N) = 0$, where N is the kernel of the action of H on Y. In other words, N is a p'-group. Thus $N \cap M \leq O_{p'}(M) = 1$ and N centralizes M. Then $N = C_G(M) \cap N \leq M \cap N = 1$, so that H acts faithfully on Y.

If H is sharply 2-transitive on Y, then the Frobenius kernel K of H consists of p-elements and hence is contained in M. Since K is a normal Sylow p-subgroup of M it is normal in G. Since H acts transitively on the nonidentity elements of K and $H < G$, it follows that some $g \in G - H$ centralizes some nonidentity element $k \in K$. Therefore, either g or gk is p-singular, and this contradicts the fact that M contains all of the p-singular elements of G.

The remaining possibility is that H^Y is the full symmetric group and $|Y| = p$. Since $H \cong H^Y$ and M is nontrivial, it follows that M is isomorphic to either the alternating or the symmetric group on Y, and the automorphisms of M induced by H constitute the full automorphism group of M. We conclude that some element of $G - H$ centralizes M, and this is again a contradiction.

Step 3. *G has no nontrivial abelian normal subgroup.*

Suppose that $V \lhd G$ is elementary abelian. Since G is primitive, V is regular on X. Then G/V is isomorphic to a point stabilizer and hence can be viewed as a permutation group of degree $n - 1$. If $p \mid |G: V|$, we have

$$1/n \geq \mu(G) \geq \mu(G/V) \geq 1/(n-1)$$

by Lemma 2.2 and the inductive hypothesis. This contradiction shows that G/V is a p'-group, and hence V is a p-group.

Let H be a point stabilizer in G. Since V is regular we see that H complements V and the conjugation action of H on the elements of V is permutation isomorphic to the action of H on X. Also, $H \cong G/V$ is a p'-group.

Each element of the form hc with $h \in H$ and $1 \neq c \in C_V(h)$ is p-singular. This gives

$$|G|/n \geq |G|\mu(G) \geq \sum_{h \in H} \left(|C_V(h)| - 1\right) = |H|k \geq |H| = |G|/n,$$

where k is the number of orbits in the conjugation action of H on $V - \{1\}$. Consequently, equality holds throughout. It follows that $k = 1$ and hence G is 2-transitive. It also follows that every p-singular element $g \in G$ has the form $g = hc = ch$, where $h \in H$ and $1 \neq c \in V$. Thus $g^p \in H$ for all p-singular elements $g \in G$, and since H was an arbitrary point stabilizer, we deduce that $g^p = 1$ and hence $h = 1$ since h has p'-order. It follows that $C_V(h) = 1$ for $1 \neq h \in H$, and hence G is sharply 2-transitive, contrary to our assumption.

Remark. By Step 3, we know that the socle of our group G is the direct product of some number $k \geq 2$ of nonabelian simple groups. There are two ways to proceed: (i) directly, providing an elementary approach leading to a contradiction; or (ii) using the (equally elementary) O'Nan-Scott Theorem [Cam] together with the classification of finite simple groups in order to prove a result—stronger than needed—that gives another view of the situation we have arrived at. We will describe both approaches: in Steps 4–9, and in Theorem 2.4, respectively. We digress to present a lemma concerning permutation representations of direct products of simple groups.

LEMMA 2.3. *Let N be a permutation group of degree n, and suppose that $N = T_1 \times T_2 \times \cdots \times T_k$, where $k \geq 2$ and each subgroup T_i is simple and nonabelian. Assume that the product of every $k - 1$ of the T_i is transitive. The following then hold.*

(a) *Some product of $k - 2$ of the T_i is intransitive.*

(b) *If some T_i is transitive, then $k = 2$.*

(c) $n \geq |N|^{1/2}$.

Proof. We recall that if a permutation group has the form $A \times B$, where A and B are both transitive, then each of these subgroups is regular and they are isomorphic. If (say) T_1 is transitive, we can write $N = T_1 \times M$, where M is the product of the remaining T_i. By hypothesis, M is transitive and hence $M \cong T_1$. It follows that $k = 2$ and (a) and (b) hold. Also in this case, $|N| = |T_1|^2 = n^2$, so (c) holds.

We proceed by induction on k. By the preceding paragraph, we may assume that no T_i is transitive and hence that $k > 2$. Writing $N = T_1 \times M$ as above, it follows from the transitivity of M that T_1 is semiregular. Let Σ denote the set of orbits of T_1 and note that $|\Sigma| = n/|T_1|$. Then M acts transitively on Σ, and we let K denote the kernel of this action, so that K is the direct product of some (possibly empty) collection of the T_i. Then M/K is a direct product of $k' < k$ nonabelian simple factors and is a transitive permutation group on Σ. If we delete any factor from M in order to obtain a product P/K of $k' - 1$ simple groups, then $T_1 \times P$ is transitive by hypothesis, so that P/K is transitive on Σ. In its action on Σ, therefore, M/K satisfies the hypotheses of the lemma.

By the inductive hypothesis, some product Q/K of all but two of the factors of M/K is not transitive on Σ, and thus $T_1 \times Q$ is not transitive in the original action, proving (a). Also by the inductive hypothesis, we have

$n/|T_1| = |\Sigma| \geq |M/K|^{1/2}$, and hence

$$n^2 \geq |T_1|^2 |M|/|K| = |N|\,|T_1|/|K|.$$

It suffices to show that $|T_1| \geq |K|$ in order to deduce (c) and complete the proof.

Let $Y \in \Sigma$ and note that $T_1 \times K$ acts faithfully on Y, since the kernel of this action is normal in N and has fixed points. By definition, T_1 is transitive on Y, and thus K is semiregular on Y. Consequently, $|K| \leq |Y| \leq |T_1|$, as required. ∎

We now return to the proof of the Main Theorem.

Step 4. *G has a unique minimal normal subgroup.*

Suppose that M and N are distinct minimal normal subgroups. As G is primitive, both M and N must be transitive, and since $M \cap N = 1$ we conclude that each is regular. It follows that G/N is isomorphic to a point stabilizer and so can be viewed as a permutation group of degree $n - 1$. Also, we claim that $|G/N|$ must be divisible by p: otherwise $p \mid |N| = n = |M|$, and yet $|M| \mid |G/N|$, a contradiction. The inductive hypothesis and Lemma 2.2 now yield

$$1/n \geq \mu(G) \geq \mu(G/N) \geq 1/(n-1),$$

a contradiction.

We establish some notation. By Steps 3 and 4, G has a unique minimal normal subgroup, and this subgroup—which we call N—is nonabelian. Write $N = T_1 \times T_2 \times \cdots \times T_k$, where the T_i are G-conjugate nonabelian simple groups; here $k \geq 2$ since we are assuming that G is not almost simple. Let H be the stabilizer G_x in G of $x \in X$, and observe that $G = NH$ since G is primitive by Step 2. It follows that H acts transitively on the T_i and hence $|H\colon \mathrm{N}_H(T_1)| = k$. Also, $|H\colon \mathrm{N}_H(T_1 \times T_2)| \leq k^2$.

Step 5. *Suppose that $N = A \times B$, and write $m = |H\colon \mathrm{N}_H(A)|$. If A is intransitive on X, with orbits of size r, then $r \leq m$. If also B is intransitive, then $n \leq m^2$.*

Write $K = \mathrm{N}_H(A)A$ and note that $x^K = x^A$. Since $H \cap K = \mathrm{N}_H(A)$, we have $|H\colon H \cap K| = m$. Also, $|K\colon H \cap K| = |x^K| = |x^A| = r$, and hence

$$|G{:}K|r = |G{:}K|\,|K{:}H \cap K| = |G{:}H|\,|H{:}H \cap K| = nm,$$

so that $|G{:}K| = nm/r$.

Since A is intransitive, so is K (because $x^A = x^K$) and thus K does not contain the unique minimal normal subgroup N of G. It follows that $\mathrm{core}_G(K) = 1$ and G has a faithful permutation representation of degree $|G{:}K|$. If $|G{:}K| < n$ then the inductive hypothesis yields $\mu(G) \geq 1/|G{:}K|$

$> 1/n$, which contradicts our standing assumption that $\mu(G) \leq 1/n$. Thus, $n \leq |G:K| = nm/r$, so that $r \leq m$, as required.

Finally, if B is intransitive with orbits of size s, apply the above argument with B in place of A in order to obtain $s \leq m$. (Note that $\mathrm{N}_H(A) = \mathrm{N}_H(B)$.) Since $A \times B$ is transitive, each B-orbit must contain some member of the A-orbit x^A, and hence $n/s \leq r$. Thus $n \leq rs \leq m^2$, as required.

Step 6. *The product of any $k - 1$ of the groups T_i is transitive.*

Otherwise, we may assume that $A = T_2 \times T_3 \times \cdots \times T_k$ has an orbit Y of size $r < n$. By Step 5, therefore, $r \leq |H:\mathrm{N}_H(A)| = |H:\mathrm{N}_H(T_1)| = k$.

If q is any odd prime divisor of $|T_1|$, then since all T_i have the same order, A contains an elementary abelian subgroup E of order q^{k-1}. Furthermore, since N is transitive on X and every normal subgroup of A is normal in N, a nontrivial normal subgroup of A can have no fixed points on X. It follows that A (and hence also E) is faithful on Y. Since the smallest possible degree for a faithful permutation representation of E is $q(k-1)$, this gives $q(k-1) \leq r \leq k$. This is a contradiction since $k \geq 2$.

Step 7. $k > 2$.

Suppose that $k = 2$ and choose a Sylow subgroup Q of T_1 with $|Q| > 2$. Write $M = \mathrm{N}_G(Q)$ and note that M does not contain T_1 since T_1 is simple and nonabelian. It follows that $\mathrm{core}_G(M) = 1$ since N is the unique minimal normal subgroup of G and $N \geq T_1$.

We claim that $|G:M| < n$. To see this, note that by Step 6, T_1 and T_2 are both transitive and hence regular, so that $|T_1| = n$. Now write $K = \mathrm{N}_G(T_1)$, and note that $|G:K| = k = 2$ and $M \leq K$. By the Frattini argument, $K = T_1M$ and so

$$|G:M| = 2|K:M| = 2|T_1:T_1 \cap M| \leq 2|T_1:Q| = 2n/|Q| < n,$$

as claimed.

The inductive hypothesis now yields $\mu(G) \geq 1/|G:M| > 1/n$, a contradiction.

Step 8. $k > 4$.

By Step 6, Lemma 2.3 applies to N, and by Step 7 and Lemma 2.3(b), T_1 is intransitive. Let Y be an orbit of T_1 of size $r < n$, so that by Step 5 we have $r \leq |H\colon \mathrm{N}_H(T_1)| = k$. But T_1 is nonabelian, simple, and faithful in its action on Y, so that $r = |Y| > 4$.

Step 9. *End of proof.*

Writing $t = |T_1|$, we have $|N| = t^k$ and so $n \geq t^{k/2}$ by Lemma 2.3(c). Thus $n > t^2$ by Step 8, and hence no $T_i \times T_j$ can be transitive. By Lemma 2.3(a), we may assume that $T_3 \times T_4 \times \cdots \times T_k$ is intransitive, so we can write $N = A \times B$, where $A = T_1 \times T_2$ and B are both intransitive. Then

$|H\colon \mathrm{N}_H(A)| \leq k^2$, and Step 5 yields $n \leq (k^2)^2$. Thus

$$60^{k/2} \leq t^{k/2} \leq n \leq k^4,$$

and this is our final contradiction. ∎

Remark. Parts of the proof of the O'Nan–Scott Theorem [Cam] are implicit in arguments used above. That theorem leads to the following observation, which is stronger than was needed in Steps 4–9 (a weaker version of which was also noted by Peter Cameron):

THEOREM 2.4. *If G is a primitive permutation group of degree n whose socle is neither abelian nor simple, then G has a faithful permutation representation of degree $\leq 2\sqrt{n}$.*

Proof. By the O'Nan–Scott Theorem, the socle N of G is the direct product $N = T_1 \times \cdots \times T_k$ of $k \geq 2$ isomorphic simple groups T_j, and there are four types of actions of G to consider.

Case I. X can be identified with X_1^k for some set X_1 on which T_1 acts, in such a way that $G \leq T_1 \mathrm{wr} S_k$ in the product action of this wreathed product.

Here G acts faithfully on the union Y of k copies of X_1, where $|Y| = k|X_1| \leq 2|X_1|^{k/2} = 2\sqrt{n}$.

Case II. $n = |T_1|^{k-b}$ for integers a, b with $ab = k$ and $a > 1$; $N_x = D_1 \times \cdots \times D_b$, where (after renumbering the T_j) D_i is a diagonal subgroup of $T_{(i-1)a+1} \times \cdots \times T_{ia}$; and G acts transitively on $\{T_1, \ldots, T_k\}$ with block system $\{\{T_{(i-1)a+1}, \ldots, T_{ia}\} | i = 1, \ldots, b\}$. In particular, $k - b \geq k/2$.

First note that, *for every nonabelian finite simple group T_1*, Aut T_1 *has a faithful permutation representation of degree* $< |T_1|^{1/2}$. The straightforward proof consists of checking all possible simple groups, using the obvious action for alternating groups, an action on a class of parabolic subgroups for groups of Lie type, and a perusal of subgroup lists of sporadic groups in the ATLAS [CCNPW] (compare [FKL]).

It follows that $G \leq \mathrm{Aut}\, T_1 \,\mathrm{wr}\, S_k$, and hence G has a faithful permutation representation of degree $< k|T_1|^{1/2} \leq 2|T_1|^{k/4} \leq 2|T_1|^{(k-b)/2} = 2\sqrt{n}$ (for the middle inequality we used the fact that $|T_1| \geq 60$).

Case III. $n = |X| = |T_1|^{k/2}$, $k \geq 2$, and N has two subgroups of order n each of which is regular and normal in G.

As in Case II we obtain a faithful permutation representation of G of degree $< k|T_1|^{1/2} \leq 2|T_1|^{k/4} = 2\sqrt{n}$.

Case IV. $n = |X| = |T_1|^k$, $k \geq 2$, and N is regular.

Proceed as in Case III. ∎

3. SOME SIMPLE GROUPS

In view of the preceding section, in order to prove the Main Theorem it suffices to assume that G is an almost simple group. Thus, we will consider all of the finite simple groups. This section concerns some relatively straightforward cases.

LEMMA 3.1. *Let $G = A_m$ or S_m with $m \geq 5$. If G lies in S_n, then $m \leq n$. If p is a prime dividing $|G|$, then $\mu_p(G) \geq 1/n$, with equality if and only if $p = m = n$ and $G = S_m$.*

Proof. The first assertion is obvious. Let x be one of the m points permuted by G in its usual permutation representation. We will count those p-singular elements $g \in G$ one of whose cycles is a p-cycle containing x. If $G = S_m$ then the number of such elements is at least $\binom{m-1}{p-1}(p!/p)(m-p)! = m!/m$, so that $\mu_p(S_m) \geq 1/m$ with equality if $m = p$; while if equality holds then every p-singular element has a p-cycle moving x, so that $m = p$. Similarly, if $G = A_m$ then the number of such elements $g \in G$ is at least $\binom{m-1}{p-1}(p!/p)|A_{m-p}| \geq |A_m|/m$ (where $A_0 = 1$); but in this situation the equality $\mu_p(A_m) = 1/m$ cannot hold, since it would imply first that $m = p$ and then that $\mu_p(A_m) \geq \binom{m-1}{p-1}(p!/p)|A_m| = 2/m$. ∎

Remark. The inequality in the preceding lemma is very weak. Lemma I of [ET] provides the following pleasant formula for the probability that an element of S_n is *not* p-singular: $1 - \mu_p(S_n) = \prod_{i=1}^{[n/p]}(1 - 1/ip)$. In particular, if $n \geq 4$ then $1 - \mu_2(A_n) = 2(1 - \mu_2(S_n)) \leq 2(\frac{1}{2})(\frac{3}{4})$, which yields the following result (easily proved directly):

LEMMA 3.2. *If $n \geq 4$ then $\mu_2(A_n) \geq \frac{1}{4}$, with equality only for $n = 4, 5$.*

LEMMA 3.3. *Let G be a subgroup of S_n having a unique simple normal subgroup G_0, and let p be a prime dividing $|G|$.*

(i) *If G_0 is sporadic then $\mu_p(G) > 1/n$ and $\mu_p(G) > \frac{1}{100}$.*

(ii) *If G_0 is A_6, $PSL(2,7)$, $PSL(2,8)$, $PSL(2,11)$, $PSL(3,4)$, $PSU(3,5)$, ${}^2F_4(2)'$ or $PSp(4,3)$ then $\mu_p(G) > 1/n$.*

(iii) *If $G = G_0$ is any of the groups in* (i) *or* (ii) *then $\mu_2(G) > \frac{1}{4}$.*

Proof.

(i) It is a straightforward matter to use the ATLAS [CCNPW] to check these assertions. It turns out that for each G and p there is always at least one p-singular conjugacy class x^G such that $|C_G(x)| \leq n$, where n is the degree of any faithful permutation representation of G. In fact, the only cases in which there is no such class having $|C_G(x)| < n$ occur when

G is a Mathieu group M_p, $n = p = 11$ or 23, where there are two p-singular classes of size $|G|/n$. Moreover, for every G and p in the lemma there is always a p-singular class of size $> |G|/100$—even for the larger sporadic groups. (Thus, the probability $\mu_p(G)$ is much greater than that given in the Main Theorem.)

(ii), (iii) Once again [CCNPW] can be used. However, a small amount of care is needed since that reference does not contain character tables for extensions of G_0 by noncyclic automorphism groups (Lemma 2.2 makes part of this check easier). ∎

At this point "only" the simple groups of Lie type remain to be considered. Some of those are handled here in (ii), instead of in the next section, because they have permutation representations of degree smaller than the lower bounds appearing in Lemmas 4.1 and 4.2. Note that, if $G = \mathrm{Aut}\, PSL(3,4)$ then $\mu_7(G) = \frac{1}{21}$, where 21 is the smallest degree of a faithful permutation representation of $G_0 = PSL(3,4)$ but not, of course, of G.

Remark. Later we will prove much more than is actually needed for the Main Theorem. It may be of some value to indicate an elementary approach to an approximation to what we need; this is implicit in Sections 5–9 and somewhat resembles the method used in Lemma 3.1. Consider the case $G = PSL(h,q)$, assume that p is a prime dividing $|G|$ but not $q(q-1)$, and let the integer $m \geq 1$ be minimal subject to $p | q^m - 1$. Let $t \in G$ have order p and arise from a linear transformation of the underlying vector space that induces the identity on a subspace of dimension $h - m$. Then $\mu_p(G)|G|$ is at least the number of elements of the form $t'u$ with t' conjugate to t and $u \in \mathrm{C}_G(t')$ unipotent. There are exactly $q^{(h-m)(h-m-1)}$ such elements $u \in \mathrm{C}_G(t')$. It follows that $\mu_p(G)|G| \geq |G{:}\, \mathrm{C}_G(t)| q^{(h-m)(h-m-1)}$, and hence $\mu_p(G) \geq q^{(h-m)(h-m-1)}/(q^m - 1)|SL(h - m, q)| \geq (q-1)/(q^h - 1)$, where $(q^h - 1)/(q-1)$ is the smallest degree of a faithful permutation representation of G (we are excluding all exceptions to the preceding statement, since they were covered in Lemma 3.3; cf. Lemma 4.1). The same type of argument can be used in other groups of Lie type: if $t \in G$ has order p, if q^k is the order of a maximal unipotent subgroup of $\mathrm{C}_G(t)$, and if $p \nmid q^k$, then $\mu_p(G) \geq q^{2k}/|\mathrm{C}_G(t)|$; this is enough to prove the Main Theorem when G is not $PSL(h,q)$, $PSU(h,q)$ and the rank of G is not too small. However, more is needed in order to prove the Main Theorem for all almost simple groups of Lie type. A hint of stronger counting arguments is given in an example toward the end of Section 5.

Similarly, if G has characteristic p and q^k is the order of a maximal unipotent subgroup of G, then G has $q^{2k} - 1$ nontrivial p-elements, and this provides a lower bound on $\mu_p(G)$ that is almost adequate for our purposes. Once again, far superior bounds will be given later (cf. Section 10).

4. LIE TYPE: PRELIMINARY REDUCTIONS

Throughout the remainder of this paper let G_0 denote a finite simple group of Lie type. In order to prove the Main Theorem, in view of Sections 2 and 3 we may assume that the group G appearing in that theorem lies between G_0 and $\operatorname{Aut} G_0$. (Note, however, that in all later sections G will denote an entirely different group!) In this section we will use results proved in Sections 5–7 concerning G_0 in order to prove the Main Theorem in this case. As usual, p will be a prime dividing $|G|$.

In this section we will exclude the following possibilities for G_0 treated in Lemma 3.3: $PSL(2,4) \cong A_5$, $PSL(2,5) \cong A_5$, $PSL(2,9) \cong PSp(4,2) \cong A_6$, $PSL(2,7) \cong PSL(3,2)$, $PSL(2,8) \cong {}^2G_2(3)'$, $PSL(2,11)$, $PSU(3,5)$, $PSL(3,4)$, $PSL(4,2) \cong A_8$, ${}^2F_4(2)'$, and $PSp(4,3) \cong PSU(4,2)$. Also, we will not have to consider the non-simple group $G_2(2) \cong \mathrm{P\Gamma U}(3,3)$.

Let h denote the *Coxeter number* of G_0: the Coxeter number of the Weyl group of a (B,N)-pair for a split form of G_0 over an algebraically closed field k of characteristic s. In view of the list of excluded groups, this (B,N)-pair is unique up to conjugation. Similarly, let l denote the rank of a split form of G_0 over k. The groups G_0 are listed in Table I. See Section 5 for a definition of the number q appearing in the table. There is an

TABLE I

G_0	$m_{G_0} \geq$
$A_{h-1}, h \geq 2$	$(q^h - 1)/(q - 1)$
${}^2A_{h-1}, h \geq 3$	q^{2h-3}
$B_l, C_l, l \geq 2$	q^{2l-1}
$D_l, l \geq 4$	q^{2l-2}
${}^2D_l, l \geq 3$	q^{2l-2}
$G_2, q \neq 4$	$q^3/2$
${}^3D_4$	$q^5/2$
F_4	q^6
${}^2F_4$	$q^{10}/2$
$E_6, {}^2E_6$	q^{10}
E_7	q^{16}
E_8	q^{28}
${}^2B_2$	q^4
${}^2G_2$	q^6

integer e such that $q = s^e$ (and then let $\delta = 1$), unless $G_0 = {}^2B_2(q^2)$, ${}^2G_2(q^2)$, or ${}^2F_4(q^2)$, in which case $q^2 = s^3$ (and then let $\delta = 2$).

Let m_G denote the smallest possible degree of a faithful permutation representation of a group G.

LEMMA 4.1. *A lower bound for m_{G_0} is given in Table* I.

Proof. If G_0 is a classical group then this holds by [Co] or [Li]; note that we have excluded those groups G_0 that would have provided counterexamples to some of these bounds.

For ${}^2B_2(q^2)$ and ${}^2G_2(q^2)$ see [Su, Wa]. In the case of the remaining exceptional or twisted groups [LS] provides a suitable lower bound on $m_{G_0} - 1$. (In fact, the bounds in [LS] are sufficient for most of what is needed for the Main Theorem even in the case of the classical groups.) ▮

It should be noted that the above bounds are crude except in the case of type A_{h-1}. Only in the latter case do any of the calculations of this section become at all delicate. Note that the case $G_2(4)$ could have been dealt with as in Lemma 3.3, but we want to indicate that excluded groups can be handled by the methods of this section if just a little more care is taken.

LEMMA 4.2.

(i) $m_{G_0} > 2h|\mathrm{Out}\ G_0|$.

(ii) *If G_0 is either $PSL(3, q)$ with $3|q - 1$ or $G_0 = PSU(3, q)$ with $3|q + 1$, then $m_{G_0} > 9|\mathrm{Out}\ G_0|$.*

(iii) *If G_0 is not $PSL(2, q)$ or $PSL(3, q)$ then $m_{G_0} > (5q^\delta/2)|\mathrm{Out}\ G_0|$.*

(iv) *If G_0 is $PSL(3, q)$ then $m_{G_0} > \{q/(1 - (3, q - 1)q^{-2})\}|\mathrm{Out}\ G_0|$.*

Proof. This involves a straightforward but tedious application of Lemma 4.1 combined with knowledge of Out G_0 [Car1], considering the various possibilities for G_0 separately. We will outline the argument only in the "hardest" case: part (i) with $G_0 = PSL(h, q)$.

By Lemma 4.1 we must show that $(q^h - 1)/(q - 1) > 2h \cdot \{\min(h - 1, 2)\}e(h, q - 1)$. If $h \geq 7$ then $q^{h-2} = s^{e(h-2)} \geq es^{h-2} \geq 4eh$. The cases $h = 4, 5, 6$ are easily handled directly. If $h = 3$ then $PSL(3, 4)$ is being excluded, and it is straightforward to check that $m_{G_0} = q^2 + q + 1 > 6e \cdot (3, q - 1)$ (and that $q^2 + q + 1 > 9 \cdot 6e$, as needed for (ii)). Finally, when $h = 2$ it is easy to check that $q + 1 > 4 \cdot e(2, q - 1)$.

This completes the argument when $G_0 = PSL(h, q)$. The unitary case is similar, though simpler. For all remaining groups of Lie type, the outer automorphism group of G_0 is smaller than we have just encountered, while m_{G_0} is noticeably larger than q^l, making all of the inequalities quite a bit easier to deal with. However, when G_0 is $G_2(q)$ with $q = 3$ or 4 the bounds in Table I are not adequate for our purposes; but in those cases, the embedding $PSU(3, q) < G_2(q)$ leads to the slight improvement $m_{G_0} \geq m_{PSU(3,q)} \geq q^3$, and this is strong enough to prove (iv). ∎

Lemma 4.3. *If $p \neq s$ and $p \mid |G_0|$ then $\mu(G) > 1/m_{G_0} \geq 1/m_G$.*

Proof. By Lemma 2.2, $\mu(G) \geq \mu(G_0)/|G{:}G_0|$. In Theorem 5.1 we will show that $\mu(G_0) \geq \frac{1}{9}$ if either $G_0 = PSL(3, q)$ with $p = 3 = (q - 1)_3$ or $G_0 = PSU(3, q)$ with $p = 3 = (q + 1)_3$, while $\mu(G_0) \geq (1 - 1/p)/h \geq 1/2h$ otherwise. Thus, Lemma 4.2(i, ii) produces the desired inequality. ∎

Lemma 4.4. *If $p \neq s$, $p \mid |G|$ and $p \nmid |G_0|$, then $\mu(G) > 1/m_{G_0} \geq 1/m_G$.*

Proof. Since $p \nmid |G_0|$, a Sylow p-subgroup of $\operatorname{Out} G_0$ is central (a property of $\operatorname{Out} G_0$ easily deduced from the description of this group in [Car1]). Then Lemma 2.2 implies that $\mu(G) \geq \mu(\mathbb{Z}_p) \geq \frac{1}{2} > 1/m_{G_0}$. ∎

Lemma 4.5. *If $s = p$ then $\mu(G) > 1/m_{G_0} \geq 1/m_G$.*

Proof. If $G_0 = PSL(2, q)$ then $\mu(G_0) = (q + 1)(q - 1)/[(q + 1)q(q - 1)/(2, q - 1)] = (2, q - 1)/q$. We must show that $\mu(G) > 1/(q + 1)$. If $p \mid |G{:}G_0|$ then, by Lemma 2.2, $\mu(G) \geq \mu(\mathbb{Z}_p) \geq \frac{1}{2}$ since $\operatorname{Aut} G_0$ is abelian. Suppose that $p \nmid |G{:}G_0|$. If G contains no field automorphism then $|G{:}G_0| \leq (2, q - 1)$ and hence $\mu(G) \geq (2, q - 1)/q|G{:}G_0| \geq 1/q$. If G contains a nontrivial field automorphism f then f centralizes some element r of order p in G_0. Then $|C_G(rf)| \leq q'|G{:}G_0|$, where $q = q'^i$ with $i \geq 2$, so that $\mu(G) \geq 1/q'|G{:}G_0| \geq 1/q'(2, q - 1)e \geq 1/q$.

If $G_0 = PSL(3, q)$ then $\mu_p(G_0) = (1 - (3, q - 1)q^{-2})/q$ by Example (b) in Section 10, and then $\mu_p(G) \geq \{(1 - (3, q - 1)q^{-2})/q\}/|\operatorname{Out} G_0| > 1/m_{G_0}$ by Lemma 4.2(iv). For all of the remaining groups, by Theorem 10.1 and Lemma 2.2 we have $\mu_p(G) \geq (2/5q^{\delta})/|\operatorname{Out} G_0|$, so that the desired inequality follows from Lemma 4.2(iii). ∎

Note that Lemmas 4.3–4.5 *complete the proof of the Main Theorem* (*modulo Theorems* 5.1 *and* 10.1 *below*).

5. GROUPS OF LIE TYPE: STATEMENTS OF RESULTS (UNEQUAL CHARACTERISTIC CASE)

It will now be convenient to change notation. As before, let G_0 be a finite simple group of Lie type and let p be a prime dividing $|G_0|$ other than the defining characteristic. However, now G will denote a simply connected simple algebraic group, defined over an algebraic closure k of a finite field, such that $G_0 \cong G^F/Z^F$ for a bijective endomorphism F of G, where $Z = Z(G)$. (As usual, for any F-invariant set X we let X^F denote the set of fixed points of F in X.) In the few cases where G_0 has non-conjugate (B, N)-pairs, any convenient choice can be used provided that its characteristic is not p. Let l and h denote the rank and Coxeter number of G. Let δ be the smallest integer such that F^{δ} is the Frobenius endomorphism corresponding to a definition of G_0 over a finite subfield $\mathbb{F}_{q^{\delta}}$ of k. (Then $\delta = 1$ unless G_0 is a Suzuki or Ree group, in which case $\delta = 2$. Thus, in the latter cases we will be dealing with ${}^2B_2(q^2)$, ${}^2G_2(q^2)$ and ${}^2F_4(q^2)$; cf. Table I.) Throughout the remainder of this paper we will no longer exclude any of the cases G_0 treated separately in Sections 3 and 4.

THEOREM 5.1. *$\mu(G_0) \geq (1/h)(1 - 1/p)$ except when $p = 3$ and $G_0 \cong PSL(3, q)$ with $(q - 1)_3 = 3$, or $G_0 \cong PSU(3, q)$ with $(q + 1)_3 = 3$, in which case $\mu(G_0) = \frac{1}{9}$.*

Theorem 5.1 is an immediate consequence of the following result.

THEOREM 5.1′. *$\mu(G^F/Z^F) \geq (1/h)(1 - 1/p)$ except when $p = 3$, $G \cong SL(3, k)$, $|Z^F| = 3$ and $|G^F/Z^F|_3 = 3$, in which case $\mu(G^F/Z^F) = \frac{1}{9}$.*

Theorem 5.1 emphasizes the type of G_0. For a fixed prime p, we can also consider the number

$$\mu_p = \inf_{G_0} \mu_p(G_0),$$

where G_0 runs over all finite simple groups of Lie type of characteristic not p having order divisible by p.

THEOREM 5.2. *$\mu_2 = \frac{1}{4}$, $\mu_3 = \frac{1}{9}$, $\mu_5 = \frac{4}{25}$, $\mu_{11} = \frac{9}{121}$, and $\mu_p \geq 1/p - 1/2p^2$ if $p \notin \{2, 3, 5, 11\}$.*

As in the case of Theorem 5.1, we will prove the corresponding result for groups of the form G^F/Z^F; see Section 9.

EXAMPLE. In order to outline the approach about to be taken, we consider the case $G_0 = PSL(h, q)$, viewed using $h \times h$ matrices. Assume that $p \nmid q - 1$ and that q is not too small. Semisimple elements are just s'-elements, where s is the prime dividing q. Each semisimple p-singular element t is a conjugate of suitable block diagonal matrix; the size of at least one of these blocks is divisible by the smallest integer m such that $p | q^m - 1$. Note that there is a partition of h arising here. Each such p-singular element t centralizes various s-elements, and each p-singular element has a Jordan decomposition $tu = ut$, where t is p-singular and semisimple while u is an s-element (i.e., unipotent). We need to estimate the number of pairs (t, u) that can arise here. This is accomplished by counting the number of pairs (t, T) with T a conjugate of an abelian group of block-diagonal matrices with blocks coming from extension fields of $\mathbb{F}_q$, where the degrees of the extensions are the members of suitable partitions of h, one of which is the aforementioned partition. This yields a formula (Theorem 6.2) for $\mu(SL(h,q))$, and this formula has a term arising from conjugacy classes in the Weyl group S_h of $SL(h,q)$. Finally, we estimate this term in order to prove the desired inequality.

We note the following amusing consequences of Lemmas 2.2, 3.2, and 3.3 and Theorem 5.2:

COROLLARY 5.3. *If G is any group having a simple homomorphic image that is neither cyclic nor Lie type of characteristic* 2, *then* $\mu_2(G) \geq \frac{1}{4}$.

6. MAXIMAL TORI

As above let G be a connected reductive algebraic group defined over an algebraic closure k of a finite field, with center Z and Weyl group W, and let F be a surjective endomorphism of G such that G^F is finite.

Let $\mathcal{T}$ be the variety of all maximal tori in G. Then F acts on W and on $\mathcal{T}$. The W-orbits in W for the action $w \cdot x = wxF(w)^{-1}$ are called *F-conjugacy classes* in W, and the set of all F-conjugacy classes in W is denoted $W/\sim_F$. The G^F-conjugacy classes in $\mathcal{T}^F$ (that is, the G^F-conjugacy classes of F-stable maximal tori) are parametrized in a natural way by the F-conjugacy classes in W, and for $C \in W/\sim_F$ we let $\mathcal{T}_C$ denote the corresponding G^F-orbit in $\mathcal{T}^F$ (see Appendix A.6). Let T_C be an element of $\mathcal{T}_C$; its stabilizer in G^F has order $|T_C^F||W||C|^{-1}$ [SS, II.1.8]. It follows that

$$\sum_{T\in\mathcal{T}_C} |T^F| = |\mathcal{T}_C||T_C^F| = \frac{|C|}{|W|}|G^F|. \tag{6.1}$$

Let T be an F-stable maximal torus of G. Then $T > Z$, and we say that T is *p-relevant* if $p \mid |T^F/Z^F|$. We say that an F-conjugacy class C in W and its elements are *p-relevant* if the tori in $\mathcal{T}_C$ are p-relevant. The probability $\mu^W(G,F) = \mu_p^W(G,F)$ that a random element in W is p-relevant is

$$\mu^W(G,F) = \frac{1}{|W|} \sum_{\substack{C\in W/\sim_F \\ C/p\text{-relevant}}} |C|.$$

THEOREM 6.2.

$$\mu(G^F/Z^F) = (1/|W|)\Sigma_{C\in W/\sim_F}(1 - 1/|T_C^F/Z^F|_p)|C|.$$

In particular,

$$(1 - 1/p)\mu^W(G,F) \leq \mu(G^F/Z^F) < \mu^W(G,F),$$

and the first inequality is an equality if and only if $|T^F/Z^F|_p \leq p$ *for every* F*-stable maximal torus* T *of* G.

Proof. In view of Lemma 2.1 this is a special case of the following result.

THEOREM 6.2′. *Let* m *be any integer and let* K *be an* F*-stable closed subgroup of* $Z(G)$. *If* m *is prime to* char(k), *then*

$$\mu_m(G^F/K^F) = \frac{1}{|W|}\sum_{C\in W/\sim_F}\mu_m(T_C^F/K^F)|C|.$$

Proof. Let X_m be the set of all elements in G^F whose images in G^F/K^F have order divisible by m. If $g \in G^F$ has Jordan decomposition su, with s semisimple and u unipotent, then $g \in X_m$ if and only if $s \in X_m$. By definition, $|X_m| = |G^F|\,\mu_m(G^F/K^F)$. For a semisimple element $t \in X_m$, let $X_m(t)$ be the set of all elements of G^F which have t as semisimple part. Then X_m is the disjoint union of its subsets $X_m(t)$ with $t \in X_m$ semisimple Moreover, $|X_m(t)|$ is the number of unipotent elements in $C_G(t)^F = C_{G^F}(t)$.

Let Y_m be the set of all pairs (t, T) consisting of a semisimple element $t \in X_m$ and an F-stable maximal torus T of G containing t. By a result of Steinberg [St, 14.14, 15.1], $C_G(t)$ has as many F-stable unipotent elements as F-stable maximal tori. Thus, by (6.1),

$$\begin{aligned}
|X_m| = |Y_m| &= \sum_{T\in\mathcal{T}^F}|T\cap X_m| \\
&= \sum_{T\in\mathcal{T}^F}\mu_m(T^F/K^F)|T^F| \\
&= \sum_{C\in W/\sim_F}\sum_{T\in\mathcal{T}_C}\mu_m(T^F/K^F)|T^F| \\
&= \sum_{C\in W/\sim_F}\mu_m(T_C^F/K^F)\sum_{T\in\mathcal{T}_C}|T^F| \\
&= |G^F|\frac{1}{|W|}\sum_{C\in W/\sim_F}\mu_m(T_C^F/K^F)|C|. \qquad\blacksquare
\end{aligned}$$

EXAMPLE. We will show how Theorem 6.2 can be used to compute $\mu(G^F/Z^F)$ when $p = 2$ and G is simply connected of type A_3 with char$(k) \neq 2$. Here F is the Frobenius endomorphism for a definition of G over a subfield $\mathbb{F}_q$ of k. Let $\varepsilon = 1$ if (G, F) is split, $\varepsilon = -1$ otherwise. The Weyl group W is isomorphic to S_4 and the F-conjugacy classes in W can be parametrized in a natural way by the partitions of 4. The F-conjugacy classes are just the conjugacy classes when (G, F) is split, while in the non-split case they are obtained by multiplying the conjugacy classes by the longest word w_0 in W; in particular, the size of each F-conjugacy class is the size of the corresponding conjugacy class. Writing C_λ for the F-conjugacy class corresponding to the partition λ, and T_λ instead of T_{C_λ}, we have

$$|C_{(1,1,1,1)}| = 1, \qquad |C_{(2,1,1)}| = 6, \qquad |C_{(2,2)}| = 3,$$
$$|C_{(3,1)}| = 8, \qquad |C_{(4)}| = 6,$$

and (cf. [Car4])

$$|T^F_{(1,1,1,1)}| = (q-\varepsilon)^3, \qquad |T^F_{(2,1,1)}| = (q+\varepsilon)(q-\varepsilon)^2,$$
$$|T^F_{(2,2)}| = (q+\varepsilon)^2(q-\varepsilon),$$
$$|T^F_{(3,1)}| = (q^2+\varepsilon q+1)(q-\varepsilon), \qquad |T^F_{(4)}| = (q^2+1)(q+\varepsilon).$$

Note that $(q^2 + \varepsilon q + 1)_2 = 1$ and $(q^2 + 1)_2 = 2$.

Suppose that $|Z^F| = 4$. Let $\pi = (q - \varepsilon)_2$. Then $\pi \geq 4$ and $(q + \varepsilon)_2 = 2$. It follows easily that

$$|T^F_{(1,1,1,1)}/Z^F|_2 = \pi^3 \cdot 2^{-2} > 1, \qquad |T^F_{(2,1,1)}/Z^F|_2 = \pi^2 \cdot 2^{-1} > 1,$$
$$|T^F_{(2,2)}/Z^F|_2 = \pi > 1, \qquad |T^F_{(3,1)}/Z^F|_2 = \pi \cdot 2^{-2} \geq 1,$$
$$|T^F_{(4)}/Z^F|_2 = 1.$$

Let $P(X,Y) = \frac{1}{24}X^3Y^{-2} + \frac{1}{4}X^2Y^{-1} + \frac{1}{8}X + \frac{1}{3}XY^{-2}$. By Theorem 6.2, $\mu(G^F/Z^F) = P(1,1) - P(1/\pi, 1/p)$.

Suppose now that $|Z^F| = 2$. Let $\pi = (q + \varepsilon)_2$. Then $\pi \geq 4$, $(q - \varepsilon)_2 = 2$, and

$$|T^F_{(1,1,1,1)}/Z^F|_2 = 2^2 > 1, \qquad |T^F_{(2,1,1)}/Z^F|_2 = \pi \cdot 2 > 1,$$
$$|T^F_{(2,2)}/Z^F|_2 = \pi^2 > 1, \qquad |T^F_{(3,1)}/Z^F|_2 = 1,$$
$$|T^F_{(4)}/Z^F|_2 = \pi > 1.$$

Let $Q(X,Y) = \frac{1}{24}Y^2 + \frac{1}{4}XY + \frac{1}{8}X^2 + \frac{1}{4}X$. By Theorem 6.2, $\mu(G^F/Z^F) = Q(1,1) - Q(1/\pi, 1/p)$.

The next section contains a generalization of this type of computation.

7. CYCLOTOMIC POLYNOMIALS AND EXCEPTIONAL GROUPS

Theorem 6.2 allows us to obtain an explicit formula for $\mu(G^F/Z^F)$ for every prime $p \neq \operatorname{char}(k)$ dividing $|G^F/Z^F|$. In this section we give results for the exceptional types (in which we include the triality twisted groups and the Suzuki and Ree groups) in the case where G is simply connected. The necessary data concerning $|Z^F|$, (F-)conjugacy classes in Weyl groups and orders of maximal tori can be found, for example, in [St, pp. 131, 192, 193], [Car2–4], [De], [DF], [DLi], [Shi1, 2], [Sho], [SS, II.1.7], or can be easily worked out.

Let δ be as in Section 5, and let $A = \mathbb{Z}[\operatorname{char}(k)^{1/\delta}]$ (so A is one of the rings $\mathbb{Z}$, $\mathbb{Z}[\sqrt{2}]$ or $\mathbb{Z}[\sqrt{3}]$). For every F-conjugacy class C in W, $|T_C^F|$ can be considered in a natural way as a polynomial in q with coefficients in A. More precisely, let $X(T_C) = \operatorname{Hom}(T_C, \mathbf{G}_m)$ (where $\mathbf{G}_m$ denotes the multiplicative group of k viewed as an algebraic group). Then $X(T_C)$ is a free abelian group, and F induces an endomorphism of it in a natural manner. Let $\chi_C \in \mathbb{R}[X]$ be the characteristic polynomial of the endomorphism $q^{-1}F$ of $X(T_C) \otimes_{\mathbb{Z}} \mathbb{R}$. Then

$$|T_C^F| = \chi_C(q) \qquad \text{and} \qquad \chi_C \in A[X]$$

(see [Car5, 3.3.5] or Appendix A.8).

When $A = \mathbb{Z}$ the irreducible factors of χ_C in $A[X]$ are cyclotomic polynomials $\varphi_m(X)$. When $A \neq \mathbb{Z}$ the factorization of χ_C is slightly more complicated: in this situation we want factorizations of χ_C which induce factorizations of $|T_C^F| = \chi_C(q)$ in $\mathbb{Z}$. For this reason we look for factors of the form $\varphi_m(X^2)$, or factors $\psi \in A[X]$ of $\varphi_m(X^2)$ such that $\psi(q) \in \mathbb{Z}$. It turns out that the only further factorizations we need are the following. If $A = \mathbb{Z}[\sqrt{2}]$ then $\varphi_8(X) = \varphi_4(X^2) = (X^2 + \sqrt{2}X + 1)(X^2 - \sqrt{2}X + 1)$ and $\varphi_{24}(X) = \varphi_{12}(X^2) = (X^4 + \sqrt{2}X^3 + X^2 + \sqrt{2}X + 1)(X^4 - \sqrt{2}X^3 + X^2 - \sqrt{2}X + 1)$. If $A = \mathbb{Z}[\sqrt{3}]$ then $\varphi_{12}(X) = \varphi_6(X^2) = (X^2 + \sqrt{3}X + 1)(X^2 - \sqrt{3}X + 1)$.

There is a cyclotomic polynomial $\varphi_m(X)$ such that $p \mid \varphi_m(q^\delta)$. We choose m minimal with respect to this condition except when $p = 2$, in which case we also require that $4 \mid \varphi_m(q^\delta)$. Let $\varphi(X) = \varphi_m(X^\delta)$, except if $A \neq \mathbb{Z}$ and $\varphi_m(X^2)$ factors as above, in which case we let φ be the unique irreducible factor of $\varphi_m(X^2)$ in $A[X]$ such that $p \mid \varphi(q)$. Let $\pi = \varphi_m(q^\delta)_p = \varphi(q)_p$.

For each F-conjugacy class C in W let i be the largest integer such that φ^i divides χ_C. Then π^i divides $\chi_C(q)$. Additional factors of $|T_C^F|_p$ arise due to the fact that $\varphi_{m'}(q^\delta)_p = p$ when $m'/m \neq 1$ is a power of p or when $p = m = 2$ and $m' = 1$, so that these contributions depend on the type of (G, F), as well as on p and m, but not directly on q. On the other hand, we must divide $|T_C^F|_p$ by $|Z^F|_p$. In this section we are only considering exceptional groups, in which case $|Z^F|_p \leq p$, so that $|Z^F|_p$ again depends on the type of (G, F), m, and p but not directly on q. (N.B. If we had been considering groups of type A_l we also would have had to take $(|Z^F|, \pi)$ into account.)

Consequently, there exist integers $i \geq 0$ and j depending on p, m, and the type of (G, F), but not on q, such that $|T_C^F/Z^F|_p = \pi^i p^j$, and we associate with C the polynomial $|C|\,|W|^{-1}X^iY^j \in \mathbb{Q}[X, Y, Y^{-1}]$. This produces a polynomial

$$P_{m,p}(X, Y) = \sum \left\{ \frac{|C|}{|W|} X^i Y^j \,\middle|\, C \in W/\sim_F \text{ and } |T_C^F/Z^F|_p = \pi^i p^j, \right.$$
$$\left. (i, j) \neq (0, 0) \right\}$$

in $\mathbb{Q}[X, Y, Y^{-1}]$ with non-negative coefficients. By Theorem 6.2 and (2.1),

$$\mu(G^F/Z^F) = P_{m,p}(1, 1) - P_{m,p}(1/\pi, 1/p).$$

Note that at $(1/\pi, 1/p)$ the monomials involved in $P_{m,p}$ take values of the form p^f with $f \leq 0$. Since $|Z^F|_p \leq p$ for exceptional groups, in the tables below $f = 0$ occurs only when $\pi = p$ and the monomial is XY^{-1}. Thus, if c is the coefficient of XY^{-1} in $P_{m,p}$ then, by definition, $\mu^W(G, F) = P_{m,p}(1, 1) - \delta_{\pi,p}c$. In fact, $c \neq 0$ only in the following cases: for $(m, p) = (1, 3)$ in type E_6, $(m, p) = (2, 3)$ in type ${}^2E_6$ and $p = 2$ in type E_7. In the corresponding polynomials the term involving XY^{-1} will be printed in boldface.

For most values of p (e.g., if $p \nmid |W|$), $P_{m,p}$ is actually a polynomial P_m in X and is independent of p. When $P_{m,p} = P_m$ we have $\mu(G^F/Z^F) = P_m(1) - P_m(1/\pi)$. We give below the polynomials P_m along with those $P_{m,p}$'s which are different from P_m. (These polynomials were obtained using a simple computer program.) For groups of type ${}^2E_6$ we write ${}^2P_{m,p}$ and ${}^2P_m$ for $P_{m,p}$ and P_m, respectively, reserving the notation $P_{m,p}$ and P_m for split groups of type E_6.

Types $E_6, {}^2E_6$. We have

$$P_1 = {}^2P_2 = \tfrac{619}{1440}X + \tfrac{4363}{17280}X^2 + \tfrac{5}{72}X^3 + \tfrac{17}{1728}X^4 + \tfrac{1}{1440}X^5 + \tfrac{1}{51840}X^6,$$

$$P_2 = {}^2P_1 = \tfrac{23}{48}X + \tfrac{95}{576}X^2 + \tfrac{1}{48}X^3 + \tfrac{1}{1152}X^4,$$

$$P_3 = {}^2P_6 = \tfrac{61}{216}X + \tfrac{1}{27}X^2 + \tfrac{1}{648}X^3,$$

$$P_4 = {}^2P_4 = \tfrac{3}{16}X + \tfrac{1}{96}X^2, \qquad P_5 = {}^2P_{10} = \tfrac{1}{5}X, \qquad P_6 = {}^2P_3 = \tfrac{2}{9}X + \tfrac{1}{72}X^2,$$

$$P_8 = {}^2P_8 = \tfrac{1}{8}X, \qquad P_9 = {}^2P_{18} = \tfrac{1}{9}X, \qquad P_{12} = {}^2P_{12} = \tfrac{1}{12}X,$$

with the following exceptions:

$$\begin{aligned}
P_{1,2} = {}^2P_{2,2} = {} & \tfrac{1}{36}Y^2 + \tfrac{19}{90}XY + \tfrac{5}{24}XY^2 + \tfrac{1}{96}XY^4 + \tfrac{59}{540}X^2 + \tfrac{23}{288}X^2Y^2 \\
& + \tfrac{1}{16}X^2Y^3 + \tfrac{1}{1152}X^2Y^4 + \tfrac{1}{36}X^3Y + \tfrac{1}{32}X^3Y^2 + \tfrac{1}{96}X^3Y^3 \\
& + \tfrac{1}{216}X^4 + \tfrac{1}{192}X^4Y^2 + \tfrac{1}{1440}X^5Y + \tfrac{1}{51840}X^6,
\end{aligned}$$

$$\begin{aligned}
P_{2,2} = {}^2P_{1,2} = {} & \tfrac{59}{540}Y^2 + \tfrac{13}{864}Y^4 + \tfrac{1}{51840}Y^6 + \tfrac{19}{90}XY + \tfrac{5}{24}XY^2 + \tfrac{1}{36}XY^3 \\
& + \tfrac{1}{32}XY^4 + \tfrac{1}{1440}XY^5 + \tfrac{1}{36}X^2 + \tfrac{5}{72}X^2Y^2 + \tfrac{1}{16}X^2Y^3 \\
& + \tfrac{1}{192}X^2Y^4 + \tfrac{1}{96}X^3Y^2 + \tfrac{1}{96}X^3Y^3 + \tfrac{1}{1152}X^4Y^2,
\end{aligned}$$

$$\begin{aligned}
P_{1,3} = {}^2P_{2,3} = {} & \tfrac{1}{648}Y^2 + \tfrac{51}{160}\mathbf{XY}^{-1} + \tfrac{1}{12}X + \tfrac{1}{36}XY + \tfrac{129}{640}X^2Y^{-1} + \tfrac{1}{24}X^2 \\
& + \tfrac{1}{108}X^2Y + \tfrac{1}{24}X^3Y^{-1} + \tfrac{1}{36}X^3 + \tfrac{1}{192}X^4Y^{-1} + \tfrac{1}{216}X^4 \\
& + \tfrac{1}{1440}X^5Y^{-1} + \tfrac{1}{51840}X^6Y^{-1},
\end{aligned}$$

$$P_{2,3} = {}^2P_{1,3} = \tfrac{1}{72}Y^2 + \tfrac{5}{16}X + \tfrac{1}{6}XY + \tfrac{7}{64}X^2 + \tfrac{1}{18}X^2Y + \tfrac{1}{48}X^3 + \tfrac{1}{1152}X^4,$$

$$\begin{aligned}
P_{1,5} = {} & \tfrac{95}{288}X + \tfrac{1}{10}XY + \tfrac{527}{3456}X^2 + \tfrac{1}{10}X^2Y + \tfrac{5}{72}X^3 + \tfrac{17}{1728}X^4 + \tfrac{1}{1440}X^5 \\
& + \tfrac{1}{51840}X^6.
\end{aligned}$$

Type E_7. We have

$$\begin{aligned}
P_1 = P_2 = {} & \tfrac{1286963}{2903040}X + \tfrac{1171}{5120}X^2 + \tfrac{7759}{138240}X^3 + \tfrac{23}{3072}X^4 + \tfrac{77}{138240}X^5 + \tfrac{1}{46080}X^6 \\
& + \tfrac{1}{2903040}X^7,
\end{aligned}$$

$$P_3 = P_6 = \tfrac{109}{432}X + \tfrac{11}{432}X^2 + \tfrac{1}{1296}X^3,$$

$$P_4 = \tfrac{3}{16}X + \tfrac{1}{96}X^2, \qquad P_5 = P_{10} = \tfrac{1}{10}X, \qquad P_7 = P_{14} = \tfrac{1}{14}X,$$

$$P_8 = \tfrac{1}{8}X, \qquad P_9 = P_{18} = \tfrac{1}{18}X, \qquad P_{12} = \tfrac{1}{12}X,$$

with the following exceptions:

$$\begin{aligned}
P_{1,2} = P_{2,2} = {} & \tfrac{227}{4320}Y^2 + \tfrac{53}{34560}Y^4 + \tfrac{1}{2903040}Y^6 + \tfrac{1189}{5670}\mathbf{XY}^{-1} + \tfrac{217}{1440}XY + \tfrac{7}{96}XY^2 \\
& + \tfrac{17}{2304}XY^3 + \tfrac{1}{384}XY^4 + \tfrac{1}{46080}XY^5 + \tfrac{43}{360}X^2 + \tfrac{7}{96}X^2Y \\
& + \tfrac{41}{2304}X^2Y^2 + \tfrac{7}{384}X^2Y^3 + \tfrac{1}{3072}X^2Y^4 + \tfrac{23}{1080}X^3Y^{-1} \\
& + \tfrac{35}{2304}X^3Y + \tfrac{7}{384}X^3Y^2 + \tfrac{13}{9216}X^3Y^3 + \tfrac{1}{288}X^4 + \tfrac{1}{384}X^4Y \\
& + \tfrac{13}{9216}X^4Y^2 + \tfrac{1}{4320}X^5Y^{-1} + \tfrac{1}{3072}X^5Y + \tfrac{1}{46080}X^6 + \tfrac{1}{2903040}X^7Y^{-1},
\end{aligned}$$

$$\begin{aligned}
P_{1,3} = P_{2,3} = {} & \tfrac{1}{144}Y^2 + \tfrac{1639}{7168}X + \tfrac{77}{360}XY + \tfrac{1}{1296}XY^3 + \tfrac{691}{5120}X^2 + \tfrac{23}{288}X^2Y \\
& + \tfrac{1}{72}X^2Y^2 + \tfrac{631}{15360}X^3 + \tfrac{1}{96}X^3Y + \tfrac{1}{216}X^3Y^2 + \tfrac{37}{9216}X^4 \\
& + \tfrac{1}{288}X^4Y + \tfrac{1}{3072}X^5 + \tfrac{1}{4320}X^5Y + \tfrac{1}{46080}X^6 + \tfrac{1}{2903040}X^7, \\
P_{1,5} = P_{2,5} = {} & \tfrac{238039}{580608}X + \tfrac{1}{30}XY + \tfrac{183}{1024}X^2 + \tfrac{1}{20}X^2Y + \tfrac{1091}{27648}X^3 + \tfrac{1}{60}X^3Y \\
& + \tfrac{23}{3072}X^4 + \tfrac{77}{138240}X^5 + \tfrac{1}{46080}X^6 + \tfrac{1}{2903040}X^7, \\
P_{1,7} = P_{2,7} = {} & \tfrac{154229}{414720}X + \tfrac{1}{14}XY + \tfrac{1171}{5120}X^2 + \tfrac{7759}{138240}X^3 + \tfrac{23}{3072}X^4 + \tfrac{77}{138240}X^5 \\
& + \tfrac{1}{46080}X^6 + \tfrac{1}{2903040}X^7.
\end{aligned}$$

Type E_8. We have

$$\begin{aligned}
P_1 = P_2 = {} & \tfrac{539179}{1161216}X + \tfrac{6504289}{34836480}X^2 + \tfrac{9799}{276480}X^3 + \tfrac{62299}{16588800}X^4 + \tfrac{13}{55296}X^5 \\
& + \tfrac{43}{4976640}X^6 + \tfrac{1}{58060800}X^7 + \tfrac{1}{696729600}X^8, \\
P_3 = P_6 = {} & \tfrac{355}{1944}X + \tfrac{77}{5184}X^2 + \tfrac{1}{1944}X^3 + \tfrac{1}{155520}X^4, \\
P_4 = {} & \tfrac{185}{768}X + \tfrac{127}{4608}X^2 + \tfrac{1}{768}X^3 + \tfrac{1}{46080}X^4,
\end{aligned}$$

$$P_5 = P_{10} = \tfrac{2}{25}X + \tfrac{1}{600}X^2, \qquad P_7 = P_{14} = \tfrac{1}{14}X, \qquad P_8 = \tfrac{5}{32}X + \tfrac{1}{192}X^2,$$

$$P_9 = P_{18} = \tfrac{1}{18}X, \qquad P_{12} = \tfrac{17}{144}X + \tfrac{1}{288}X^2, \quad P_{15} = P_{30} = \tfrac{1}{30}X,$$

$$P_{20} = \tfrac{1}{20}X, \qquad P_{24} = \tfrac{1}{24}X,$$

with the following exceptions:

$$\begin{aligned}
P_{1,2} = P_{2,2} = {} & \tfrac{48427}{544320}Y^2 + \tfrac{22373}{20736000}Y^4 + \tfrac{143}{2488320}Y^6 + \tfrac{1}{696729600}Y^8 + \tfrac{1189}{5670}XY \\
& + \tfrac{53}{360}XY^2 + \tfrac{827}{8640}XY^3 + \tfrac{25}{2304}XY^4 + \tfrac{53}{69120}XY^5 + \tfrac{1}{15360}XY^6 \\
& + \tfrac{1}{5806080}XY^7 + \tfrac{2377}{34020}X^2 + \tfrac{133}{1920}X^2Y^2 + \tfrac{7}{192}X^2Y^3 \\
& + \tfrac{271}{27648}X^2Y^4 + \tfrac{1}{768}X^2Y^5 + \tfrac{1}{184320}X^2Y^6 + \tfrac{19}{1080}X^3Y \\
& + \tfrac{11}{1152}X^3Y^2 + \tfrac{19}{4608}X^3Y^3 + \tfrac{19}{4608}X^3Y^4 + \tfrac{1}{18432}X^3Y^5 \\
& + \tfrac{79}{64800}X^4 + \tfrac{59}{55296}X^4Y^2 + \tfrac{1}{768}X^4Y^3 + \tfrac{37}{221184}X^4Y^4 \\
& + \tfrac{1}{8640}X^5Y + \tfrac{1}{15360}X^5Y^2 + \tfrac{1}{18432}X^5Y^3 + \tfrac{1}{311040}X^6 \\
& + \tfrac{1}{184320}X^6Y^2 + \tfrac{1}{5806080}X^7Y + \tfrac{1}{696729600}X^8,
\end{aligned}$$

$$\begin{aligned} P_{1,3} = P_{2,3} = {} & \tfrac{41}{1728}Y^2 + \tfrac{1}{155520}Y^4 + \tfrac{24967}{71680}X + \tfrac{323}{2880}XY + \tfrac{1}{288}XY^2 + \tfrac{1}{2592}XY^3 \\ & + \tfrac{15627}{143360}X^2 + \tfrac{2561}{34560}X^2Y + \tfrac{1}{288}X^2Y^2 + \tfrac{1}{7776}X^2Y^3 + \tfrac{197}{10240}X^3 \\ & + \tfrac{1}{72}X^3Y + \tfrac{1}{432}X^3Y^2 + \tfrac{1537}{614400}X^4 + \tfrac{1}{1152}X^4Y + \tfrac{1}{2592}X^4Y^2 \\ & + \tfrac{11}{92160}X^5 + \tfrac{1}{8640}X^5Y + \tfrac{1}{184320}X^6 + \tfrac{1}{311040}X^6Y \\ & + \tfrac{1}{5806080}X^7 + \tfrac{1}{696729600}X^8, \end{aligned}$$

$$\begin{aligned} P_{1,5} = P_{2,5} = {} & \tfrac{1}{600}Y^2 + \tfrac{490795}{1161216}X + \tfrac{1}{24}XY + \tfrac{1097645}{6967296}X^2 + \tfrac{7}{240}X^2Y \\ & + \tfrac{1409}{55296}X^3 + \tfrac{1}{120}X^3Y + \tfrac{1939}{663552}X^4 + \tfrac{1}{1200}X^4Y + \tfrac{13}{55296}X^5 \\ & + \tfrac{43}{4976640}X^6 + \tfrac{1}{5806080}X^7 + \tfrac{1}{696729600}X^8, \end{aligned}$$

$$P_{4,5} = \tfrac{1}{20}Y + \tfrac{185}{768}X + \tfrac{127}{4608}X^2 + \tfrac{1}{768}X^3 + \tfrac{1}{46080}X^4,$$

$$\begin{aligned} P_{1,7} = P_{2,7} = {} & \tfrac{71101}{165888}X + \tfrac{1}{28}XY + \tfrac{751447}{4976640}X^2 + \tfrac{1}{28}X^2Y + \tfrac{9799}{276480}X^3 \\ & + \tfrac{62299}{165888000}X^4 + \tfrac{13}{55296}X^5 + \tfrac{43}{4976640}X^6 + \tfrac{1}{5806080}X^7 \\ & + \tfrac{1}{696729600}X^8. \end{aligned}$$

Type F_4. We have

$$\begin{aligned} & P_1 = P_2 = \tfrac{23}{48}X + \tfrac{95}{576}X^2 + \tfrac{1}{48}X^3 + \tfrac{1}{1152}X^4, \\ & P_3 = P_6 = \tfrac{2}{9}X + \tfrac{1}{72}X^2, \\ & P_4 = \tfrac{3}{16}X + \tfrac{1}{96}X^2, \\ & P_8 = \tfrac{1}{8}X, \qquad P_{12} = \tfrac{1}{12}X, \end{aligned}$$

with the exceptions:

$$\begin{aligned} P_{1,2} = P_{2,2} = {} & \tfrac{1}{8}Y + \tfrac{19}{288}Y^2 + \tfrac{1}{32}Y^3 + \tfrac{1}{1152}Y^4 + \tfrac{1}{3}XY + \tfrac{1}{8}XY^2 + \tfrac{1}{48}XY^3 \\ & + \tfrac{1}{18}X^2 + \tfrac{1}{32}X^2Y + \tfrac{5}{64}X^2Y^2 + \tfrac{1}{48}X^3Y + \tfrac{1}{1152}X^4, \end{aligned}$$

$$P_{1,3} = P_{2,3} = \tfrac{1}{72}Y^2 + \tfrac{5}{16}X + \tfrac{1}{6}XY + \tfrac{7}{64}X^2 + \tfrac{1}{18}X^2Y + \tfrac{1}{48}X^3 + \tfrac{1}{1152}X^4.$$

Type G_2. We have

$$P_1 = P_2 = \tfrac{1}{2}X + \tfrac{1}{12}X^2 \qquad \text{and} \qquad P_3 = P_6 = \tfrac{1}{6}X,$$

with the exceptions

$$P_{1,2} = P_{2,2} = \tfrac{1}{12}Y^2 + \tfrac{1}{2}XY + \tfrac{1}{12}X^2$$

and

$$P_{1,3} = P_{2,3} = \tfrac{1}{6}Y + \tfrac{1}{2}X + \tfrac{1}{12}X^2.$$

Type ${}^3D_4$. We have

$$\begin{aligned} P_1 &= P_2 = \tfrac{1}{2}X + \tfrac{1}{12}X^2, \\ P_3 &= P_6 = \tfrac{1}{3}X + \tfrac{1}{24}X^2, \\ P_{12} &= \tfrac{1}{4}X, \end{aligned}$$

with the exceptions

$$P_{1,2} = P_{2,2} = \tfrac{1}{12}X^2 + \tfrac{1}{2}XY + \tfrac{1}{12}Y^2$$

and

$$P_{1,3} = P_{2,3} = \tfrac{1}{4}X + \tfrac{1}{4}XY + \tfrac{1}{24}Y^2 + \tfrac{1}{12}X^2Y.$$

Type ${}^2B_2$. We have

$$P_1 = \tfrac{1}{2}X \qquad \text{and} \qquad P_4 = \tfrac{1}{4}X.$$

Type ${}^2F_4$. We have

$$\begin{aligned} P_1 &= \tfrac{1}{2}X + \tfrac{1}{16}X^2, \\ P_2 &= \tfrac{1}{4}X + \tfrac{1}{48}X^2, \\ P_4 &= \tfrac{3}{16}X + \tfrac{1}{96}X^2, \\ P_6 &= \tfrac{1}{6}X, \qquad P_{12} = \tfrac{1}{12}X, \end{aligned}$$

with the exception

$$P_{2,3} = \tfrac{1}{4}X + \tfrac{1}{48}X^2 + \tfrac{1}{6}Y.$$

Type ${}^2G_2$. We have

$$P_1 = \tfrac{1}{2}X \qquad \text{and} \qquad P_2 = P_6 = \tfrac{1}{6}X,$$

with the exception

$$P_{2,2} = \tfrac{1}{6}X + \tfrac{1}{2}Y.$$

Remark. Let Ψ be an opposition automorphism of G commuting with F (cf. [Kaw]). Then $F' = F \circ \Psi$ is also a Frobenius endomorphism and many polynomials which express properties of the finite group $G^{F'}$ are obtained up to sign by substituting $-q$ for q in the corresponding polynomials for G^F. This procedure, known as *Ennola duality*, is easily shown to give information for the problem under consideration. Namely, let w_0 be the longest element in W. Then $F'(w) = w_0 F(w) w_0^{-1}$. Since $w_0^2 = 1$ and $F(w_0) = w_0$, the map $w \mapsto w_0 w$ induces a bijection from $W/\sim_F$ to $W/\sim_{F'}$ (indeed, $w_0 \cdot (xwF(x)^{-1}) = (w_0 x w_0^{-1})(w_0 w) F'(w_0 x w_0^{-1})^{-1}$). Moreover, if $C \in W/\sim_F$ and $C' = w_0 C \in W/\sim_{F'}$, then the action of F' on $X(T_{C'})$ corresponds to the action of F on $X(T_C)$. It follows that we have the polynomial identity $\chi_{C'}(X) = (-1)^l \chi_C(-X)$, where $l = \dim T_C$ is the rank of G. A similar relation holds for $|Z^F|$ and $|Z^{F'}|$. For example, if G is of type E_6 and F is a split Frobenius endomorphism, then $|Z^F| = (3, q-1)$ and $|Z^{F'}| = (3, q+1)$, where $q+1$ should be read as $-(-q-1)$. Note that the effect of Ennola duality on cyclotomic polynomials is to interchange φ_n and φ_{2n} if n is odd and to fix φ_n if $4|n$. If $-1 \in W$, then (G, F) and (G, F') are isomorphic and we get equalities of the form $P_m = P_{2m}$ for m odd. When $-1 \notin W$, as is the case for G of type E_6 in the tables above, we get equalities of the form $P_m = {}^2P_{2m}$ and $P_{2m} = {}^2P_m$ for m odd, as well as $P_m = {}^2P_m$ when $4|m$.

8. A LOWER BOUND FOR $\mu^W(G, F)$

In view of Theorem 6.2, the following result will complete the proof of Theorem 5.1:

THEOREM 8.1. *Let G be simple and simply connected and let h be the Coxeter number of G. If $p \neq$ char(k) is a prime factor of $|G^F/Z^F|$, then*

$$\mu^W(G, F) \geq \frac{1}{h}$$

except when $p = 3$, G is of type A_2, $|Z^F| = 3$, and $9 \nmid |G^F/Z^F|$, in which case $\mu^W(G, F) = \frac{1}{6}$.

Proof. The various possibilities for (G, F) are known up to isomorphism [Car5]. For the groups considered in Section 7, the possibilities for $\mu^W(G, F)$ can be read from the polynomials $P_{m,p}$. In all cases, a straightforward hand calculation yields $\mu^W(G, F) \geq 1/h$. (In many cases, one term of $P_{m,p}$ is large enough to prove the desired inequality.)

We are therefore left with the types A_l, ${}^2A_l$, B_l, C_l, D_l, and ${}^2D_l$. In each of these cases F is the Frobenius endomorphism corresponding to a definition of G over some subfield $\mathbb{F}_q$ of k.

Types A_l, ${}^2A_l$. Here $h = l + 1$ is the dimension of the underlying vector space. Let $\varepsilon = 1$ if (G, F) is split, $\varepsilon = -1$ if (G, F) is twisted. The F-conjugacy classes in W are parametrized by the partitions of h. If C_λ is the F-conjugacy class corresponding to the partition $\lambda = (\lambda_1, \lambda_2, \dots)$, then $|C_\lambda|$ coincides with the size of the conjugacy class determined by λ: either C_λ is that conjugacy class, or it is obtained from that conjugacy class by multiplication by w_0. Moreover, since $|Z^F| = (h, q - \varepsilon)$ we have

$$|T^F_{C_\lambda}/Z^F| = \{(q^{\lambda_1} - \varepsilon^{\lambda_1})(q^{\lambda_2} - \varepsilon^{\lambda_2}) \cdots\}/\{(q - \varepsilon)(h, q - \varepsilon)\} \tag{8.2}$$

(see, e.g., [Car4]). For every $m \geq 1$, $q^m - \varepsilon^m$ is a multiple of $q - \varepsilon$, hence also of $|Z^F|_p$.

Suppose first that $|Z^F|_p > 1$. By (8.2), every partition with at least three parts corresponds to a p-relevant F-conjugacy class in W. Looking at the remaining partitions, we get

$$1 - \mu^W \leq \frac{1}{h} + \frac{1}{2} \sum_{i=1}^{h-1} \frac{1}{i(h-i)}. \tag{8.3}$$

In particular, for $h = 4$ we have $\mu^W > \frac{1}{4}$. For $h \geq 5$, we have

$$\begin{aligned} \frac{1}{h} + \frac{1}{2} \sum_{i=1}^{h-1} \frac{1}{i(h-i)} &\leq \frac{1}{h} + \frac{1}{h-1} + \frac{1}{2} \sum_{i=2}^{h-2} \frac{1}{i(h-i)} \\ &< \frac{1}{h} + \frac{1}{h-1} + \frac{1}{2} \int_2^{h-1} \frac{dx}{(x-1)(h-x)} \\ &= \frac{1}{h} + \frac{1 + \log(h-2)}{h-1} \\ &\leq \frac{1}{5} + \frac{1 + \log(5-2)}{5-1} \leq \frac{3}{4}. \end{aligned} \tag{8.4}$$

Consequently, if $h \geq 4$ then $\mu^W \geq \frac{1}{4} \geq 1/h$.

If $h = 2$ then $p = 2$ (since $|Z^F|_p > 1$), one of the two classes of F-stable maximal tori is p-relevant, and therefore $\mu^W = \frac{1}{2}$ by definition. Suppose that $h = 3$. Then $p = 3$. If $(q - \varepsilon)_3 \geq 9$ then, by (8.2), $C_{(1,1,1)}$ and $C_{(2,1)}$ are p-relevant and therefore $\mu^W = \frac{1}{6} + \frac{3}{6} = \frac{2}{3}$. On the other hand, if $(q - \varepsilon)_3 = 3$ then $C_{(1,1,1)}$ is the only p-relevant F-conjugacy class, and therefore $\mu^W = \frac{1}{6}$.

Suppose now that $|Z^F|_p = 1$. Let m be the smallest positive integer such that $p | q^m - \varepsilon^m$.

Assume that $m \geq 2$. If m is a part of λ, then C_λ is p-relevant by (8.2). By using inclusion–exclusion to count the number of elements of W corresponding to such partitions λ, we find that

$$\mu^W \geq \sum_{1 \leq i \leq h/m} (-1)^{i-1} \frac{1}{i! m^i}. \tag{8.5}$$

If $m > h/2$, the right side is $1/m \geq 1/h$. If $m \leq h/2$, then

$$\mu^W \geq \frac{1}{m}\left(1 - \frac{1}{2m}\right) \geq \frac{1}{m} \cdot \frac{1}{2} \geq \frac{1}{h}.$$

Finally, assume that $m = 1$. Then $C_{(h)}$ is the only F-conjugacy class which is not p-relevant. Therefore

$$\mu^W = 1 - \frac{1}{h} \geq \frac{1}{h}.$$

Remark. One can do better than (8.5): by imitating the proof of [ET, Lemma I] in order to take into account all of the permutations having a cycle of length *divisible* by m, we find that $1 - \mu^W = \prod_{k=1}^{[h/m]}(1 - 1/km)$ when $m > 1$.

Types B_l, C_l, with $l \geq 2$. Here $h = 2l$, $|Z^F| = (2, q-1)$—so in particular $|Z^F|_p = 1$ if p is odd—and F acts trivially on W. The conjugacy classes in W are parametrized by all of the pairs (α, β) of partitions of l (so $|\alpha| + |\beta| = l$), and if C corresponds to (α, β), then (cf. [Car4])

$$|T_C^F| = (q^{\alpha_1} - 1)(q^{\alpha_2} - 1) \cdots (q^{\beta_1} + 1)(q^{\beta_2} + 1) \cdots. \tag{8.6}$$

There exist $\sigma = \pm 1$ and a positive integer $m \leq l$ such that $q^m - \sigma$ is a multiple of $p|Z^F|_p$. (We do not need to make any minimality assumption this time, although we will need to do so in the next section.) If $\sigma = 1$ (resp. -1), every pair (α, β) of partitions in which m is a part of α (resp. β) gives a p-relevant class. It follows that

$$\mu^W \geq \sum_{1 \leq i \leq l/m} \frac{(-1)^{i-1}}{i!(2m)^i}.$$

If $m > l/2$, the right hand side is $1/2m \geq 1/2l = 1/h$. If $m \leq l/2$, then

$$\mu^W \geq \frac{1}{2m}\left(1 - \frac{1}{4m}\right) \geq \frac{1}{2m} \cdot \frac{1}{2} \geq \frac{1}{h}.$$

Types $D_l, {}^2D_l$, with $l \geq 4$. Here $h = 2l - 2$. If $\varepsilon = 1$ when (G, F) is split and $\varepsilon = -1$ otherwise, then $|Z^F| = (4, q^l - \varepsilon)$. In particular, $|Z^F|_p = 1$ if p is odd. To every F-conjugacy class C in W one can associate a pair (α, β) of partitions of l, and (8.6) holds [Car4]. The pairs which occur in the split case are exactly those in which β has an even number of parts and the pairs which occur in the twisted case are exactly those in which β has an odd number of parts. In the split case, pairs in which all parts of α

are even and β is the empty partition correspond to two conjugacy classes in W; these two classes are simultaneously p-relevant or p-irrelevant (they are interchanged by a graph automorphism).

There exist $\sigma = \pm 1$ and a positive integer $m \leq l$ such that $q^m - \sigma$ is a multiple of $p|Z^F|_p$, with $\sigma = \varepsilon$ if $m = l$. If $\sigma = 1$ (resp. -1), every F-conjugacy class corresponding to a pair (α, β) of partitions in which m is a part of α (resp. β) is p-relevant. Counting the elements of W in these F-conjugacy classes, we get

$$\mu^W \geq \sum_{1 \leq i < l/m} \frac{(-1)^{i-1}}{i!(2m)^i} + R_W, \tag{8.7}$$

where

$$R_W = \begin{cases} 2\dfrac{(-1)^{(l/m)-1}}{(l/m)!(2m)^{l/m}} & \text{if } m|l \text{ and } \sigma^{l/m} = \varepsilon, \\ 0 & \text{otherwise.} \end{cases}$$

If $l = m$, then $\sigma = \varepsilon$ and $\mu^W \geq 1/l \geq 1/h$. If $l/2 < m \leq l - 1$, or if $m = l/2$ and $\sigma^2 \neq \varepsilon$, the right side of (8.7) is $1/2m \geq 1/(2l - 2) = 1/h$. If $m < l/2$, then

$$\mu^W \geq \frac{1}{2m}\left(1 - \frac{1}{4m}\right) \geq \frac{1}{2m} \cdot \frac{1}{2} \geq \frac{1}{h}.$$

It remains to consider the case where $m = l/2$ and $\sigma^2 = \varepsilon$. Here $q^m - \sigma | q^l - \varepsilon$, and we can instead use $m = l$, $\sigma = \varepsilon$. ∎

9. PROOF OF THEOREM 5.2

In this section we will prove Theorem 5.2, using the same method as that used for proving Theorem 8.1. Let ν_p be the lower bound indicated for μ_p in Theorem 5.2. It is straightforward to use the tables in Section 7 to check that $\mu(G^F/Z^F) \geq \nu_p$ when G is exceptional (not many cases need to be checked: since $h \leq 30$ the bound in Theorem 5.2 is better than the one in Theorem 5.1 only for small p). Moreover, we have $\mu_5(G^F/Z^F) = \frac{4}{25}$ if G is of type E_6 over $\mathbb{F}_q$ with $(q^2 + 1)_5 = 5$, and $\mu_{11}(G^F/Z^F) = \frac{9}{121}$ if G is of type E_8 over $\mathbb{F}_q$ with $(q^8 + q^6 + q^4 + q^2 + 1)_{11} = 11$ ($q = 2$ gives an example in each of these cases). Note also that $\mu_2(G^F/Z^F) = \frac{1}{4}$ for $G = PSL(2, q)$ with $(q^2 - 1)_2 = 8$, and that Theorem 5.1′ gives examples with $\mu_3(G^F/Z^F) = \frac{1}{9}$. Thus,

$$\mu_p \leq \nu_p \text{ for } p \in \{2, 3, 5, 11\}. \tag{9.1}$$

Now consider a simply connected group G of classical type, set $\mu = \mu_p(G^F/Z^F)$, $\mu^W = \mu_p^W(G, F)$ and define q as in Section 5. We will make use of the fact that the function $f(x) = x - \frac{1}{2}x^2$, $x \leq 1$, is increasing. Note also that $f(1/p) = \nu_p$ if $p \notin \{2,3,5,11\}$ and $f(1/p) \geq \nu_p$ in all cases.

Types $A_l, {}^2A_l$. Let $h = l + 1$ and ε be as in Section 8. Suppose first that $|Z^F|_p \geq p$. If $h \geq 4$, we have seen in Section 8 that $\mu^W \geq \frac{1}{4}$, and for $p \neq 2$ it follows that $\mu \geq (1 - 1/p)(\frac{1}{4}) \geq \nu_p$. By Theorem 5.1′ we have $\mu \geq \nu_p$ if $h = p = 3$. Now assume that $p = 2$, so that h is even. We claim that $\mu^W \geq \frac{1}{2}$ if $h \neq 4$. By (8.4) this holds for $h \geq 10$ since then $1 - \mu^W \leq \frac{1}{10} + (1 + \log 8)/9 \leq \frac{1}{2}$. Also, $\mu^W = \frac{1}{2}$ if $h = 2$, we can use (8.3) if $h = 8$, and if $h = 6$ we can use the additional fact that the partition $(4,2)$ corresponds to a 2-relevant F-conjugacy class (in view of (8.2) and the fact that $|Z^F|_2 = 2$). This proves our claim, and hence by Theorem 6.2 we have $\mu \geq \frac{1}{2}\mu^W \geq \nu_2$ for $h \neq 4$. Finally, if $p = 2$ and $h = 4$, the formulas for $\mu_2(G^F/Z^F)$ given in Section 6 show that $\mu \geq \frac{45}{128} \geq \nu_2$ when $|Z^F| = 4$ and $\mu \geq \frac{71}{128} \geq \nu_2$ when $|Z^F| = 2$.

Suppose now that $|Z^F|_p = 1$. Let m be the smallest positive integer such that $p | q^m - \varepsilon^m$. Note that $m \leq p - 1$. If $m = 1$ then we saw in Section 8 that $\mu^W \geq 1 - 1/h$, so that $\mu \geq \nu_p$ by Theorem 6.2 and the definition of ν_p. Suppose that $m \geq 2$. If λ has r parts equal to m, then $\mu(T^F_{C_\lambda}/Z^F) \geq 1 - 1/p^r$ (cf. Lemma 2.1). By a variation on the inclusion–exclusion formula we find that

$$\mu \geq \sum_{1 \leq i \leq h/m} (-1)^{i-1} \frac{1}{i! m^i} \left(1 - \frac{1}{p}\right)^i .$$

Therefore

$$\mu \geq \frac{1}{m}\left(1 - \frac{1}{p}\right) - \frac{1}{2m^2}\left(1 - \frac{1}{p}\right)^2 = f\left(\frac{1}{m}\left(1 - \frac{1}{p}\right)\right)$$
$$\geq f\left(\frac{1}{p-1}\left(1 - \frac{1}{p}\right)\right) = f\left(\frac{1}{p}\right) = \frac{1}{p} - \frac{1}{2p^2} \geq \nu_p .$$

Types B_l, C_l, *with* $l \geq 2$. Suppose first that p is odd. Let $m \geq 1$ and $\sigma = \pm 1$ be such that $p | q^m - \sigma$, with m as small as possible. Note that $m \leq (p - 1)/2$. We have

$$\mu \geq \sum_{1 \leq i \leq l/m} (-1)^{i-1} \frac{1}{i!(2m)^i} \left(1 - \frac{1}{p}\right)^i ,$$

so that

$$\mu \geq \frac{1}{2m}\left(1-\frac{1}{p}\right)-\frac{1}{8m^2}\left(1-\frac{1}{p}\right)^2=f\left(\frac{1}{2m}\left(1-\frac{1}{p}\right)\right)$$
$$\geq f\left(\frac{1}{2(p-1)/2}\left(1-\frac{1}{p}\right)\right)=f\left(\frac{1}{p}\right)=\frac{1}{p}-\frac{1}{2p^2}\geq \nu_p.$$

Suppose that $p=2$. Since q is odd and $|Z^F|=2$, it follows from (8.6) that any conjugacy class in W corresponding to a pair (α, β) of partitions in which α and β together have at least two parts is necessarily 2-relevant. Therefore $\mu^W \geq 1-1/n \geq \frac{1}{2}$, and hence $\mu \geq \frac{1}{4}$ by Theorem 6.2.

Types $D_l, {}^2D_l$, with $l \geq 4$. Suppose first that p is odd. Let ε, σ, and $m \leq l$ be as in Section 8, with $\sigma=\varepsilon$ if $m=l$, and m minimal subject to these conditions. In particular, $m \leq (p-1)/2$. We have

$$\mu \geq \sum_{1\leq i<l/m}\frac{(-1)^{i-1}}{i!(2m)^i}\left(1-\frac{1}{p}\right)^i+R,$$

where

$$R=\begin{cases} 2\dfrac{(-1)^{(l/m)-1}}{(l/m)!(2m)^{l/m}}\left(1-\dfrac{1}{p}\right)^{l/m} & \text{if } m|l \text{ and } \sigma^{l/m}=\varepsilon,\\ 0 & \text{otherwise.}\end{cases}$$

If $m=l$, then $\sigma=\varepsilon$ and $\mu \geq R=(1-1/p)/m \geq 2/p \geq \nu_p$.

If $m \leq l-1$ and $m \neq l/2$, or if $m=l/2$ and $\sigma^2 \neq \varepsilon$, then, as in the case of type B_l, we have

$$\mu \geq \sum_{1\leq i\leq l/m}(-1)^{i-1}\frac{1}{i!(2m)^i}\left(1-\frac{1}{p}\right)^i$$
$$\geq \frac{1}{2m}\left(1-\frac{1}{p}\right)-\frac{1}{8m^2}\left(1-\frac{1}{p}\right)^2=f\left(\frac{1}{2m}\left(1-\frac{1}{p}\right)\right)\geq f\left(\frac{1}{p}\right)\geq \nu_p.$$

If $m=l/2$ and $\varepsilon=1$, then $l \leq p-1$. We have already seen in Section 8 that $\mu^W \geq 1/l$, and therefore $\mu \geq 1/p \geq \nu_p$.

Suppose now that $p = 2$. Recall that $|Z^F| \leq 4$. If an F-conjugacy class C in W corresponds to the pair (α, β) of partitions, let $r(C)$ denote the total number of parts of α and β. For $r \geq 1$, let W_r be the union of the F-conjugacy classes C with $r(C) = r$, and let $W_{\geq 3}$ be the union of the W_r's with $r \geq 3$. We have $|W_1| = |W|/l \leq |W|/4$, and it follows from (8.3) and (8.4) that $|W_{\geq 3}| \geq |W|/4$. In view of (8.6), every element in $W_{\geq 3}$ is 2-relevant, and so are at least half of the elements in W_2. This implies that $\mu^W \geq \frac{1}{2}$ and therefore $\mu \geq \frac{1}{4}$.

We have checked that $\mu \geq \nu_p$ in all cases. In view of (9.1), this proves Theorem 5.2. ▌

10. EQUAL CHARACTERISTIC CASE

Throughout this section the notation and the assumptions will be the same as those in section 5, except that p is now assumed to be the characteristic s of G.

THEOREM 10.1. $\mu_p(G^F) \geq \frac{2}{5}q^{-8}$.

Remarks. We need this only in the case where G^F/Z^F is simple and hence only when G is simply connected. The result holds as long as G is reductive, connected, but not just a torus. However, the case presented here is slightly easier to prove.

The constant $\frac{2}{5}$ is not best possible. It is likely that it could be replaced by $\frac{3}{4}$.

The proof is based on a case by case analysis together with the next proposition. For $t \in G^F$ semisimple, let $n_G(t) = \dim C_G(t) - l(G)$, where $l(G)$ is the rank of G, and let $f_G(t) = 1 - q^{-n_G(t)}$. If T is a maximal torus containing t, then $n_G(t)$ is the number of roots α of T in G such that $\alpha(t) = 1$ (compare [Car5, p. 92]). For an F-stable maximal torus T in G, let $\varphi_G(T)$ be the average of the function f_G on T^F. If T' is G^F-conjugate to T, then $\varphi_G(T') = \varphi_G(T)$. Therefore we can also view φ_G as a function on W, constant on F-conjugacy classes: $\varphi_G(w) = \varphi_G(T_C)$ where C is the F-conjugacy class containing w.

PROPOSITION 10.2. $\mu_p(G^F) = (1/|W|)\Sigma_{w \in W}\varphi_G(w)$.

Proof. Every element $x \in G^F$ has a Jordan decomposition $x = tu = ut$ with $t \in G^F$ semisimple and $u \in G^F$ unipotent, and x is p-singular if and only if $u \neq 1$. There are exactly $q^{n_G(t)}$ unipotent elements in G^F that commute with t (compare the proof of Theorem 6.2′). The total number of p-singular elements in G^F is therefore equal to the sum of $q^{n_G(t)} - 1$ over all semisimple elements $t \in G^F$. As there are exactly $q^{n_G(t)}$ F-stable maximal tori in G that contain t, this is also the sum of $f_G(t)$ over all pairs (t, T) consisting of a semisimple element $t \in G^F$ and an F-stable maximal torus T containing t. Thus, exactly as in the proof of Theorem 6.2′ (cf. (6.1)),

$$\begin{aligned}\mu_p(G^F) &= \frac{1}{|G^F|}\sum_{(t,T)} f_G(t) = \frac{1}{|G^F|}\sum_{T}|T^F|\varphi_G(T)\\ &= \frac{1}{|G^F|}\sum_{C\in W/\sim_F}\ \sum_{T\in\mathscr{T}_C}|T^F|\varphi_G(T)\\ &= \frac{1}{|G^F|}\sum_{C\in W/\sim_F}\varphi_G(T_C)\sum_{T\in\mathscr{T}_C}|T^F|\\ &= \sum_{C\in W/\sim_F}\frac{|C|}{|W|}\varphi_G(T_C) = \frac{1}{|W|}\sum_{w\in W}\varphi_G(w). \quad ▮\end{aligned}$$

EXAMPLES.

(a) Let $G = SL_2$. Let $z = (q-1, 2) = |Z^F|$. We have $|W| = 2$. The identity corresponds to a maximal torus $T_{(1,1)}$ with $|T^F_{(1,1)}| = q-1$. Let $t \in T^F_{(1,1)}$. If t is central then $n_G(t) = 2$, and $n_G(t) = 0$ otherwise. It follows that $\varphi_G(T_{(1,1)}) = z(1-q^{-2})/(q-1) = z(1+q^{-1})q^{-1}$.

The other element in W corresponds to a maximal torus $T_{(2)}$ with $|T^F_{(2)}| = q+1$, and $\varphi_G(T_{(2)}) = z(1-q^{-2})/(q+1) = z(1-q^{-1})q^{-1}$. The average of $z(1+q^{-1})q^{-1}$ and $z(1-q^{-1})q^{-1}$ is zq^{-1}. Thus, by the proposition, $\mu_p(G^F) = (q-1, 2)q^{-1}$. (Of course, it is easy to check directly that there are $z(q^2-1)$ p-singular elements in G^F.)

(b) Let $G = SL_3$ and $z = (3, q-1) = |Z^F|$.

Let $T_{(1,1,1)}$ be the diagonal subgroup of SL_3. If $t \in T_{(1,1,1)}$ is $\operatorname{diag}(a, b, c)$ then $abc = 1$ and $F(t) = \operatorname{diag}(a^q, b^q, c^q)$. In particular, $|T^F_{(1,1,1)}| = (q-1)^2$. Moreover, $n_G(t) = 0$ if a, b, and c are all distinct, $n_G(t) = 2$ if two of a, b, c are equal, and $n_G(t) = 6$ if $a = b = c$. Thus, there are z elements $t \in T^F_{(1,1,1)}$ such that $n_G(t) = 6$ and $3(q-z-1)$ such that $n_G(t) = 2$, while $n_G(t) = 0$ for all remaining elements in $T^F_{(1,1,1)}$.

Let $T_{(2,1)}$ be an F-stable maximal torus corresponding to a transposition. Choosing a basis consisting of eigenvectors, we see that $T_{(2,1)}$ is represented by diagonal matrices $\operatorname{diag}(a, b, c)$ with $abc = 1$. After permuting the elements of the basis if necessary, F induces the map $\operatorname{diag}(a, b, c) \mapsto \operatorname{diag}(b^q, a^q, c^q)$. Let $t \in T_{(2,1)}$ be $\operatorname{diag}(a, b, c)$. Then $t \in T^F_{(2,1)}$ if and only if $a \in \mathbb{F}^*_{q^2}$, $b = a^q$, and $c = a^{-(q+1)}$. In particular $|T^F_{(2,1)}| = q^2 - 1$. Moreover, $n_G(t) = 0$ if $a \notin \mathbb{F}_q$, $n_G(t) = 2$ if $a \in \mathbb{F}_q$ and $a^3 \neq 1$, and

$n_G(t) = 6$ if $a \in \mathbb{F}_q$ and $a^3 = 1$. Thus, there are exactly z elements $t \in T^F_{(2,1)}$ with $n_G(t) = 6$ and $q - 1 - z$ with $n_G(t) = 2$, while $n_G(t) = 0$ for the remaining elements.

Let $T_{(3)}$ be an F-stable maximal torus corresponding to a 3-cycle. Choosing a basis consisting of eigenvectors of $T_{(3)}$, we see that $T_{(3)}$ is represented by diagonal matrices $\operatorname{diag}(a, b, c)$ such that $abc = 1$. After permuting the elements of the basis if necessary, F induces the map $\operatorname{diag}(a, b, c) \mapsto \operatorname{diag}(b^q, c^q, a^q)$. Let $t \in T_{(3)}$ be $\operatorname{diag}(a, b, c)$. Then $t \in T^F_{(3)}$ if and only if $a^{q^2+q+1} = 1$, $b = a^q$, and $c = a^{q^2}$. In particular $|T^F_{(3)}| = q^2 + q + 1$. Moreover, $n_G(t) = 0$ unless $a \in \mathbb{F}_q$ and $a^3 = 1$, in which case $n_G(t) = 6$. Thus, there are exactly z elements $t \in T^F_{(3)}$ with $n_G(t) = 6$, and $n_G(t) = 0$ for the remaining elements.

It follows that

$$\begin{aligned}
\mu_p(G^F) &= \frac{1}{6}\left\{\varphi_G(T_{(1,1,1)}) + 3\varphi_G(T_{(2,1)}) + 2\varphi_G(T_{(3)})\right\} \\
&= \frac{1}{6}\Big\{\big(z(1 - q^{-6}) + 3(q - z - 1)(1 - q^{-2})\big)(q - 1)^{-2} \\
&\qquad + 3\big(z(1 - q^{-6}) + (q - z - 1)(1 - q^{-2})\big)(q^2 - 1)^{-1} \\
&\qquad + 2\big(z(1 - q^{-6})\big)(q^2 + q + 1)^{-1}\Big\} \\
&= q^{-1} - zq^{-3} = \left(1 - \frac{(q - 1, 3)}{q^2}\right)q^{-1}.
\end{aligned}$$

We turn now to the proof of Theorem 10.1. For the sake of clarity we will assume for now that $\delta = 1$, and indicate later how the argument can be modified to cover the Ree and Suzuki groups.

Our task is to find sufficiently many elements in W for which $\varphi_G(w)$ is large enough. Consider an F-stable maximal torus T. Let α be a root of T in G and let $(\operatorname{Ker}(\alpha))^0 = S$. Then $n_G(t) \geq 2$ for every $t \in S$ since $-\alpha$ also has S in its kernel. Throughout the proof we will consider those roots α fixed or inverted by F (cf. Appendix A.9). If F fixes α then $|S^F|/|T^F| = 1/(q - 1)$, and it follows that $\varphi_G(T) \geq (1 - q^{-2})|S^F|/|T^F| = (1 + q^{-1})q^{-1}$. Similarly, if F inverts α, then $|S^F|/|T^F| = 1/(q + 1)$, and $\varphi_G(T) \geq (1 - q^{-2})(q + 1)^{-1} = (1 - q^{-1})q^{-1}$. In many cases this approximation is sufficient to prove Theorem 10.1. First we consider the easy cases.

Types B_l, C_l. We can think of W as a group of linear transformations of $\mathbb{Q}^l$ consisting of transformations obtained by permuting the coordinates and multiplying any number of them by -1. The roots correspond to the vectors $\pm e_i$ $(1 \leq i \leq l)$ and $\pm e_i \pm e_j$ $(1 \leq i < j \leq l)$, and the action of F

on a torus corresponding to the element w: $\mathbb{Q}^l \to \mathbb{Q}^l$ is qw^{-1} (see Appendix A.7). Thus, if there exists i such that $w(e_i) = e_i$, then $\varphi_G(w) \geq (1 + q)q^{-1}$; and if there exists i such that $w(e_i) = -e_i$, then $\varphi_G(w) \geq (1 - q)q^{-1}$. Let x_l be the probability that an element in W fixes at least one basis vector e_i, and let y_l be the probability that an element in W inverts at least one basis vector e_i but does not fix any e_i. Then $x_l \geq y_l$ (if $w \in W$ inverts some basis vector e_i then $-w$ fixes e_i) and $\mu_p(G^F) \geq x_l(1 + q^{-1})q^{-1} + y_l(1 - q^{-1})q^{-1} \geq (x_l + y_l)q^{-1}$. But $x_l + y_l$ is the probability that an element in W fixes or inverts at least one basis vector, and this is the same as the probability that an element of S_l fixes at least one of the l points it permutes. For each l, this probability is at least $\frac{1}{2}$. Thus $\mu_p(G^F) \geq \frac{1}{2}q^{-1}$.

Type G_2. The only elements that do not fix at least one root are the rotations. Thus, half of the elements in W fix at least one root, and $\mu_p(G^F) \geq \frac{1}{2}(1 + q^{-1})q^{-1} \geq \frac{1}{2}q^{-1}$.

Type F_4. An explicit computation shows that 575 of the 1152 elements of W fix at least one root. It follows that

$$\mu_p(G^F) \geq \tfrac{575}{1152}(1 + q^{-1})q^{-1} \geq \tfrac{575}{1152}q^{-1} \geq \tfrac{2}{5}q^{-1}.$$

Type E_6. An explicit computation shows that 17,371 of the 51,840 elements of W fix at least one root and that 10,809 elements fix no root but invert at least one. It follows that

$$\mu_p(G^F) \geq \tfrac{17371}{51840}(1 + q^{-1})q^{-1} + \tfrac{10809}{51840}(1 - q^{-1})q^{-1} \geq \tfrac{28180}{51840}q^{-1} \geq \tfrac{2}{5}q^{-1}.$$

Type E_7. In this case 952,435 of the 2,903,040 elements of W fix at least one root and 733,069 elements fix no root but invert at least one. It follows that

$$\begin{aligned}\mu_p(G^F) &\geq \tfrac{952435}{2903040}(1 + q^{-1})q^{-1} + \tfrac{733069}{2903040}(1 - q^{-1})q^{-1} \geq \tfrac{1685504}{2903040}q^{-1} \\ &\geq \tfrac{2}{5}q^{-1}.\end{aligned}$$

Type E_8. In this case 228,350,039 of the 696,729,600 elements of W fix at least one root and 150,831,449 elements fix no root but invert at least one. It follows that

$$\begin{aligned}\mu_p(G^F) &\geq \tfrac{228350039}{696729600}(1 + q^{-1})q^{-1} + \tfrac{150831449}{696729600}(1 - q^{-1})q^{-1} \geq \tfrac{379181488}{696729600}q^{-1} \\ &\geq \tfrac{2}{5}q^{-1}.\end{aligned}$$

Type ${}^2E_6$. In this case the action of F involves multiplication by $-q$ instead of q as in the case of type E_6. It follows that to detect F-stable maximal tori with roots fixed (resp. inverted) by F, we must consider elements of the ordinary Weyl group of type E_6 that invert (resp. fix) some roots. There are 16,335 elements that invert at least one root and 11,845 that do not invert any root but fix at least one of them. This leads to the estimate

$$\mu_p(G^F) \geq \tfrac{16335}{51840}(1 + q^{-1})q^{-1} + \tfrac{11845}{51840}(1 - q^{-1})q^{-1} \geq \tfrac{28180}{51840}q^{-1} \geq \tfrac{2}{5}q^{-1}.$$

Type ${}^3D_4$. In this case the action of F can be described using a coset determining an element of order 3 of $W(F_4)/W(D_4) \cong S_3$ (see Appendix A.5). It turns out that 64 of the 192 elements of W give fixed roots, 64 elements give inverted roots, and there is no overlap. Therefore we get

$$\mu_p(G^F) \geq \tfrac{64}{192}(1 + q^{-1})q^{-1} + \tfrac{64}{192}(1 - q^{-1})q^{-1} = \tfrac{2}{3}q^{-1} \geq \tfrac{2}{5}q^{-1}.$$

When $\delta = 1$ we are left with the harder cases: groups of type A or D.

Types $A_l, {}^2A_l$. Let $h = l + 1$. We identify W with S_h. Let $w \in S_h$. In the split (resp. twisted) case, w corresponds to tori in which F fixes (resp. inverts) at least one root if and only if w has two 1-cycles, and F inverts (resp. fixes) at least one root if and only if w has at least one 2-cycle.

First, consider the split case. Let a_h be the probability that a permutation in S_h has no 1-cycle and b_h the probability that a permutation in S_h has no 1- or 2-cycle. The probability x_h that an element $w \in S_h$, has at least two 1-cycles is then $1 - a_h - a_{h-1}$, and the probability y_h that w has at most one 1-cycle and at least one 2-cycle is $a_h + a_{h-1} - b_h - b_{h-1}$. From the inclusion–exclusion formula we obtain

$$a_h = \sum_{i=0}^{h} \frac{(-1)^i}{i!}, \qquad b_h = \sum_{\substack{i,j \geq 0 \\ i+2j \leq h}} \frac{(-1)^{i+j}}{2^j i! j!}.$$

As $h \to \infty$, $a_h \to e^{-1}$ and $b_h \to e^{-3/2}$. These imply that $\frac{11}{30} \leq a_h \leq \frac{3}{8}$ and $b_h \leq \frac{1}{4}$ for large enough h, and it is easily checked that these inequalities hold for all $h \geq 4$. It follows that $x_h \geq \frac{1}{4}$ and $y_h \geq \frac{7}{30}$ for $h \geq 5$. For $h \geq 5$ we thus have $\mu_p(G^F) \geq \frac{1}{4}(1 + q^{-1})q^{-1} + \frac{7}{30}(1 - q^{-1})q^{-1} \geq \frac{29}{60}q^{-1} \geq \frac{2}{5}q^{-1}$. For $h = 4$ we have $x_h = \frac{7}{24}$ and $y_h = \frac{1}{8}$, and therefore $\mu_p(G^F) \geq \frac{7}{24}(1 + q^{-1})q^{-1} + \frac{1}{8}(1 - q^{-1})q^{-1} \geq \frac{5}{12}q^{-1} \geq \frac{2}{5}q^{-1}$. For $h = 2$, $x_h = y_h = \frac{1}{2}$ and therefore $\mu_p(G^F) \geq q^{-1}$. Finally, for $h = 3$, $x_h = \frac{1}{6}$, $y_h = \frac{1}{2}$, and $\mu_p(G^F) \geq \frac{1}{6}(1 + q^{-1})q^{-1} + \frac{1}{2}(1 - q^{-1})q^{-1} = (\frac{2}{3} - \frac{1}{3}q^{-1})q^{-1} \geq (\frac{2}{3} - \frac{1}{6})q^{-1} = \frac{1}{2}q^{-1}$.

In the twisted case, let ${}^2x_h$ be the probability that an element of S_h has at least one 2-cycle, ${}^2y_h$ the probability that an element of S_h has at least two 1-cycles and no 2-cycle, and c_h the probability that an element of S_h has no 2-cycle. Let b_h be as above. Then ${}^2x_h = 1 - c_h$, ${}^2y_h = c_h - b_h - b_{h-1}$, and

$$c_h = \sum_{j=0}^{[h/2]} \frac{(-1)^j}{2^j j!}.$$

It follows easily that $\frac{1}{2} \leq c_h \leq \frac{5}{8}$ for $h \geq 3$. Since $b_h \leq \frac{1}{4}$ for $h \geq 4$, while $b_2 = 0$ and $b_3 = \frac{1}{3}$, we have $b_h + b_{h-1} \leq \frac{7}{12}$ for every $h \geq 3$. Then

$$\begin{aligned}\mu_p(G^F) &\geq {}^2x_h(1+q^{-1})q^{-1} + {}^2y_h(1-q^{-1})q^{-1}\\ &= (1-c_h)(1+q^{-1})q^{-1} + (c_h - b_h - b_{h-1})(1-q^{-1})q^{-1}\\ &= (1-\tfrac{5}{8})(1+q^{-1})q^{-1} + (\tfrac{5}{8} - b_h - b_{h-1})(1-q^{-1})q^{-1}\\ &\quad + 2(\tfrac{5}{8} - c_h)q^{-2}\\ &\geq \tfrac{3}{8}(1+q^{-1})q^{-1} + \tfrac{1}{24}(1-q^{-1})q^{-1} \geq \tfrac{5}{12}q^{-1} \geq \tfrac{2}{5}q^{-1}.\end{aligned}$$

Type D_l or ${}^2D_l$. The method used so far fails because when the Frobenius endomorphism of an F-stable maximal torus fixes or inverts a pair of roots, it fixes or inverts at least one other pair of roots. There are therefore fewer elements in W that correspond to such a behavior of F.

Let T be an F-stable maximal torus. There is an isomorphism of the character group $X(T) = \mathrm{Hom}(T, \mathbf{G}_m)$ of T with the subgroup of $\mathbb{Q}^l$ generated by the standard basis $(e_i)_{1 \leq i \leq l}$ and the vector $\frac{1}{2}(1, 1, \ldots, 1)$, and such that the roots correspond to the vectors $\pm e_i \pm e_j$ $(1 \leq i < j \leq l)$. Let $\alpha_{\pm i \pm j}$ be the root corresponding to $\pm e_i \pm e_j$. The Frobenius endomorphism F acts on $X(T)$ by multiplication by q composed with a signed permutation σ_T of the basis that is determined up to conjugacy under W (see Appendix A.7). The group $Y(T) = \mathrm{Hom}(\mathbf{G}_m, T)$ is dual to $X(T)$ (over $\mathbb{Z}$) and can be identified with the subgroup of $\mathbb{Q}^l$ generated by the vectors $\pm e_i \pm e_j$ $(1 \leq i < j \leq l)$. These vectors correspond in $Y(T)$ to the coroots $\alpha^\vee_{\pm i \pm j}$ [Bou, pp. 256–257].

With this description, the Weyl group W becomes a subgroup of index 2 in a Weyl group $\tilde{W}_l$ of type B_l. Let $W_l^+ = W$ and $W_l^- = \tilde{W}_l - W$ be the two cosets. Note that $\tilde{W}_l$, W, and $W_l^{\pm}$ can all be defined in this manner for $l \geq 1$. When dealing with F-conjugacy classes and the action of F on F-stable maximal tori and their character groups, it is natural to replace W by a suitable coset of W in $\tilde{W}$, namely by W_l^- in the twisted case, and W_l^+ $(= W)$ in the split case (see Appendix A.5). We handle the two cases simultaneously, and we let ε stand for $+$ or $-$ for G split or twisted, respectively.

Suppose that T and σ are as before, and that $\sigma_T \in W_l^\varepsilon$ stabilizes the subset $\{e_1, -e_1, e_2, -e_2\}$. Then the subgroup S of T generated by the images in T of the coroots $\alpha^\vee_{\pm i \pm j}$ with $3 \le i < j \le l$ is an F-stable subtorus of T. Consider the quotient group T/S. Its character group $X(T/S)$ can be identified with the subgroup of $X(T)$ consisting of all elements orthogonal to the vectors $\pm e_i \pm e_j$, $3 \le i < j \le l$. Thus, $X(T/S)$ is the subgroup generated by e_1 and e_2. The subtorus S is contained in the kernel of the roots α_{1+2} and α_{1-2}. By Lang's theorem the natural homomorphism $T^F \to (T/S)^F$ is surjective and its kernel is S^F (Appendix A.1). Let K be the union of the kernels of the characters of T/S induced by α_{1-2} and α_{1+2}. Observe that $\mathrm{Ker}(\alpha_{1-2})/S$ and $\mathrm{Ker}(\alpha_{1+2})/S$ look like the subgroups $\{\mathrm{diag}(t,t) | t \in k^*\}$ and $\{\mathrm{diag}(t,t^{-1}) | t \in k^*\}$ of GL_2, respectively. Then $\varphi_G(T) \ge (|K^F|/|(T/S)^F|)(1 - q^{-2})$. Moreover, since K is the union of two one-dimensional tori that intersect in $(q-1, 2) \le 2$ points, we get the following.

(i) If $\sigma(1) = 1$ and $\sigma(2) = 2$, then $\varphi_G(T) \ge ((2(q-1) - 2)/(q-1)^2) \cdot (1 - q^{-2})$.

(ii) If $\sigma(1) = 2$ and $\sigma(2) = 1$, then $\varphi_G(T) \ge (((q-1) + (q+1) - 2)/(q^2 - 1)) \cdot (1 - q^{-2})$.

(iii) If $\sigma(1) = -1$ and $\sigma(2) = -2$, then $\varphi_G(T) \ge ((2(q+1) - 2)/(q+1)^2) \cdot (1 - q^{-2})$.

If $q = 2$ these values can be replaced by $\frac{15}{16}$, $\frac{13}{16}$, and $\frac{7}{16}$, respectively, as can be seen by direct computation. For example, in case (iii), $|(T/S)^F| = 9$, each of the two kernels which constitute K contains three F-stable points, and the identity is the only common point. Thus $|K^F| = 5$, and $n_G(s) \ge 4$ for every $s \in S$, so that $\varphi_G(T) \ge ((1 - 2^{-4}) + 4 \cdot (1 - 2^{-2}))/9 = \frac{7}{16}$.

Given a signed permutation $\sigma \in \tilde{W}$, let $m_i^+(\sigma)$ and $m_i^-(\sigma)$ be respectively the numbers of positive and negative cycles of σ of length i. The discussion above shows the following.

(i) If $m_1^+(\sigma) \ge 2$, then $\varphi_G(\sigma) \ge ((2q - 4)/(q-1)^2) \cdot (1 - q^{-2})$.

(ii) If $m_2^+(\sigma) \ge 1$, then $\varphi_G(\sigma) \ge ((2q - 2)/(q^2 - 1)) \cdot (1 - q^{-2})$.

(iii) If $m_1^-(\sigma) \ge 2$, then $\varphi_G(\sigma) \ge (2q/(q+1)^2) \cdot (1 - q^{-2})$.

If $q = 2$ these values can be replaced by $\frac{15}{16}$, $\frac{13}{16}$, and $\frac{7}{16}$, respectively.

For $l \ge 0$, let $w_l = 2^{l-1} l!$. Then $w_l = |W_l^+| = |W_l^-|$ for $l \ge 1$. Define $W_0^+ = 1$ and $W_0^- = \varnothing$, so that W_l^ε is now defined for all $l \ge 0$. For a subset X of W_l^ε, let $P_l^\varepsilon(X) = |X|/w_l$ (thus, for $l \ge 1$, P_l^ε is the obvious probability measure on W_l^ε). Let

$$x_l^\varepsilon = P_l^\varepsilon\{\sigma \mid m_1^+(\sigma) \geq 2\}, \qquad y_l^\varepsilon = P_l^\varepsilon\{\sigma \mid m_1^+(\sigma) \leq 1, m_2^+(\sigma) \geq 1\},$$
$$z_l^\varepsilon = P_l^\varepsilon\{\sigma \mid m_1^+(\sigma) \leq 1, m_2^+(\sigma) = 0 \text{ and } m_1^-(\sigma) \geq 2\}.$$

In view of (i)–(iii), we have

$$\mu_p(G^F) \geq x_l^\varepsilon \frac{2q-4}{(q-1)^2} \cdot (1 - q^{-2}) + y_l^\varepsilon \frac{2q-2}{q^2-1} \cdot (1 - q^{-2})$$
$$+ z_l^\varepsilon \frac{2q}{(q+1)^2} \cdot (1 - q^{-2})$$

if $q \geq 3$, and

$$\mu_p(G^F) \geq x_l^\varepsilon \cdot \tfrac{15}{16} + y_l^\varepsilon \cdot \tfrac{3}{16} + z_l^\varepsilon \cdot \tfrac{7}{16}$$

if $q = 2$. In order to prove that $\mu_p(G^F) \geq \frac{2}{5} q^{-1}$, it is therefore enough to prove that

$$x_l^\varepsilon(2q-4)(q+1)^2 + y_l^\varepsilon(2q-2)(q^2-1) + z_l^\varepsilon(2q)(q-1)^2$$
$$\geq \tfrac{2}{5} q(q^2-1), \tag{10.3}$$

if $q \geq 3$, or

$$\tfrac{15}{16} x_l^\varepsilon + \tfrac{13}{16} y_l^\varepsilon + \tfrac{7}{16} z_l^\varepsilon \geq \tfrac{1}{5} \tag{10.3$'$}$$

if $q = 2$. It is easily checked that

$$x_4^+ = \tfrac{19}{192}, \qquad y_4^+ = \tfrac{1}{8}, \qquad z_4^+ = \tfrac{1}{192}, \qquad x_4^- = \tfrac{1}{12}, \qquad y_4^- = \tfrac{1}{4}, \qquad z_4^- = \tfrac{1}{12}.$$

When $l = 4$, an easy computation shows that (10.3) or (10.3$'$) holds, except in the split case if $q \leq 5$. Explicit computations using tori of codimension 2 (instead of codimension 1) show that, if σ is an element of W_4^+ corresponding to a positive 4-cycle, then $\varphi_G(\sigma) \geq 2q^{-2} - q^{-4}$. There are 48 such elements in W. Taking their contribution into account gives the required lower bound for $\mu_p(G^F)$ in the remaining cases. (When q is 2 or 3, one can also use [CCNPW] to determine $\mu_p(G^F)$ exactly.)

Suppose now that $l \geq 5$. It is easily checked that if

$$x_l^\varepsilon \geq \tfrac{1}{12}, \qquad y_l^\varepsilon \geq \tfrac{11}{60}, \qquad z_l^\varepsilon \geq \tfrac{1}{20}, \tag{10.4}$$

then (10.3) and (10.3′) hold. We claim that (10.4) actually holds for every $l \geq 5$. For $l \geq 0$ let

$$a_l^\varepsilon = P_l^\varepsilon\{\sigma | m_1^+(\sigma) = 0\}, \qquad b_l^\varepsilon = P_l^\varepsilon\{\sigma | m_1^+(\sigma) = 0, m_2^+(\sigma) = 0\},$$
$$c_l^\varepsilon = P_l^\varepsilon\{\sigma | m_1^+(\sigma) = 0, m_2^+(\sigma) = 0, m_1^-(\sigma) = 0\},$$
$$u_l^\varepsilon = P_l^\varepsilon\{\sigma | m_1^+(\sigma) = 0, m_2^+(\sigma) \geq 1\},$$
$$v_l^\varepsilon = P_l^\varepsilon\{\sigma | m_1^+(\sigma) = 0, m_2^+(\sigma) = 0, m_1^-(\sigma) \geq 2\}.$$

For $l \geq 1$ we then have

$$x_l^\varepsilon = 1 - a_l^\varepsilon - \tfrac{1}{2}a_{l-1}^\varepsilon, \qquad y_l^\varepsilon = u_l^\varepsilon + \tfrac{1}{2}u_{l-1}^\varepsilon, \qquad z_l^\varepsilon = v_l^\varepsilon + \tfrac{1}{2}v_{l-1}^\varepsilon,$$
$$u_l^\varepsilon = a_l^\varepsilon - b_l^\varepsilon, \qquad v_l^\varepsilon = b_l^\varepsilon - c_l^\varepsilon - \tfrac{1}{2}c_{l-1}^{-\varepsilon}.$$

For $l \geq 6$ we have

$$\tfrac{3}{5} \leq a_l^\varepsilon \leq \tfrac{11}{18}, \qquad \tfrac{8}{17} \leq b_l^\varepsilon \leq \tfrac{43}{90}, \qquad c_l^\varepsilon \leq \tfrac{223}{765}. \tag{10.5}$$

These can be checked, for example, as follows. Partitioning W_l^ε according to the type and the length of the cycle containing 1, we get for $l \geq 2$ the recursion relations

$$a_l^\varepsilon = \frac{1}{2l}\left(a_{l-1}^{-\varepsilon} + \sum_{i=0}^{l-2}(a_i^+ + a_i^-)\right),$$
$$b_l^\varepsilon = \frac{1}{2l}\left(b_{l-1}^{-\varepsilon} + b_{l-2}^{-\varepsilon} + \sum_{i=0}^{l-3}(b_i^+ + b_i^-)\right),$$
$$c_l^\varepsilon = \frac{1}{2l}\left(c_{l-2}^{-\varepsilon} + \sum_{i=0}^{l-3}(c_i^+ + c_i^-)\right),$$

and therefore for $l \geq 4$,

$$a_l^\varepsilon = \frac{1}{2l}(a_{l-1}^{-\varepsilon} + a_{l-2}^\varepsilon + (2l-2)a_{l-1}^\varepsilon),$$
$$b_l^\varepsilon = \frac{1}{2l}(b_{l-1}^{-\varepsilon} + b_{l-3}^\varepsilon + (2l-2)b_{l-1}^\varepsilon),$$
$$c_l^\varepsilon = \frac{1}{2l}(c_{l-2}^{-\varepsilon} + c_{l-3}^\varepsilon + (2l-2)c_{l-1}^\varepsilon).$$

These formulas show that once the inequalities in (10.5) hold for three consecutive values of l, they hold for all larger values of l. In this way we find that (10.5) holds for $l \geq 6$, and it follows that (10.4) holds for every $l \geq 8$. Explicit computations show that (10.4) actually holds for every $l \geq 5$.

Finally, we turn to the cases in which $\delta = 2$.

Type ${}^2B_2$. There are q^8 unipotent elements in G^F, hence at least (in fact, exactly) $q^8 - 1$ p-singular elements (where $p = 2$). Since $|G^F| = q^4(q^2 - 1)(q^4 + 1)$, we have $\mu_p(G^F) > (q^2 + 1)q^{-4} \geq q^{-\delta}$.

Type ${}^2G_2$. There are q^{12} unipotent elements in G^F, hence at least $q^{12} - 1$ p-singular elements (where $p = 3$). Since $|G^F| = q^6(q^2 - 1)(q^6 + 1)$, we have $\mu_p(G^F) \geq (q^4 + q^2 + 1)q^{-6} \geq q^{-\delta}$.

Type ${}^2F_4$. In this case it is not enough to count unipotent elements, and we cannot simply look for maximal tori T that have some roots fixed or inverted by F. Indeed, F interchanges short and long roots. What we can look for are subsystems of rank 2 in the root system which are globally F-stable. Such subsystems occur in particular when we have a root α which is fixed by F^2. There is then a 2-dimensional F-stable torus $S < T$ contained in $\mathrm{Ker}(\alpha)$, and $|T^F/S^F| = q^2 - 1$. For such an F-stable maximal torus we therefore have $\varphi_G(T) \geq (1 - q^{-2})/(q^2 - 1) = q^{-2}$. Representatives of the F-conjugacy classes in W and the orders of their stabilizers are given in [Shi2]. With the notation used there, we find that the maximal tori corresponding to the F-conjugacy classes of the elements w_1, w_2, w_3, and w_4 satisfy this condition, and we therefore have $\mu_2(G^F) \geq \frac{1}{16}q^{-2} + \frac{1}{4}q^{-2} + \frac{1}{8}q^{-2} + \frac{1}{8}q^{-2} = \frac{9}{16}q^{-\delta}$.

APPENDIX

We discuss here various issues pertaining to F-stable maximal tori, F-conjugacy classes in the Weyl group, and related matters. We consider a simple algebraic group G defined over an algebraic closure k of a finite field of characteristic $p > 0$, equipped with an endomorphism F such that for some $m \geq 1$, F^m is the Frobenius endomorphism of some definition of G over a finite field (with the notation of Section 5, we can take $m = \delta$).

A.1. A crucial tool in the study of (G, F) is a theorem of Lang, generalized by Steinberg, which asserts in particular that for any F-stable closed connected subgroup H of G, the map $H \to H$, $x \mapsto xF(x)^{-1}$ is surjective [SS, I.2.2]. A typical application of Lang's theorem is that if $H < K$ are F-stable closed subgroups of G and H is connected, then the natural map $K^F/H^F \to (K/H)^F$ is surjective. Indeed, if $F(kH) = kH$, with $k \in K$, then $k^{-1}F(k) \in H$, and by Lang's theorem applied to H there exists therefore $h \in H$ such that $hF(h)^{-1} = k^{-1}F(k)$. Then $kh \in kH$ and $F(kh) = kh$.

A.2. Given a maximal torus $T < G$, we can consider the Weyl group of T in G, $W(T) = N_G(T)/T$. This Weyl group is adequate for many purposes, but it is not canonical. If T' is a second maximal torus, then there exists $g \in G$ such that $T' = {}^gT$, and conjugation by g induces an isomorphism from $W(T)$ to $W(T')$. This isomorphism is not unique in general. For example, when $T' = T$, we get the inner automorphism of $W(T)$ induced by the coset gT. A way to remedy this situation is to consider the set $\mathscr{S}$ of all pairs (T, B) consisting of a maximal torus T and a Borel subgroup $B > T$ of G. Then G acts transitively on $\mathscr{S}$. If $(T, B), (T', B') \in \mathscr{S}$ and $g \in G$ are such that $T' = gTg^{-1}$ and $B' = gBg^{-1}$, then the isomorphisms $i_{T,B}^{T',B'}$: $W(T) \to W(T')$ and $j_{T,B}^{T',B'}$: $T \to T'$ induced by conjugation by g are well-defined. We get in this way direct systems $(W(T)_{(T,B)})_{(T,B) \in \mathscr{S}}$ and $(T_{(T,B)})_{(T,B) \in \mathscr{S}}$ in which all maps are isomorphisms. We define the Weyl group $\mathbf{W}$ and the maximal torus $\mathbf{T}$ of G to be the respective direct limits of these constant direct systems. Moreover, every $\mathbf{w} \in \mathbf{W}$ induces an automorphism $\Theta_{\mathbf{w}}$ of $\mathbf{T}$. Let $i_{T,B}$: $W(T) \to \mathbf{W}$ and $j_{T,B}$: $T \to \mathbf{T}$ be the isomorphisms determined by (T, B). (This approach is similar to that used in [DL, pp. 105–106].)

A.3. There are Frobenius maps on $\mathbf{T}$ and $\mathbf{W}$, defined as follows. Let $(T, B) \in \mathscr{S}$. Then $\mathbf{F}$: $\mathbf{T} \to \mathbf{T}$ is defined by $\mathbf{F} = j_{F(T),F(B)} \circ F \circ (j_{T,B})^{-1}$, and if $\mathbf{w} = i_{T,B}(nT)$ with $n \in N_G(T)$, then $\mathbf{F}(\mathbf{w}) = i_{F(T),F(B)}(F(nT))$. For $\mathbf{w} \in \mathbf{W}$ we have then $\Theta_{\mathbf{F}(\mathbf{w})} \circ \mathbf{F} = \mathbf{F} \circ \Theta_{\mathbf{w}}$. If both T and B are F-stable, then the $\mathbf{F}$-actions on $\mathbf{T}$ and $\mathbf{W}$ can be read directly from the F-actions on T and $W(T)$ (in [Car5, pp. 84 ff.], an F-stable pair (T_0, B_0) is chosen once and for all).

A.4. The character group of an algebraic torus T is the abelian group $X(T) = \mathrm{Hom}(T, \mathbf{G}_m)$. If φ: $S \to T$ is a homomorphism of algebraic tori, we let φ^* denote the homomorphism of abelian groups $X(T) \to X(S)$ defined by $\lambda \mapsto \lambda \circ \varphi$.

Consider the endomorphism $\mathbf{F}^*$ of $X(\mathbf{T})$ induced by $\mathbf{F}$: $\mathbf{T} \to \mathbf{T}$. If F corresponds to a split structure of G over a finite field of order q, then $\mathbf{F}^*$ is just multiplication by q. In general there exists $n \geq 1$ such that F^n corresponds to a split structure, and $\mathbf{F}^{*n}$ is then multiplication by some power q_n of p. We set $q = (q_n)^{1/n}$ (this definition of q agrees with that given in Section 5). Then $\mathbf{F}^* = q\mathbf{f}$, where $\mathbf{f}$ is an automorphism of finite order arising from the graph automorphism induced by F (when $\delta = 2$, this actually holds only in $X(\mathbf{T}) \otimes_{\mathbb{Z}} \mathbb{R}$). In particular $\mathbf{f}$ belongs to the group $\mathbf{A}$ of all automorphisms of the root system of G (again, when $\delta = 2$ this definition needs some stretching; root lengths must be ignored).

A.5. The action of $\mathbf{W}$ on $X(\mathbf{T})$ defined by $\mathbf{w}\boldsymbol{\lambda} = \Theta_{\mathbf{w}^{-1}}^*(\boldsymbol{\lambda})$ is faithful and allows us to think of $\mathbf{W}$ as a normal subgroup of $\mathbf{A}$. Since $\Theta_{\mathbf{F}(\mathbf{w})} \circ \mathbf{F} = \mathbf{F} \circ \Theta_{\mathbf{w}}$, we have $\mathbf{F}(\mathbf{w}) = \mathbf{f}^{-1}\mathbf{w}\mathbf{f}$ in $\mathbf{A}$. The map $\mathbf{W} \to \mathbf{W}\mathbf{f}^{-1}$, $\mathbf{w} \mapsto \mathbf{w}\mathbf{f}^{-1}$ induces therefore a bijection between the **F**-conjugacy classes in **W** and the **W**-conjugacy classes in $\mathbf{Wf}^{-1}$. For twisted groups of type A_l or E_6, we have $\mathbf{f} = -\mathbf{w}_0$, where $\mathbf{w}_0$ is the longest element in the Weyl group. For groups of type D_l, $l \geq 5$, **A** is a Weyl group of type B_l, and (G, F) can actually be realized as an F-stable subgroup in a group of type B_l. For D_4, **A** is a Weyl group of type F_4, and (G, F) itself can be realized as an F-stable subgroup of a group of type F_4 (we need this only for ${}^3D_4$; for other groups of type D_4 it is more convenient to use a subgroup of **A** which is a Weyl group of type B_4).

A.6. Consider now an F-stable maximal torus T of G. Let $B > T$ be a Borel subgroup. Then $F(B)$ is also a Borel subgroup containing T, and there exists $v \in N_G(T)$ such that $F(B) = vBv^{-1}$. Then vT depends only on (T, B), and we get therefore an element $\mathbf{w} = i_{T,B}(vT) \in \mathbf{W}$ which depends only on T and B. This element **w** is also characterized by the condition $\Theta_{\mathbf{w}} = j_{T,B} \circ (j_{T,F(B)})^{-1}$. It depends on the choice of B, but its **F**-conjugacy class in **W** is independent of B. We get in this way a map from the set $\mathscr{T}^F$ of all F-stable maximal tori in G to the set $W/\sim_F$ of all **F**-conjugacy classes in the Weyl group. This map is obviously constant on G^F-conjugacy classes in $\mathscr{T}^F$, hence induces a map $\mathscr{T}^F/G^F \to W/\sim_F$, which can be shown to be a bijection by repeated use of Lang's theorem (compare [Car5, p. 84]).

A.7. Given an F-stable maximal torus T in G and a Borel subgroup $B > T$, the maps $F^*\colon X(T) \to X(T)$ and $\mathbf{F}^*\colon X(\mathbf{T}) \to X(\mathbf{T})$ satisfy $(j^*_{T,B})^{-1} \circ F^* \circ j^*_{T,B} = \mathbf{F}^* \circ \Theta^*_{\mathbf{w}}$, where **w** is the element of **W** such that $\Theta_{\mathbf{w}} = j_{T,B} \circ (j_{T,F(B)})^{-1}$. Indeed, since $\mathbf{F} = j_{F(T),F(B)} \circ F \circ (j_{T,B})^{-1}$, we have $F^* = j^*_{T,B} \circ \mathbf{F}^* \circ (j^*_{F(T),F(B)})^{-1}$, and since $\Theta_{\mathbf{w}} = j_{T,B} \circ (j_{T,F(B)})^{-1}$, we have $\Theta^*_{\mathbf{w}} = (j^*_{T,F(B)})^{-1} \circ j^*_{T,B}$. Thus

$$(j^*_{T,B})^{-1} \circ F^* \circ j^*_{T,B} = (j^*_{T,B})^{-1} \circ j^*_{T,B} \circ \mathbf{F}^* \circ (j^*_{F(T),F(B)})^{-1} \circ j^*_{T,B} = \mathbf{F}^* \circ \Theta^*_{\mathbf{w}}.$$

This means that F^* acts on $X(T)$ in the same way in which $\mathbf{F}^* \circ \mathbf{w}^{-1}$ acts on $X(\mathbf{T})$.

A.8. This shows also that qF^{*-1} has finite order. Let χ_T be the characteristic polynomial of qF^{*-1}, or equivalently the characteristic polynomial of $\mathbf{wf}^{-1} \in \mathbf{A}$. We show that $|T^F| = \chi_T(q)$. We note first that the definition of the character group makes sense for every algebraic group. This notion is extremely powerful for *diagonalizable* algebraic groups, that is, algebraic groups which can be embedded as closed subgroups of algebraic tori. Indeed, X induces a contravariant equivalence of categories between diagonalizable algebraic groups and finitely generated abelian groups without p-torsion [Bor, p. 113]. Since T^F is a closed subgroup of T,

$X(T^F)$ is a quotient of $X(T)$. It is clear that for every $\lambda \in (F^* - 1)X(T)$ and every $t \in T^F$ we have $\lambda(t) = 1$. Conversely, it is easily checked that if $t \in T$ is such that $\lambda(t) = 1$ for every $\lambda \in (F^* - 1)X(T)$, then $t \in T^F$. It follows that $X(T^F)$ and $X(T)/(F^* - 1)X(T)$ are isomorphic up to p-torsion. However, $X(T)/(F^* - 1)X(T)$ has no p-torsion (since $q^{-1}F^*$ has finite order, F^* is nilpotent mod p, hence $F^* - 1$ is invertible mod p). Thus $X(T^F) \cong X(T)/(F^* - 1)X(T)$ and therefore

$$\begin{aligned}|T^F| &= |X(T^F)| = |X(T)/(F^* - 1)X(T)| = |\det(F^* - 1)|\\ &= |\det(F^* - 1)|\,|\det(qF^{*-1})|\\ &= |\det(q - qF^{*-1})| = \det(q - qF^{*-1}) = \chi_T(q).\end{aligned}$$

(Observe that $q - qF^{*-1}$ is orientation preserving since $q > 1$ and qF^{*-1} has finite order.)

A.9. Suppose now that $\alpha \in X(T)$ is a root of G. Let U_α be the root subgroup corresponding to α. Then there exists a root β such that $F(U_\alpha) = U_\beta$. This root is characterized by the property that $\beta \circ F = q_\alpha \alpha$ for some positive integer q_α (q_α is always a power of the characteristic, $q_\alpha = q$ if $\delta = 1$ and $q_\alpha q_\beta = q^2$ if $\delta = 2$). We say that F fixes α if $\alpha \circ F$ is a positive multiple of α, or equivalently if $F(U_\alpha) = U_\alpha$. If $B > T$ and $\mathbf{w} \in \mathbf{W}$ are as above and $\alpha = j^*_{T,B}(\boldsymbol{\alpha})$, then α is fixed by F if and only if $(\mathbf{wf}^{-1})(\boldsymbol{\alpha}) = \boldsymbol{\alpha}$. Thus α is fixed by F if and only the element of $\mathbf{Wf}^{-1}$ associated to (T, B) fixes $\boldsymbol{\alpha}$ in the usual sense. Similarly, we say that F inverts α if $\alpha \circ F$ is a negative multiple of α, or equivalently if $F(U_\alpha) = U_{-\alpha}$. This is also equivalent to the requirement that $(\mathbf{wf}^{-1})(\boldsymbol{\alpha}) = -\boldsymbol{\alpha}$.

ACKNOWLEDGMENT

We are grateful to the thorough and helpful referee.

REFERENCES

[Bor] A. Borel, "Linear Algebraic Groups," Springer-Verlag, Berlin/Heidelberg/New York, 1991.

[Bou] N. Bourbaki, "Groupes et algèbres de Lie," Chaps. 4–6, Hermann, Paris, 1968.

[Cam] P. J. Cameron, Finite permutation groups and finite simple groups, *Bull. London Math. Soc.* **13** (1981), 1–22.

[Car1] R. W. Carter, "Simple Groups of Lie Type," Wiley, London/New York, 1972.

[Car2] R. W. Carter, Conjugacy classes in the Weyl group, *in* "Seminar on Algebraic Groups and Related Finite Groups," (A. Borel *et al.*), G1–G22, Lecture Notes in Mathematics, Vol. 131, Springer-Verlag, Berlin/New York, 1970.

[Car3] R. W. Carter, Conjugacy classes in the Weyl group, *Compositio Math.* **25** (1972), 1–59.

[Car4] R. W. Carter, Centralizers of semisimple elements in the finite classical groups. *Proc. London Math. Soc.* **42** (1981), 1–41.

[Car5] R. W. Carter, Simple Groups of Lie Type: Conjugacy Classes and Characters," Wiley, New York, 1985.

[CCNPW] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson, "Atlas of Finite Groups," Clarendon Press, Oxford, 1985.

[Co] B. N. Cooperstein, Minimal degree for a permutation representation of a classical group, *Israel J. Math.* **30** (1978), 213–235.

[De] D. I. Deriziotis, The centralizers of semisimple elements of the Chevalley groups E_7 and E_8, *Tokyo J. Math.* **6** (1983), 191–216.

[DF] D. I. Deriziotis and A. P. Fakiolas, The maximal tori of the finite Chevalley groups of type E_6, E_7 and E_8, *Comm. Algebra* **19** (1991), 889–903.

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

[DLi] D. I. Deriziotis and M. W. Liebeck, Centralizers of semisimple elements in finite twisted groups of Lie type, *J. London Math. Soc.* **31** (1985), 48–54.

[ET] P. Erdös and P. Turán, On some problems of a statistical group-theory, II, *Acta Math. Acad. Sci. Hungar.* **18** (1967), 151–163.

[FKL] L. Finkelstein, D. Kleitman, and T. Leighton, Applying the classification theorem for finite simple groups to minimize pin count in uniform permutation architectures, *in* "Proceedings, 3rd International Workshop on Parallel Computation and VLSI Theory," pp. 247–256, Lecture Notes in Computer Science, Vol. 319, Springer-Verlag, Berlin/New York.

[Ga] A. Gambini, Ph.D. thesis, Universität Freiburg, 1992.

[Ka] W. M. Kantor, Polynomial-time algorithms for finding elements of prime order and Sylow subgroups, *J. Algorithms* **6** (1985), 478–514.

[Kaw] N. Kawanaka, Generalized Gelfand–Graev representations and Ennola duality, *in* "Algebraic Groups and Related Topics," pp. 175–206, Adv. Stud. Pure Math., Vol. 6, North-Holland, Amsterdam, 1985.

[LS] V. Landazuri and G. M. Seitz, On the minimal degrees of projective representations of the finite Chevalley groups, *J. Algebra* **32** (1974), 418–443.

[Li] M. W. Liebeck, On the orders of maximal subgroups of the finite classical groups, *Proc. London Math. Soc.* **50** (1985), 426–446.

[Sh1] K. Shinoda, The conjugacy classes of Chevalley groups of type F_4 over finite fields of characteristic 2, *J. Fac. Sci. Univ. Tokyo* **21** (1974), 133–159.

[Sh2] K. Shinoda, The conjugacy classes of the finite Ree groups of type F_4, *J. Fac. Sci. Univ. Tokyo* **22** (1973), 1–15.

[Sho] T. Shoji, The conjugacy classes of Chevalley groups of type F_4 over finite fields of characteristic $p \neq 2$, *J. Fac. Sci. Univ. Tokyo* **21** (1974), 1–17.

[SS] T. A. Springer and R. Steinberg, Conjugacy classes, *in* "Seminar on Algebraic Groups and Related Finite Groups," (A. Borel *et al.*), E1–E100, Lecture Notes in Mathematics, Vol. 131, Springer-Verlag, Berlin/New York, 1970.

[St] R. Steinberg, "Lectures on Chevalley Groups," Lecture Notes, Yale University, 1967.

[Su] M. Suzuki, On a class of doubly transitive groups, *Ann. of Math.* **75** (1962), 105–145.

[Wa] H. N. Ward, On Ree's series of simple groups. *Trans. Amer. Math. Soc.* **121** (1966), 62–89.