JOURNAL OF ALGEBRA **194,** 79–112 (1997)
ARTICLE NO. JA966980

Short Presentations for Finite Groups

L. Babai*[†]

University of Chicago, Computer Science Department,
1100 E. 58th St., Chicago, Illinois 60637 and
Mathematical Institute of the Hungarian Academy of Science, Budapest, Hungary

A. J. Goodman*

University of Missouri-Rolla, Mathematics and Statistics Department,
Rolla, Missouri 65409

W. M. Kantor*

University of Oregon, Mathematics Department, Eugene, Oregon 97403

E. M. Luks*

University of Oregon, Computer and Information Science Department,
Eugene, Oregon 97403

and

P. P. Pálfy[‡]

Mathematical Institute of the Hungarian Academy of Science,
Budapest, P. O. Box 127, H-1364 Hungary

Communicated by Robert Steinberg

Received February 8, 1996

We conjecture that every finite group G has a short presentation (in terms of generators and relations) in the sense that the total *length* of the relations is $(\log|G|)^{O(1)}$.

* Research supported in part by NSF grants.

[†] E-mail: laci@cs.uchicago.edu, goodman@umr.edu, kantor@math.uoregon.edu, luks@cs.uoregon.edu, ppp@math-inst.hu.

[‡] Research supported in part by Hungarian Research Fund Grant OTKA-1903.

0021-8693/97 $25.00

We show that it suffices to prove this conjecture for simple groups.

Motivated by applications in computational complexity theory, we conjecture that for finite simple groups, such a short presentation is computable in polynomial time from the standard name of G, assuming in the case of Lie type simple groups over $GF(p^m)$ that an irreducible polynomial f of degree m over $GF(p)$ and a primitive root of $GF(p^m)$ are given.

We verify this (stronger) conjecture for all finite simple groups except for the three families of rank 1 twisted groups: we do not handle the unitary groups $PSU(3, q) = {}^2A_2(q)$, the Suzuki groups $Sz(q) = {}^2B_2(q)$, and the Ree groups $R(q) = {}^2G_2(q)$. In particular, *all finite groups G without composition factors of these types have presentations of length $O((\log|G|)^3)$.*

For groups of Lie type (normal or twisted) of rank ≥ 2, we use a reduced version of the Curtis–Steinberg–Tits presentation.

CONTENTS

1. INTRODUCTION

A *presentation* of a group is a description by generators and relations which defines the group; we allow relations written as equations using powers of the generators. For general background and a number of specific presentations, we refer to Coxeter and Moser [CM72].

We define the *length* of a presentation to be the total number of characters required to write all the relations. We write exponents in binary and include the number of digits in this count. We shall only be interested in the asymptotic order of magnitude of presentation lengths, so it makes no difference whether or not we count parentheses, equal signs, and minus signs; for definiteness, we shall not count them. We use log to denote base 2 logarithms. Each generator will count as a single symbol.

EXAMPLE 1.1. The presentation $\langle x : x^m = 1\rangle$ of the *cyclic group* of order m has length $\log m + O(1)$ (since it takes at most $1 + \log m$ digits to write the exponent m in binary). The presentation of $PSL(2, p)$ given in Remark 7.2 has length $O(\log p)$.

Remark 1.2. A natural alternative to counting each generator as a single symbol is to assume that the generators are indexed symbols $x_1, \ldots, x_k$ and to count the digits in these subscripts as part of the length. However, this would only increase the length of a presentation by a factor of $\log k$, an insignificant factor from our point of view since $\log k$ is $O(\log\log|G|)$ in the short presentations we are interested in.

Remark 1.3. We note that insisting that no exponents be used would not force an increase in the length by more than a factor of 4 (asymptotically), since we can introduce extra generators for the repeated squares of expressions to be raised to large powers. To illustrate this, consider again the cyclic group of order m. Written without exponents, the above presentation $\langle x : x^m = 1\rangle$ turns into $\langle x : xxx \cdots x = 1\rangle$ which has length $m + O(1)$. But another presentation for the same group is the following: $\langle x_0, x_1, \ldots, x_k : x_{i+1} = x_i^2 \ (0 \le i < k), x_0^{e_0}x_1^{e_1} \cdots x_k^{e_k} = 1\rangle$, where each e_i is 0 or 1 and $m = \sum e_i 2^i$ expresses m in base 2 (so $k = \lfloor \log m \rfloor$). This presentation, written without exponents, has length at most $3k + (k + 1) + O(1) \le 4\log m + O(1)$. (The factor of 4 would become a factor of 5 if we counted the number of generators as well as the total length of all the relations.) The same method can be used to remove the exponents from any presentation.

In this paper we shall give evidence to support the following conjecture.

Conjecture 1 (Short Presentation Conjecture). There exists a constant C such that every finite group G has a presentation of length $O(\log^C|G|)$.

We conjecture that $C = 3$ suffices.

We show that it suffices to prove the conjecture for simple groups.

THEOREM 1.4. *If Conjecture* 1 *holds for all finite simple groups with some* $C \ge 2$, *then every finite group G has a presentation of length* $O(\log^{C+1}|G|)$.

We postpone the proof of Theorem 1.4 to the last section and focus on the case of simple groups in Sections 2 through 7. We shall verify the conjecture for most classes of finite simple groups. The groups we miss are the rank 1 twisted groups of Lie type, i.e., the following three families: the unitary groups $PSU(3, q) = {}^2A_2(q)$, the Suzuki groups $Sz(q) = {}^2B_2(q)$, and the Ree groups $R(q) = {}^2G_2(q)$.

THEOREM 1.5. *The short presentation conjecture holds, with* $C = 2$, *for all finite simple groups, with the possible exception of the rank* 1 *twisted groups of Lie type*.

Part of the motivation for this work comes from the complexity theory of algorithmic problems in finite matrix groups [BS84, Ba91, Ba92, Ba97]. What those applications actually require is more than the mere *existence* of short presentations for finite simple groups. The presentations must be "efficiently verifiable" in a well-defined sense. The reader interested only in the group-theoretic aspects of these questions may skip the rest of this section.

Every finite simple group G has a *standard name* of length $O(\log|G|)$ (see below).

DEFINITION 1.1. Assume that "standard names" are associated with the members of a class $\mathcal{F}$ of finite groups. We say that $\mathcal{F}$ has *efficiently verifiable* presentations if there exists a constant C and a deterministic multitape Turing machine M accepting triples (G, P, W) of strings such that

(i) whenever G is the standard name of some member of $\mathcal{F}$, there exist P and W such that P is a presentation of G and M accepts (W is the "witness" of correctness of the presentation);

(ii) if G is a standard name of some member of $\mathcal{F}$ and P is not a presentation of G, then M rejects (no "false witness" exists);

(iii) if G is the standard name of some member of $\mathcal{F}$, then M halts in $O(\log^C|G|)$ steps.

Informally this means that short presentations with short proofs of correctness exist. (Note that efficiently verifiable presentations are automatically short: M has at least to read P before accepting.) The following version of Conjecture 1 suffices for the applications we have in mind.

Conjecture 2 (Efficiently Verifiable Presentation Conjecture). The class of all finite simple groups admits efficiently verifiable presentations.

It would seem desirable to make the conjecture still more effective; one might want to see presentations that are *uniform* in the sense that they are *computable* (on a multitape Turing machine) in time $O(\log^C|G|)$, given the standard name of G. One obstacle to this might be the following. In the case of groups of Lie type, the standard name of G includes a prime power

$q = p^m$. An explicit representation of the field $GF(q)$, i.e., an irreducible polynomial of degree m over $GF(p)$, is likely to be needed for the computation. No deterministic algorithm, running in $(\log q)^{O(1)}$ time, is known at present to produce such a polynomial except under the extended Riemann hypothesis (Adleman and Lenstra [AL86]). (Randomized algorithms do exist for this purpose, and the verification that a given polynomial is irreducible can be performed deterministically in $(\log q)^{O(1)}$ time [Ber70, CZ81, Rab80].) We shall not be concerned with this difficulty and get around it by the following convention.

CONVENTION $(*)$. In case of a group G of Lie type over the field $GF(p^m)$, we require that the standard name of G include an irreducible polynomial of degree m over $GF(p)$.

This convention seems quite natural and amounts to saying that the underlying field itself rather than its order has to be specified as part of the name of the group. It is less natural, but it seems necessary in some cases, that we have, in addition, a primitive root (generator of the multiplicative group) of $GF(q)$ at hand. Without this, we presently are not able to turn a presentation of $SL(n, q)$ into one of $PSL(n, q)$. (However, this step and its analogues for the other Lie type groups are the only places where we will need the primitive root.) Unfortunately, no deterministic algorithm is known to find a primitive root of $GF(q)$ in time $(\log q)^{O(1)}$, even for the case $m = 1$ ($q = p$ a prime). Given these additional tools, however, we believe that presentations can be computed efficiently from the names of the finite simple groups. (For twisted groups the relevant field may be $GF(q^2)$ or $GF(q^3)$ in terms of the q given in the standard name, and we assume that an irreducible polynomial and primitive root are given for this larger field.)

Conjecture 3 (Uniform Short Presentation Conjecture). There exists a constant C and an algorithm which computes a presentation of each finite simple group G in time $O(\log^C|G|)$, given the following input:

(a) the standard name of G;

(b) for groups of Lie type over $GF(q)$, where $q = p^m$ with p prime,

(b1) an irreducible polynomial f of degree m over $GF(p)$ and

(b2) a primitive root of $GF(q)$.

More precisely, we conjecture that $C = 2$ suffices.

(Given f as in (b1), we represent $GF(q)$ as $GF(p)[x]/(f)$; and (b2) is understood to conform to this representation, i.e., the primitive root is given by a polynomial in $GF(p)[x]$ whose class mod (f) is the primitive root in $GF(p)[x]/(f)$.)

Each of the three conjectures is stronger than its predecessor: Conjecture 3 implies Conjecture 2 even without assuming $(*)$, because the additional input (b1) and (b2) may be contained in the witness W (the machine in Definition 1.1 may use its inputs G and W together to compute a presentation which it then compares with the given presentation P). Thus it suffices to consider Conjecture 3. For clarity, the proofs will be worded to stress the validity of Conjecture 1; the presentations obtained will be so explicit that the algorithmic requirements in the other conjectures will evidently be met.

THEOREM 1.6. *The uniform short presentation conjecture holds, with* $C = 2$, *for all finite simple groups, with the possible exception of the rank* 1 *twisted groups of Lie type*.

The *complexity theoretic application* are of a qualitative nature. First stated in [BS84], Conjecture 2 is known to imply that certain problems for matrix groups G over finite fields belong to the complexity class NP (cf. [GJ79] for information regarding this complexity class), where a group is given by a list of generating matrices. These problems include *nonmembership* of a given matrix in G, verification of the *order* of G [BS84], and *isomorphism* of two such groups G, H (cf. [Ba92, Propositions 4.9 and 4.10] for this result of Luks). (The proof that the *membership* problem belongs to NP is elementary; it is an immediate consequence of the reachability lemma [BS84], stated in Section 8 of this paper.) For a list of further consequences of Conjecture 2, see [Ba92, Corollary 12.1].

Conjecture 3 is expected to be useful in efficient algorithms for matrix groups (cf. [Ba97]).

The algorithmic parts of the present paper deal only with simple groups. We do not address algorithmic questions for the nonsimple groups in Conjecture 1, nor the question of finding the standard name of a simple group given in some other way.

2. GROUPS OTHER THAN LIE TYPE

As customary, for two sequences of positive real numbers $\{a_n\}$ and $\{b_n\}$, we use the expression $a_n = \Theta(b_n)$ to denote that $a_n = O(b_n)$ and $b_n = O(a_n)$.

First we note that the *sporadic groups* do not matter for us. Even if we use their full Cayley tables for presentations, this will not influence the validity of the conjectures.

Example 1.1 takes care of *cyclic groups* G: their natural presentations have length $\Theta(\log|G|)$.

A number of suitable presentations for the alternating groups $G = A_n$ are given in [CM72, Sect. 6.3]. All of them have length $\Theta(n^2) = \Theta((\log|G|/\log\log|G|)^2)$. We quote a presentation contained in [Car23].

Generators: $x_i \quad (i = 1, \ldots, n-2)$,

Relations: $x_1^3 = \cdots = x_{n-2}^3 = (x_i x_j)^2 = 1 \quad (1 \le i < j \le n-2)$.

3. CHEVALLEY GROUPS

The remaining finite simple groups are of *Lie type*. They fall into two categories: the *normal* and the *twisted* types. We shall refer to the former as *Chevalley groups*. Presentations for the groups of Lie type were first given by Steinberg [Ste 62, Ste 81] (cf. Sections 4 and 6 below).

In Sections 3 to 5 we consider the case of Chevalley groups, postponing the discussion of the twisted types until Section 6. Each Chevalley group G is associated with a finite field of order q and a parameter N (the number of "positive roots"). Steinberg uses $2Nq$ generators. The Curtis−Steinberg−Tits presentation [Cur65] reduces this number somewhat, but its length is still proportional to q, not $\log q$ (cf. Section 4). This is not satisfactory from our point of view since $\log|G|$ is approximately $2N \log q$. We shall see, however, that it is a relatively simple matter to reduce the presentation in terms of these generators, for groups of rank $n \ge 2$. The rank 1 case, i.e., the groups $PSL(2, q)$, was settled in [Tod36].

We will need some standard terminology involving Chevalley groups. Our main references are Steinberg's lecture notes [Ste67] and Carter's book [Car72].

Each Chevalley group G has a central extension $\hat{G}$ called a *universal Chevalley group*. If Z denotes the center of $\hat{G}$, then $G = \hat{G}/Z$; moreover $\hat{G}' = \hat{G}$ (where $'$ denotes the commutator subgroup). First we describe presentations for the universal Chevalley groups, and then presentations of the simple Chevalley groups are obtained by killing Z.

Let L stand for one of the letters A, B, C, D, E, F, G. Let L_n denote the corresponding complex simple Lie algebra of rank n. (For type A, $n \ge 1$; for types B and C, $n \ge 2$ (here $B_2 = C_2$); for type D, $n \ge 4$; for type E, $6 \le n \le 8$; for type F, $n = 4$; and for type G, $n = 2$.) Types A, B, C, and D are called *classical*, the others *exceptional*.

Let $\mathsf{L}_n(q)$ denote the universal Chevalley group over the field of order $q = p^m$ (p prime), corresponding to L_n. The number n is the *rank* of the group.

Remark 3.1. In this classification, the classical simple groups appear as follows. The left column indicates universal Chevalley groups in Lie notation; the middle column their central quotient, a classical simple group; the right column their classical name.

$A_n(q)$	$PSL(n+1,q)$	linear
$B_n(q)$	$P\Omega(2n+1,q)$	orthogonal
$C_n(q)$	$PSp(2n,q)$	symplectic
$D_n(q)$	$P\Omega^+(2n,q)$	orthogonal

There is another class of simple orthogonal groups of even dimension, $P\Omega^-(2n,q)$. They, along with the unitary groups, correspond to twisted types:

${}^2A_n(q)$	$PSU(n+1,q)$	unitary
${}^2D_n(q)$	$P\Omega^-(2n,q)$	orthogonal

where, however, n is not the rank of the twisted group.

Remark 3.2. In almost all cases the universal Chevalley group $\hat{G}$ is the *universal central extension* of the simple group G, but there are a few cases where $\hat{G}$ has nontrivial Schur multiplier [Ste81]. For convenience, we treat these finitely many exceptions as sporadic groups (they cannot affect our asymptotic results) and exclude them in the following sections, because the Curtis–Steinberg–Tits presentation is known [Cur65] to describe a central extension of $\hat{G}$, which therefore must be $\hat{G}$ if $\hat{G}$ has trivial Schur multiplier. If the rank is 2, then the Curtis–Steinberg–Tits presentation is the same as the Steinberg presentation for $\hat{G}$, so we only need to exclude the cases of rank ≥ 3 where $\hat{G}$ has nontrivial Schur multiplier, namely $A_3(2)$, $B_3(2) \cong C_3(2)$, $B_3(3)$, $D_4(2)$, and $F_4(2)$.

The same remark applies to the central extensions of twisted groups considered in Section 6, where we exclude ${}^2A_5(2)$ and ${}^2E_6(2)$.

4. PRESENTATIONS FOR THE UNIVERSAL CHEVALLEY GROUPS OF RANK ≥ 2

In order to simplify the notation, in Sections 4 and 5 we will *change notation* and now let G stand for the universal Chevalley group $L_n(q)$. Let Φ denote the set of roots of the simple Lie algebra L_n. Let Φ^+ be a positive system of roots; $N = |\Phi^+|$ and $|\Phi| = 2N$. Let $d_1, \dots, d_n$ denote the degrees of the algebraically independent homogeneous polynomials that generate the ideal of invariant polynomials of the Weyl group of L_n. Then [Ste67, Theorem 25, p. 130]

$$N = \sum_{i=1}^{n} (d_i - 1) \tag{1}$$

and

$$|G| = q^N \prod (q^{d_i} - 1). \tag{2}$$

Consequently,

$$q^{2N} < |G| < q^{2N+n} \leq q^{3N}. \tag{3}$$

These inequalities can also be checked directly from explicit order formulas (e.g. [Gor80, p. 490]).

4.1. *The Steinberg and Curtis–Steinberg–Tits Presentations*

Steinberg introduces the set $\{x_\alpha(t) : \alpha \in \Phi, t \in GF(q)\}$ of $2Nq$ generators for G and shows the following set of relations to be sufficient to define G in the case $n \geq 2$ [Ste67, Theorem 9, p. 72].

(A) $\qquad x_\alpha(t+u) = x_\alpha(t)x_\alpha(u) \qquad (t, u \in GF(q),\ \alpha \in \Phi);$

(B) $\quad \left[x_\alpha(t), x_\beta(u)\right] = \prod_{i,j>0} x_{i\alpha+j\beta}\left(C_{i,j,\alpha,\beta} t^i u^j\right)$

for $\alpha, \beta \in \Phi$, $\alpha \neq \pm\beta$, $t, u \in GF(q)$.

Here $[g, h]$ denotes the commutator $g^{-1}h^{-1}gh$; the product is taken over those values $i, j > 0$ for which $i\alpha + j\beta$ is a root; therefore the product never has more than four terms and the relevant values of i, j are not greater than 3; the coefficients $C_{i,j,\alpha,\beta}$ thus form a finite list (independent of n and q) for given α, β. These coefficients are integers of absolute value ≤ 3 (most often ± 1) and are explicitly computed in [Ste67, Sect. 10]. In all cases except $G_2(q)$, the terms in the product in (B) commute and thus their order does not matter. The case $G_2(q)$ is described in detail in [Ste67, Sect. 10].

A subset of the above generators and relations will suffice for a presentation of G (see Remark 3.2), namely the Curtis–Steinberg–Tits presentation [Cur65], which can be described as follows.

For $1 \leq i < j \leq n$, let Φ_{ij} denote the rank 2 subsystem of Φ spanned by the ith and jth fundamental roots (we are assuming a fixed choice of a positive system of roots Φ^+, or equivalently a fixed base consisting of n fundamental roots). Let $\Psi = \bigcup_{i<j} \Phi_{ij}$, let $\Upsilon_{ij} = \{(\alpha, \beta) \in \Phi_{ij} \times \Phi_{ij}$: $\alpha \neq \pm\beta\}$, and let $\Upsilon = \bigcup_{i<j} \Upsilon_{ij}$. (Note that Ψ is a subset of Φ and Υ is a set of pairs of roots from Ψ.) Then use only those of Steinberg's generators $x_\alpha(t)$ for which $\alpha \in \Psi$; and use only those of the above relations of type (A) for which $\alpha \in \Psi$ and those of type (B) for which the pair (α, β) is in Υ.

In other words, the Curtis–Steinberg–Tits presentation involves only those relations that come from certain rank 2 subgroups of G (namely the rank 2 subgroups $G_{ij} = \langle x_\alpha(t) : \alpha \in \Phi_{ij},\ t \in GF(q)\rangle$ corresponding to pairs of fundamental roots, i.e., edges or nonedges in the Dynkin diagram).

EXAMPLE 4.1 [Ste67, p. 72]. For $n \geq 3$, the group $SL(n, q) = A_{n-1}(q)$ is defined in terms of the generators $x_{ij}(t)$ $(1 \leq i, j \leq n,\ i \neq j,\ t \in GF(q))$ by the following relations:

$$x_{ij}(t+u) = x_{ij}(t)x_{ij}(u),$$

$$\left[x_{ij}(t), x_{jk}(u)\right] = x_{ik}(tu) \qquad \text{if } i, j, k \text{ are different},$$

$$\left[x_{ij}(t), x_{kl}(u)\right] = 1 \qquad \text{if } j \neq k,\ i \neq l.$$

This presentation encodes the relations among the usual matrices $I + tE_{ij}$, where E_{ij} is the matrix whose (i, j) entry is 1 with all other entries 0.

The Curtis–Steinberg–Tits presentation of $SL(n, q)$ uses only the generators $x_{ij}(t)$ with $|i - j| \leq 2$. Of the three types of relations given above, use the first type for these same pairs i, j (i.e., whenever $|i - j| \leq 2$); use the second type whenever i, j, k are three consecutive integers (in some order); and use the third type whenever i, j, k, l all differ by at most one from some fixed value (these, like the second type, come from rank 2 subsystems $\Phi_{i,i+1}$ of type A_2) as well as whenever $|i - j| = 1$ and $|k - l| = 1$ (these come from subsystems Φ_{ik} of type $A_1 \times A_1$). This uses $O(nq)$ generators and $O(n^2q^2)$ relations, whereas the original Steinberg presentation used $O(n^2q)$ generators and $O(n^4q^2)$ relations (cf. Section 4.3).

4.2. *The Reduced Set*

We further reduce the Curtis–Steinberg–Tits presentation as follows. Let $B = \{b_1, \dots, b_m\}$ be a basis of $GF(q)$ over the prime field $GF(p)$. Use $|\Psi|m$ generators $y_\alpha(b_\nu)$ $(\alpha \in \Psi,\ 1 \leq \nu \leq m)$, and define

(A0)
$$y_\alpha(t) := \prod_\nu y_\alpha(b_\nu)^{k_\nu} \qquad \text{whenever } t = \sum_\nu k_\nu b_\nu \ (0 \leq k_\nu < p).$$

We define a group G_0, generated by the symbols $y_\alpha(b_\nu)$ $(\alpha \in \Psi,\ 1 \leq \nu \leq m)$, subject to the following set of relations.

(A1) $$y_\alpha(b_\nu)^p = 1 \qquad (\alpha \in \Psi, 1 \leq \nu \leq m);$$

(A2) $$\left[y_\alpha(b_\nu), y_\alpha(b_\mu)\right] = 1 \qquad (\alpha \in \Psi, 1 \leq \nu < \mu \leq m);$$

(B0) $$\left[y_\alpha(b_\nu), y_\beta(b_\mu)\right] = \prod_{i,j>0} y_{i\alpha+j\beta}\left(C_{i,j,\alpha,\beta} b_\nu^i b_\mu^j\right)$$

$$((\alpha, \beta) \in \Upsilon, 1 \leq \nu, \mu \leq m).$$

(Note: The right-hand side of (B0) must be expanded, using the definition (A0), into an expression involving the generators $y_\alpha(b_\nu)$ $(\alpha \in \Psi,\ 1 \le \nu \le m)$.)

THEOREM 4.2. $G_0 \cong G$: *each universal Chevalley group of rank* $n \ge 2$ *is defined by a presentation using just the relations* (A1), (A2), *and* (B0).

Proof. Using (A0), the correspondence $y_\alpha(t) \mapsto x_\alpha(t)$ $(\alpha \in \Psi,\ t \in GF(q))$ defines a homomorphism ϕ from G_0 onto G. We have to show it is one-to-one.

It suffices to show that ϕ is one-to-one on each of the rank 2 subgroups corresponding to a pair of fundamental roots, because every relation in the Curtis–Steinberg–Tits presentation occurs inside one of these rank 2 subgroups. More precisely, we proceed as follows:

Consider one of the rank 2 subsystems Φ_{ij}. Let Φ_{ij}^+ be an arbitrary positive system of roots in Φ_{ij} (not necessarily the one induced by our fixed positive system Φ^+). Let $G_0(\Phi_{ij}^+)$ be the subgroup of G_0 generated by the set $S = \{y_\alpha(b_\nu) : \alpha \in \Phi_{ij}^+, 1 \le \nu \le m\}$.

Claim 4.3. ϕ is one-to-one on $G_0(\Phi_{ij}^+)$.

Proof. Let α_i and α_j be the two fundamental roots in Φ_{ij}^+. Then every member of Φ_{ij}^+ is a unique linear combination of α_i and α_j with nonnegative integer coefficients [Ste67, Appendix I, Proposition 9]. Order Φ_{ij}^+ in increasing order of the sum of these coefficients and extend this ordering to S. Then, by (A2) and (B0) (with (A0)) together with (A1), it follows readily that $G_0(\Phi_{ij}^+)$ is a p-group of order at most $q^{|\Phi_{ij}^+|}$.

On the other hand, the image of $G_0(\Phi_{ij}^+)$ under ϕ is a Sylow p-subgroup of the rank 2 subgroup of G corresponding to Φ_{ij} ([Ste67, corollary after Lemma 54, p. 132; cf. Lemma 21, p. 31], applied to the rank 2 Chevalley group G_{ij}) and thus its order is $q^{|\Phi_{ij}^+|}$ (cf. (2)). This proves Claim 4.3. ∎

Now consider any relation (B) for a pair $(\alpha, \beta) \in \Upsilon_{ij}$. We claim that the corresponding relation $[y_\alpha(t), y_\beta(u)] = \prod_{i,j>0} y_{i\alpha+j\beta}(C_{i,j,\alpha,\beta}t^iu^j)$ holds in G_0. Indeed, since $\alpha \ne \pm\beta$, there exists a positive system Φ_{ij}^+ in Φ_{ij} such that $\alpha, \beta \in \Phi_{ij}^+$. By Claim 4.3, any relation involving only these positive roots is satisfied simultaneously in $G_0(\Phi_{ij}^+)$ and its image in G.

Thus, we have shown that all relations (B) for $(\alpha, \beta) \in \Upsilon$ hold in G_0. Similarly, each relation (A) for $\alpha \in \Psi$ occurs in some subsystem Φ_{ij}^+ (alternatively, these relations also follow more directly from the relations (A1), (A2), and (A0)).

Thus, we have shown that all relations of the Curtis–Steinberg–Tits presentation for G also hold in G_0, completing the proof of Theorem 4.2. ∎

4.3. *Length of the Presentation*

The presentation (A1), (A2), (B0) involves $|\Psi|m$ generators. We show that $|\Psi|$ is $O(n)$ (as opposed to $|\Phi| = 2N$ which is $\Theta(n^2)$). This can be proved as follows. First observe that any rank 2 root system (in particular any one of the subsystems Φ_{ij}) contains a bounded number of roots (at most 12). Next observe that there are only $O(n)$ edges in the Dynkin diagram for any root system Φ of rank n, i.e., only $O(n)$ "nontrivial" subsystems Φ_{ij} if we call a subsystem of type $A_1 \times A_1$ "trivial." Then observe that the "trivial" subsystems Φ_{ij} corresponding to nonedges in the Dynkin diagram (i.e., the $O(n^2)$ different subsystems of type $A_1 \times A_1$, corresponding to pairs of root subgroups which centralize each other) contribute nothing new to the set $\Psi = \bigcup_{i<j} \Phi_{ij}$ since $\Phi_{ij} = \{\pm \alpha_i, \pm \alpha_j\}$ already occurs in the union of the "nontrivial" subsystems. Therefore $|\Psi|$ is $O(n)$.

Also observe that $|\Upsilon|$ is $\Theta(n^2)$, because each $|\Phi_{ij}|$ and hence each $|\Upsilon_{ij}|$ is bounded.

The length of each of the $|\Psi|m$ relations (A1) is $\log p + O(1)$; each of the $O(|\Psi|m^2)$ relations (A2) has bounded length. The number of (B0) relations is $|\Upsilon|m^2$, which is $\Theta(n^2m^2)$; the length of each is dominated by the length of the expansion of $y_\gamma(t)$ according to (A0), i.e., $O(m \log p)$. The total length of this presentation is therefore

$$\begin{aligned} &O(|\Psi|m \cdot \log p) + O(|\Psi|m^2) + \Theta\big(|\Upsilon|m^2 \cdot m \log p\big) \\ &\quad = \Theta\big(nm \log p + nm^2 + n^2m^3 \log p\big) \\ &\quad = \Theta\big(n^2m^3 \log p\big) = \Theta\big(m^2 \log|G|\big) \end{aligned}$$

(using (3) and the fact that N is $\Theta(n^2)$). This proves

COROLLARY 4.4. *Each universal Chevalley group G of rank $n \geq 2$ (given by its standard name, using $(*)$) has an explicit presentation of length*

$$O\big(m^2 \log|G|\big) \leq O\big(m \log^2|G|\big) \leq O\big(\log^3|G|\big).$$

While we use the term "explicit" in an informal sense here, we certainly mean to say that Conjecture 3 holds for these (nonsimple) groups, even without assuming (b2) (primitive root as part of the input). The primitive roots will be used in order to eliminate the center (Section 5.2).

5. DECREASING THE LENGTH AND ELIMINATING THE CENTER

In this section we let θ denote a generator of $GF(q)$ over $GF(p)$, and choose our basis B of $GF(q)$ over $GF(p)$ to consist of the first m powers of θ. Thus we let $b_\nu = \theta^{\nu-1}$ $(1 \le \nu \le m)$, and our generators $y_\alpha(b_\nu)$ will correspond to $x_\alpha(\theta^{\nu-1})$.

5.1. $O(\log|G|)$ *Presentation*

For $\alpha \in \Psi$ and $t \in GF(q)$, $t \neq 0$, following [Ste67, p. 27], we let

$$w_\alpha(t) = y_\alpha(t)\, y_{-\alpha}(-t^{-1})\, y_\alpha(t) \tag{4}$$

and

$$h_\alpha(t) = w_\alpha(t) w_\alpha(1)^{-1}. \tag{5}$$

(Here again we implicitly use (A0) to expand $y_\alpha(t)$, etc.) Steinberg proves that

$$h_\alpha(tu) = h_\alpha(t) h_\alpha(u) \qquad (\alpha \in \Psi, t, u \in GF(q) - \{0\}) \tag{6}$$

follows from his relations [Ste67, Theorem 9, p. 72], hence it follows from our reduced system of relations as well. Moreover, all of the $h_\alpha(t)$ commute; they generate a *Cartan subgroup* H of G.

Furthermore, the relations

$$h_\beta(\theta) x_\alpha(t) h_\beta(\theta)^{-1} = x_\alpha(\theta^{2(\alpha,\beta)/(\beta,\beta)} t) \tag{7}$$

also hold in the Chevalley group [Ste67, Lemma 20(c), p. 29], for any $\alpha, \beta \in \Phi$, not necessarily distinct (here (α, β) denotes the usual inner product of roots, so that $2(\alpha,\beta)/(\beta,\beta)$ is always an integer of absolute value ≤ 3).

We now make use of these relations to further shorten our presentation (at the same time adapting it to allow the center to be easily eliminated).

We use $|\Psi|$ additional generators denoted h_α (for $\alpha \in \Psi$). We add the relations

(H0)

$$h_\alpha = y_\alpha(\theta) y_{-\alpha}(-\theta^{-1}) y_\alpha(\theta) \big(y_\alpha(1) y_{-\alpha}(-1) y_\alpha(1)\big)^{-1} \qquad (\alpha \in \Psi);$$

note that the right-hand side is just $h_\alpha(\theta)$, expanded according to (4) and (5), and it must be further expanded according to definition (A0) (but $b_1 = 1$ and $b_2 = \theta$ so most of the above terms do not need any further expansion); after that expansion each of these $|\Psi|$ relations has length $O(m \log p)$. We also add the relations

(H1)

$$[h_\alpha, h_\beta] = 1 \qquad (\alpha, \beta \in \Psi);$$

(H2)

$$h_\beta y_\alpha(b_\nu) h_\beta^{-1} = y_\alpha\big(\theta^{2(\alpha,\beta)/(\beta,\beta)} b_\nu\big) \qquad (\alpha, \beta \in \Psi, 1 \le \nu \le m).$$

Here we expand the right-hand side of (H2) using (A0) if necessary. However, $\theta^d b_\nu = b_{\nu+d}$ for any ν and d such that $1 \le \nu + d \le m$; in (H2) this occurs with $d = 2(\alpha, \beta)/(\beta, \beta)$, hence $|d| \le 3$, so, for each α, β, expansion is required only for a *bounded* number of values of ν. Thus, most of the relations (H2) have bounded length.

Now with these additional relations we may omit many of the relations from our earlier set (A1), (A2), and (B0), because most of those will follow from a few of them together with conjugation by the various h_β using (H2).

THEOREM 5.1. *Each universal Chevalley group G of rank $n \ge 2$ has an explicit presentation of length $O(\log|G|)$ using the following relations*:

(A1′) $$y_\alpha(b_\nu)^p = 1 \qquad (\alpha \in \Psi, \nu \in \{1,2\});$$

(A2′) $$[y_\alpha(b_\nu), y_\alpha(b_\mu)] = 1 \qquad (\alpha \in \Psi, \nu \in \{1,2\}, 1 \le \mu \le m);$$

(B0′) $$[y_\alpha(b_\nu), y_\beta(b_\mu)] = \prod_{i,j>0} y_{i\alpha+j\beta}(C_{i,j,\alpha,\beta} b_\nu^i b_\mu^j)$$
$$((\alpha,\beta) \in \Upsilon, \nu, \mu \in \{1,2\}),$$

together with the relations (H0), (H1), *and* (H2) *above*.

Proof. The relations (A1′), (A2′), and (B0′) are subsets of the relations (A1), (A2), and (B0) from Section 4.2, obtained by simply restricting the values of ν (as well as μ in (B0′)). The additional relations (H0)–(H2) hold in the Chevalley group G (as noted earlier in this section), thus to prove that the above relations define G it suffices (by Theorem 4.2) to show that the remaining relations (A1), (A2), and (B0) follow from the ones above.

In the case $\alpha = \beta$, (H2) says that $h_\alpha y_\alpha(b_\nu) h_\alpha^{-1} = y_\alpha(\theta^2 b_\nu) = y_\alpha(b_{\nu+2})$ (assuming that $\nu + 2 \le m$). Thus, the relations (A1′) together with (H2) imply all of the relations (A1), and similarly the relations (A2′) imply (A2). Finally, the relations (B0′) imply all relations (B0), because all other pairs (ν, μ) can be obtained from the four pairs in $\{1,2\} \times \{1,2\}$, using conjugation by $h_\beta^{c_1} h_\alpha^{c_2}$, where the integers c_1 and c_2 are chosen as in the following lemma.

LEMMA 5.2. *Let α and β be two roots from any root system, with $\alpha \ne \pm\beta$, and let (ν, μ) be an arbitrary pair of integers. Then there are integers c_1, c_2 such that*

$$(c_1 \cdot 2(\alpha,\beta)/(\beta,\beta) + 2c_2, 2c_1 + c_2 \cdot 2(\beta,\alpha)/(\alpha,\alpha))$$
$$\in \{\nu - 1, \nu - 2\} \times \{\mu - 1, \mu - 2\}.$$

Proof. Let $\langle \alpha, \beta \rangle$ denote $2(\alpha, \beta)/(\beta, \beta)$ and similarly for $\langle \beta, \alpha \rangle$, and write $v(c_1, c_2) = (c_1\langle \alpha, \beta \rangle + 2c_2, 2c_1 + c_2\langle \beta, \alpha \rangle)$. Then we are looking at the lattice $\Lambda = \{v(c_1, c_2) \mid c_1, c_2 \in \mathbb{Z}\}$. Because of the way they arise from a root system, the numbers $\langle \alpha, \beta \rangle$ and $\langle \beta, \alpha \rangle$ are either both 0, or else one of them is ± 1 and the other is 1, 2, or 3 times that one (see [Hum72, Table 1, Sect. 9.4, p. 45]). If they are both 0 the conclusion is clear. Otherwise, replacing β by $-\beta$ if needed, we may assume that $\langle \alpha, \beta \rangle = 1$.

If $\langle \beta, \alpha \rangle = 1$, then $v(2, -1) = (0, 3)$, $v(-1, 2) = (3, 0)$, and $v(1, -1) = (-1, 1)$. In this case $\Lambda = \{(x, y) \mid x + y \equiv 0 \pmod 3\}$.

If $\langle \beta, \alpha \rangle = 2$, then $v(2, -1) = (0, 2)$ and $v(-2, 2) = (2, 0)$, so this time we can reach any pair of even integers.

Finally, if $\langle \beta, \alpha \rangle = 3$, then $v(2, -1) = (0, 1)$ and $v(-3, 2) = (1, 0)$, so we can reach *every* pair of integers. ▮

Thus we have a presentation for G using only $2|\Psi| = O(n)$ relations (A1′) (each of length $O(\log p)$), only $O(|\Psi|m) = O(nm)$ relations (A2′) (each of bounded length), and only $4|\Upsilon| = O(n^2)$ relations (B0′) (each of length $O(m \log p)$ as before). In addition, we have $|\Psi| = O(n)$ relations (H0), each of length $O(m \log p)$, together with $|\Psi|^2 = O(n^2)$ relations (H1) (each of bounded length) and $|\Psi|^2 m = O(n^2 m)$ relations (H2). As noted above, the relations (H2) have bounded length, except for a bounded number of values of ν (for each pair $\alpha, \beta \in \Psi$) where (H2) has length $O(m \log p)$ because of the expansion (A0). Thus, the total length of all the relations (H2) is $O(n^2 m) + O(n^2 \cdot m \log p) = O(n^2 m \log p)$. Thus the length of the entire presentation (A1′), (A2′), (B0′), (H0), (H1), and (H2) is

$$O\big(n \cdot \log p + nm + n^2 \cdot m \log p + n \cdot m \log p + n^2 + n^2 m \log p\big) = O\big(n^2 m \log p\big) = O(\log|G|). \quad ▮$$

5.2. *Killing the Center*

Finally, we now add one or two more relations (of length $O(\log|G|)$) to eliminate the center Z and turn the presentation into one for the simple Chevalley group G/Z. (Here $|Z| \le n + 1$ is negligible compared to $|G|$, so the total length of the resulting presentation is both $O(\log|G/Z|)$ and $O(\log|G|)$.) Now we shall make the additional assumption that the element θ used above is actually a primitive root of $GF(q)$.

Specifically, the center Z is contained in the Cartan subgroup H and thus [Car72, Sect. 12.1] consists of elements of the form

$$h_{\alpha_1}(t_1) \cdots h_{\alpha_n}(t_n) \tag{8}$$

(recall (5)), where $\alpha_1, \dots, \alpha_n$ are the fundamental roots in Φ^+ and the elements $t_i \in GF(q)^*$ are chosen as discussed below. Since θ is a primitive root (so every element of $GF(q)^*$ has the form θ^k with $0 \le k < q-1$) and $h_\alpha(\theta^k) = h_\alpha(\theta)^k$ by (6), we can express any element (8) in terms of our generators h_α (recall (H0)) with a relation of length $O(n \log q) = O(nm \log p) \le O(\log|G|)$. Furthermore, Z is cyclic except in some cases of type $D_n(q)$ (orthogonal groups) where it may be noncyclic of order 4, so we can eliminate Z by adding only one or two such additional relations (but even if we added one for each element of Z we would still have total length $O(\log|G|)$).

So it only remains to consider which values of t_i put (8) into the center Z. These are easily found by matrix computations in specific cases; in general (see [Car72, Sect. 12.1]), Z consists of exactly those elements (8) for which

$$\prod_{i=1}^{n} t_i^{2(\alpha_i,\,\alpha_j)/(\alpha_i,\,\alpha_i)} = 1 \qquad \text{for } j = 1, \dots, n. \tag{9}$$

EXAMPLE 5.3. In the case $A_n(q)$ the equation (9) are $t_1^2 t_2^{-1} = 1$, $t_{i-1}^{-1} t_i^2 t_{i+1}^{-1} = 1$ $(2 \le i \le n-1)$, and $t_{n-1}^{-1} t_n^2 = 1$; or, equivalently, $t_i = t_1^i$ $(2 \le i \le n)$ and $t_1^{n+1} = 1$. From this, given θ, it is easy to find the appropriate $t_1 \in GF(q)$ that produces a generator of Z. The other groups $\mathsf{L}_n(q)$ can be handled in a similar manner.

Since θ is a primitive root of $GF(q)$, (9) amounts to a system of linear equations. Namely, if we let z_i denote the integer for which

$$t_i = \theta^{z_i} \qquad \text{and} \qquad 0 \le z_i \le q-2,$$

then (9) is replaced by the following system of n congruences:

$$\sum_{i=1}^{n} \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)} z_i \equiv 0 \qquad (\mathrm{mod}\ q-1)\ (j = 1, \dots, n). \tag{10}$$

This system can be solved (mod $q-1$) for the z_i (this is especially easy since most of the Cartan integers $2(\alpha_i, \alpha_j)/(\alpha_i, \alpha_i)$ are 0), and the solutions give the (at most $n+1$) elements of the form (8) in Z. Using these (or enough to generate Z) together with the set of relations in Theorem 5.1, we obtain a presentation of G/Z. This completes the proof of

THEOREM 5.4. *Conjecture* 3 *holds* (*and consequently Conjectures* 1 *and* 2 *hold*), *with* $C = 1$, *for simple Chevalley groups of rank* ≥ 2.

6. TWISTED GROUPS OF RANK ≥ 2

The twisted groups have presentations similar to those discussed above for the untwisted groups. Details of Steinberg's presentations for central extensions of the twisted simple groups can be found in [Gri73]. These presentations can be reduced in length, and supplemented to eliminate the center, in essentially the same way as described in the previous sections. However, a few additional complications arise, especially for the odd-dimensional unitary groups $PSU(2k+1, q) = {}^2A_{2k}(q)$ and the Ree groups ${}^2F_4(q)$, where we only obtain a presentation of length $O(\log^2|G|)$, not $O(\log|G|)$ as in the other cases.

THEOREM 6.1. *Conjecture* 3 *holds for the twisted simple groups of rank* $\geq$ 2. *It holds with* $C = 1$, *except perhaps for the groups* ${}^2A_{2k}(q)$ *and* ${}^2F_4(q)$, *where it holds with* $C = 2$.

This is proved in the following subsections.

6.1. *Easier Twisted Groups*

First, we consider types ${}^2A_n(q)$ with n odd (the case of even n will be discussed in Section 6.3), ${}^2D_n(q)$ (for any $n \geq 4$), and ${}^2E_6(q)$. In these cases the "twisted" root system Φ has type C_{k+1} (where $n = 2k+1$), B_{n-1}, and F_4, respectively. In each of these cases the long roots of Φ correspond to root subgroups of order q, while the short roots correspond to root subgroups of order q^2 (and all the root subgroups are elementary abelian as in the untwisted case). Thus, in these cases we start with generators $x_\alpha(t)$, where $\alpha \in \Phi$, and $t \in GF(q)$ if α is long while $t \in GF(q^2)$ if α is short. The Steinberg relations are then as follows:

(A)
$$x_\alpha(t+u) = x_\alpha(t)x_\alpha(u)$$
$$\big(\alpha \in \Phi, t, u \in GF(q) \text{ or } GF(q^2) \text{ for } \alpha \text{ long resp. short}\big);$$

(B)
$$[x_\alpha(t), x_\beta(u)]$$
$$= \begin{cases} 1 & \text{for } \alpha+\beta \notin \Phi, \\ x_{\alpha+\beta}(\epsilon_{\alpha\beta}tu) & \text{for } \alpha, \beta, \alpha+\beta \text{ all short or all long}, \\ x_{\alpha+\beta}\big(\epsilon_{\alpha\beta}(t\bar{u} + \bar{t}u)\big) & \text{for } \alpha, \beta \text{ short}, \alpha+\beta \text{ long}, \\ x_{\alpha+\beta}(\epsilon_{\alpha\beta}tu)x_{\alpha+2\beta}(\eta_{\alpha\beta}tu\bar{u}) & \text{for } \alpha, \alpha+2\beta \text{ long}, \beta, \alpha+\beta \text{ short}. \end{cases}$$

Here $u \mapsto \bar{u}$ denotes the involutory field automorphism of $GF(q^2)$, and the coefficients $\epsilon_{\alpha\beta}$ and $\eta_{\alpha\beta}$ are ± 1 and depend only on α, β (and Φ).

As in the untwisted case, we only need a subset of the above relations (by [Cur65]; see Remark 3.2), where we consider only roots $\alpha \in \Psi$ and, for (B), only pairs of roots $(\alpha, \beta) \in \Upsilon$, where the subsets $\Psi \subseteq \Phi$ and $\Upsilon \subseteq \Psi \times \Psi$ are defined as in Section 4.

We then shorten this presentation essentially as in Sections 4.2 and 5. This time we use a basis $\{b_1, \dots, b_{2m}\}$ of $GF(q^2)$ over $GF(p)$ (where $q = p^m$), chosen so that its first half $\{b_1, \dots, b_m\}$ is a basis of $GF(q)$ over $GF(p)$. Specifically, let θ be a primitive root of $GF(q^2)$, so $\theta^{q+1} = \theta\bar{\theta}$ is a primitive root of $GF(q)$; then, for $1 \le \nu \le m$, let $b_\nu = (\theta\bar{\theta})^{\nu-1}$ and $b_{m+\nu} = \theta b_\nu$. As before we define

(A0)

$$y_\alpha(t) := \prod_\nu y_\alpha(b_\nu)^{k_\nu} \qquad \text{whenever } t = \sum_\nu k_\nu b_\nu \ (0 \le k_\nu < p),$$

where $t \in GF(q)$ and $1 \le \nu \le m$ if α is long, but $t \in GF(q^2)$ and $1 \le \nu \le 2m$ if α is short.

We define a group G_0, generated by the symbols $y_\alpha(b_\nu)$ ($\alpha \in \Psi$, and $1 \le \nu \le m$ for α long, $1 \le \nu \le 2m$ for α short), with additional generators h_α ($\alpha \in \Psi$) and the following relations:

(A1) $y_\alpha(b_\nu)^p = 1$ ($\alpha \in \Psi$, $\nu = 1, 2$, and also $\nu = m+1, m+2$ if α is short);

(A2) $[y_\alpha(b_\nu), y_\alpha(b_\mu)] = 1$ ($\alpha \in \Psi$, ν as in (A1), $1 \le \mu \le m$ or $1 \le \mu \le 2m$ for α long resp. short);

(B0) The relations (B) above for pairs $(\alpha, \beta) \in \Upsilon$, except with x replaced by y throughout and t, u taking only the values b_ν, b_μ, with $\nu, \mu \in \{1, 2\}$ or $\{1, 2, m+1, m+2\}$ as appropriate (as before, the right-hand sides of (B0) must be expanded, using (A0), into expressions involving the generators $y_\alpha(b_\nu)$);

(H0) $h_\alpha = h_\alpha(\theta\bar{\theta})$ ($\alpha \in \Psi$), where the right-hand side of (H0) is expanded according to the definitions (5), (4), and (A0) as before;

(H1) $[h_\alpha, h_\beta] = 1$ ($\alpha, \beta \in \Psi$);

(H2) $h_\beta y_\alpha(b_\nu) h_\beta^{-1} = y_\alpha((\theta\bar{\theta})^{2(\alpha,\beta)/(\beta,\beta)} b_\nu)$ ($\alpha, \beta \in \Psi$; $1 \le \nu \le m$ or $1 \le \nu \le 2m$ for α long resp. short)

As before we expand the right-hand side of (H2) using (A0) if necessary, but $(\theta\bar{\theta})^d b_\nu = b_{\nu+d}$ in most cases (whenever ν and $\nu + d$ are either both between 1 and m or both between $m+1$ and $2m$), thus, as before, expansion is required only for a bounded number of values of ν.

To see that the above relations hold in G, only (H2) requires some additional remarks. Each twisted group is defined as a subgroup of the corresponding untwisted Chevalley group (e.g., ${}^2A_n(q) \subset A_n(q^2)$, and similarly in the other cases). Furthermore, each “root subgroup” of the twisted group corresponds to parts of one or more root subgroups in the untwisted group (cf. [Car72, Sect. 13.6]; for α short, our $x_\alpha(t)$ here corresponds to $x_S(t) = x_r(t)x_{\bar{r}}(\bar{t})$ in Carter's notation [Car72, 13.6.4(ii)], and our $h_\alpha(u)$ corresponds to $h_r(u)h_{\bar{r}}(\bar{u})$ [Car72, 13.7.2]). Thus the relations (H2) can be verified in G by straightforward calculations using the corresponding relations for the untwisted group. However, we can also see that they must have the simple form given above by reasoning as follows: The conjugate $h_\beta(u)x_\alpha(t)h_\beta(u)^{-1}$ has the form $x_\alpha(f(u)t)$ for some function f, because of the way this occurs inside a larger untwisted group. In (H2) we only need to know $f(u)$ for $u \in GF(q)$ (though $u \in GF(q^2)$ may occur in general), and this case has the same formula as in the untwisted cases (i.e., $f(u) = u^{2(\alpha,\beta)/(\beta,\beta)}$ for $u \in GF(q)$; see (7) in Section 5.1). This holds because this formula is determined inside another untwisted Chevalley group, namely the subgroup of our twisted group G obtained by restricting all scalars to $GF(q)$. That gives an untwisted group with root system Φ (e.g., $C_{k+1}(q) \subset {}^2A_{2k+1}(q) \subset A_{2k+1}(q^2)$, or $F_4(q) \subset {}^2E_6(q) \subset E_6(q^2)$). This can be seen from the fact that it is the intersection of the (untwisted) subgroups of fixed points of a pure graph automorphism and a pure field automorphism (whereas the twisted group is only fixed by the composition of the graph and field automorphisms; cf. [Ste67, p. 171]). Alternatively, observe that restricting all scalars to $GF(q)$ in the above Steinberg presentation for the twisted group produces the Steinberg presentation for the untwisted group with root system Φ.

Thus, all of the above relations hold in the twisted group G. Then, just as in the previous sections, it is easy to prove that the group G_0 defined by the above presentation is isomorphic to the group presented by the Steinberg relations above (which is a central extension of the twisted simple group in question).

The center Z can be eliminated as before: [Ste67, proof of Theorem 35(b), p. 192; cf. Theorem 34 and following exercise, p. 187] shows that the center of the twisted group G is contained in the center of the corresponding untwisted universal Chevalley group. Therefore the elements of Z can be expressed as in Section 5.2, in terms of generators for the Cartan subgroup of the untwisted group (which is defined over $GF(q^2)$). But Z lies inside G, so in fact it can be expressed in terms of generators for the Cartan subgroup of G, namely [Car72, 13.7.2] the elements $h_\alpha(\theta)$ for α short (as noted above, these $h_\alpha(t)$ correspond to $h_r(t)h_{\bar{r}}(\bar{t})$ in Carter's notation) and $h_\alpha(\theta\bar{\theta})$ for α long (these $h_\alpha(t)$ correspond to those $h_r(t)$ which are contained in the twisted group if and only if $t = \bar{t}$). Thus, for α long, we may use our generators h_α from (H0) above, but for α short we introduce additional generators h'_α and additional relations

(H0′) $h'_\alpha = h_\alpha(\theta)$ ($\alpha \in \Psi$, α short), with the right-hand side of (H0′) expanded according to (5), (4), and (A0) as before. Then, using these generators h'_α for α short and h_α for α long, we may kill the center as in Section 5.2.

Finally, observe that this presentation again has length $O(\log|G|)$. This follows essentially as before, with the following differences: the length of the presentation involves $|\Psi|$ and $|\Upsilon|$, which are determined by the root system Φ, but the parameter N for which $q^{2N} < |G| < q^{3N}$ is the number of positive roots of the untwisted root system, which is now larger than Φ since we are using Φ to denote the twisted root system. But $|\Phi|$ and N are both still $\Theta(n^2)$. Also, the presentation now involves $2m$ elements b_ν instead of just m of them, and additional relations (H0′) of length similar to (H0), but these differences only change the total length by a constant factor. Therefore, the presentation still has length $O(\log|G|)$.

6.2. *Triality Twisted Groups*

The group ${}^3D_4(q)$, with "twisted" root system Φ of type G_2, is handled similarly. Since Φ now has rank 2, we have $\Psi = \Phi$ in this case. This time we start with generators $x_\alpha(t)$, where either α is long and $t \in GF(q)$, or α is short and $t \in GF(q^3)$. We use $t \mapsto t^\sigma$ to denote the field automorphism of $GF(q^3)$ of order 3, and we use the notation $\mathrm{Tr}(t) = t + t^\sigma + t^{\sigma^2}$ for the trace map Tr: $GF(q^3) \to GF(q)$. Then the Steinberg relations are:

(A)
$$x_\alpha(t+u) = x_\alpha(t)x_\alpha(u)$$
$$\big(\alpha \in \Phi,\, t, u \in GF(q) \text{ or } GF(q^3) \text{ for } \alpha \text{ long resp. short}\big);$$

(B)
$$[x_\alpha(t), x_\beta(u)]$$
$$= \begin{cases} 1 & \text{for } \alpha+\beta \notin \Phi, \\ x_{\alpha+\beta}(\epsilon_{\alpha\beta} tu) & \text{for } \alpha, \beta, \alpha+\beta \text{ long}, \\ x_{\alpha+\beta}(\epsilon_{\alpha\beta}\mathrm{Tr}(tu)) & \text{for } \alpha, \beta \text{ short}, \alpha+\beta \text{ long}, \\ x_{\alpha+\beta}\big(\epsilon_{\alpha\beta}(t^\sigma u^{\sigma^2} + t^{\sigma^2}u^\sigma)\big)x_{2\alpha+\beta}\big(\eta_{\alpha\beta}\mathrm{Tr}(tt^\sigma u^{\sigma^2})\big) & \\ \quad \times x_{\alpha+2\beta}\big(\delta_{\alpha\beta}\mathrm{Tr}(tu^\sigma u^{\sigma^2})\big) & \text{for } \alpha, \beta, \alpha+\beta \text{ short}, \\ & \qquad 2\alpha+\beta, \alpha+2\beta \text{ long}, \\ x_{\alpha+\beta}(\epsilon_{\alpha\beta}tu)x_{2\alpha+\beta}\big(\eta_{\alpha\beta}t^\sigma t^{\sigma^2}u\big)x_{3\alpha+\beta}\big(\delta_{\alpha\beta}tt^\sigma t^{\sigma^2}u\big) & \\ \quad \times x_{3\alpha+2\beta}\big(2\gamma_{\alpha\beta}tt^\sigma t^{\sigma^2}u^2\big) & \text{for } \alpha, \alpha+\beta, 2\alpha+\beta \text{ short}, \\ & \qquad \beta, 3\alpha+\beta, 3\alpha+2\beta \text{ long}. \end{cases}$$

Again the coefficients $\epsilon_{\alpha\beta}, \eta_{\alpha\beta}, \delta_{\alpha\beta}, \gamma_{\alpha\beta}$ are ± 1 and depend only on α and β.

We shorten this presentation in a way completely analogous to the cases in Section 6.1. This time let θ be a primitive root of $GF(q^3)$, so $\theta\theta^{\sigma}\theta^{\sigma^2}$ is a primitive root of $GF(q)$, and use the basis $b_{\nu} = (\theta\theta^{\sigma}\theta^{\sigma^2})^{\nu-1}$, $b_{m+\nu} = \theta b_{\nu}$, and $b_{2m+\nu} = \theta^2 b_{\nu}$ (for $1 \leq \nu \leq m$). Then use relations (A1) through (H2) just as above, except replacing $2m$ by $3m$ (and $\{1, 2, m+1, m+2\}$ by $\{1, 2, m+1, m+2, 2m+1, 2m+2\}$) in the obvious places, and replacing $\theta\bar{\theta}$ by $\theta\theta^{\sigma}\theta^{\sigma^2}$ in (H0) and (H2).

In this case the universal central extension of ${}^3D_4(q)$ is actually the simple group itself, since ${}^3D_4(q)$ has trivial Schur multiplier [KL90, Theorem 5.1.4, p. 173]. Thus, we may omit the generators h'_{α} and relations (H0′) in this case. But even if we include them, counting as before, we find that our presentation has length $O(\log q) = O(\log|G|)$.

6.3. *Odd-Dimensional Unitary Groups*

The groups ${}^2A_n(q)$ with n even are somewhat harder (these are the odd-dimensional unitary groups $PSU(n+1, q)$). Let $n = 2k$ ($k \geq 2$ since we are only considering rank ≥ 2 in this section). For many purposes the root system of this twisted group is best viewed as BC_k, namely, the union of a B_k and C_k system. However, following Griess [Gri73], we will just use a C_k system Φ.

An additional complication in this case is that the root subgroups are not all abelian. Those corresponding to short roots of Φ are elementary abelian of order q^2, but those corresponding to long roots are nonabelian of order q^3. We start with generators $x_{\alpha}(t)$ for $\alpha \in \Phi$ short and $t \in GF(q^2)$, and $x_{\alpha}(t, u)$ for $\alpha \in \Phi$ long with $t, u \in G(q^2)$ and $u + \bar{u} = \epsilon_{\alpha} t\bar{t}$, where $\epsilon_{\alpha} = \pm 1$ are constants depending on the root system and $u \mapsto \bar{u}$ again denotes the involutory field automorphism of $GF(q^2)$. For relations (A) we have

(A) $x_{\alpha}(t+u) = x_{\alpha}(t) x_{\alpha}(u)$ (as before) for α short,

but for α long we have

$$(\mathrm{A}') \qquad \begin{aligned} x_{\alpha}(t, u) x_{\alpha}(v, w) &= x_{\alpha}\big(t + v, u + w + \epsilon_{\alpha}\bar{t}v\big), \\ \big[x_{\alpha}(t, u), x_{\alpha}(v, w)\big] &= x_{\alpha}\big(0, \epsilon_{\alpha}(\bar{t}v - t\bar{v})\big). \end{aligned}$$

(See [Car72, 13.6.4(iv)]; the version in [Gri73] occasionally has $\mp\epsilon$ instead of our $\pm\epsilon_{\alpha}$.)

The relations (B) for this case (G of type ${}^2A_{2k}(q)$) are as follows:

(B)

$$[x_\alpha(t), x_\beta(u)]$$
$$= \begin{cases} 1 & \text{for } \alpha, \beta \text{ short, } \alpha+\beta \notin \Phi, \\ x_{\alpha+\beta}(\epsilon_{\alpha\beta} tu) & \text{for } \alpha, \beta, \alpha+\beta \text{ short,} \\ x_{\alpha+\beta}\big(0, \epsilon_{\alpha\beta}(t\bar{u} - \bar{t}u)\big) & \text{for } \alpha, \beta \text{ short, } \alpha+\beta \text{ long,} \end{cases}$$

$$[x_\alpha(t,u), x_\beta(v,w)]$$
$$= \begin{cases} 1 & \text{for } \alpha, \beta \text{ long, } \tfrac{1}{2}(\alpha+\beta) \notin \Phi, \\ x_{\frac{1}{2}(\alpha+\beta)}(\epsilon_{\alpha\beta} tv) & \text{for } \alpha, \beta \text{ long, } \tfrac{1}{2}(\alpha+\beta) \text{ short,} \end{cases}$$

$$[x_\alpha(t,u), x_\beta(v)]$$
$$= \begin{cases} 1 \quad \text{for } \alpha \text{ long, } \beta \text{ short, } \alpha+\beta \notin \Phi, \\ x_{\alpha+\beta}(\epsilon_{\alpha\beta}\bar{u}v) x_{\alpha+2\beta}(\eta_{\alpha\beta} tv, \delta_{\alpha\beta} uv\bar{v}) \\ \qquad \text{for } \alpha, \alpha+2\beta \text{ long, } \beta, \alpha+\beta \text{ short.} \end{cases}$$

(Again the coefficients $\epsilon_{\alpha\beta}$, $\eta_{\alpha\beta}$, $\delta_{\alpha\beta}$ are ± 1 and depend only on α and β.)

We will shorten these relations essentially as in Section 6.1, but some differences arise. The reduction from Φ to Ψ (and Υ) is as before. Let θ be a primitive root of $GF(q^2)$ and let $\{b_\nu : 1 \le \nu \le 2m\}$ be a basis defined in terms of θ as in Section 6.1. Let ξ denote a fixed nonzero element of $GF(q^2)$ for which $\bar{\xi} = -\xi$ (if q is odd, then let $\xi = \theta^{(q+1)/2}$, while if q is even let $\xi = 1$).

Then, for a long root α, the root subgroup elements of the form $x_\alpha(0,u)$ can be written as $x_\alpha(0, t\xi)$ with $t \in GF(q)$; these elements form a subgroup isomorphic to the additive group of $GF(q)$. Modulo that subgroup, the root subgroup is isomorphic to the additive group of $GF(q^2)$. Moreover, the commutator relations in the root subgroup are determined by the $GF(q)$-bilinear map $\bar{t}v - t\bar{v}$, and hence are completely determined on a basis of $GF(q^2)$ over $GF(p)$.

For each long root α, let $u_{\alpha,1}$ denote a fixed element of $GF(q^2)$ for which $u_{\alpha,1} + \bar{u}_{\alpha,1} = \epsilon_\alpha b_1 \bar{b}_1 = \epsilon_\alpha$ (recall that $b_1 = 1$), and similarly let $u_{\alpha,m+1}$ be an element for which $u_{\alpha,m+1} + \bar{u}_{\alpha,m+1} = \epsilon_\alpha b_{m+1}\bar{b}_{m+1} = \epsilon_\alpha \theta\bar{\theta}$. (Elements $u_{\alpha,1}$ and $u_{\alpha,m+1}$ with the desired trace can easily be computed.) Then define

$$u_{\alpha,\nu} = (\theta\bar{\theta})^{2\nu-2} u_{\alpha,1} \qquad \text{and} \qquad u_{\alpha,m+\nu} = (\theta\bar{\theta})^{2\nu-2} u_{\alpha,m+1}$$
$$\text{for } 2 \le \nu \le m.$$

Thus, we have $u_{\alpha,\nu} + \bar{u}_{\alpha,\nu} = \epsilon_\alpha b_\nu \bar{b}_\nu$ for all ν $(1 \le \nu \le 2m)$, so $x_\alpha(b_\nu, u_{\alpha,\nu})$ is an element of G (cf. (A0$''$) below).

Now for our short presentation we use the following generators. For each short root $\alpha \in \Psi$ we use $2m$ generators $y_\alpha(b_\nu)$ $(1 \le \nu \le 2m)$. For each long root $\alpha \in \Psi$ we use $3m$ generators, namely $y_\alpha(0, b_\nu \xi)$ for $1 \le \nu \le m$ and $y_\alpha(b_\nu, u_{\alpha,\nu})$ for $1 \le \nu \le 2m$. In addition, our presentation will use generators h_α and h'_α for each $\alpha \in \Psi$.

For short roots $\alpha \in \Psi$ we define (as before)

$$\text{(A0)} \qquad y_\alpha(t) := \prod_\nu y_\alpha(b_\nu)^{k_\nu}$$

$$\text{whenever } t = \sum_\nu k_\nu b_\nu \in GF(q^2) \ (0 \le k_\nu < p, 1 \le \nu \le 2m);$$

whereas for long roots $\alpha \in \Psi$ we define

$$\text{(A0}'\text{)} \qquad y_\alpha(0, u) := \prod_\nu y_\alpha(0, b_\nu \xi)^{k_\nu}$$

$$\text{whenever } u\xi^{-1} = \sum_\nu k_\nu b_\nu \in GF(q) \ (0 \le k_\nu < p, 1 \le \nu \le m);$$

$$\text{(A0}''\text{)} \qquad y_\alpha(t, u) := y_\alpha(0, v) \prod_\nu y_\alpha(b_\nu, u_{\alpha,\nu})^{k_\nu}$$

whenever $t = \sum_\nu k_\nu b_\nu \in GF(q^2)$ $(0 \le k_\nu < p;\ 1 \le \nu \le 2m)$ and $u + \bar{u} = \epsilon_\alpha t\bar{t}$, where b_ν and $u_{\alpha,\nu}$ are as defined above and $v \in GF(q^2)$ is the appropriate element to make this relation hold in G (i.e., $\prod_\nu x_\alpha(b_\nu, u_{\alpha,\nu})^{k_\nu}$ $= x_\alpha(t, w)$ for some $w \in GF(q^2)$ which has, from (A$'$), a known expression in terms of the b_ν, k_ν, $u_{\alpha,\nu}$, and ϵ_α; then let $v = u - w$ so that $x_\alpha(0, v) x_\alpha(t, w) = x_\alpha(t, u)$ as desired); the factor $y_\alpha(0, v)$ is itself expanded according to (A0$'$).

For each short root $\alpha \in \Psi$ we use the relations:

$$\text{(A1)} \qquad y_\alpha(b_\nu)^p = 1 \qquad (\nu = 1, 2, m+1, m+2);$$

$$\text{(A2)} \qquad [y_\alpha(b_\nu), y_\alpha(b_\mu)] = 1 \qquad (\nu = 1, 2, m+1, m+2, 1 \le \mu \le 2m).$$

On the other hand, for each long root $\alpha \in \Psi$ we use the following relations:

$$\text{(A3)} \qquad y_\alpha(0, b_\nu \xi)^p = 1 \qquad (\nu = 1, 2, 3, 4);$$

$$\text{(A4)} \qquad [y_\alpha(0, b_\nu \xi), y_\alpha(0, b_\mu \xi)] = 1 \qquad (\nu = 1, 2, 3, 4, 1 \le \mu \le m);$$

(A5)
$$y_\alpha(b_\nu, u_{\alpha,\nu})^p = \begin{cases} 1 & \text{for } p > 2, \\ y_\alpha\left(0, \epsilon_\alpha b_\nu \bar{b}_\nu\right) & \text{for } p = 2 \end{cases}$$
$$(\nu = 1, 2, m+1, m+2),$$

where, for $p = 2$, $y_\alpha(0, \epsilon_\alpha b_\nu \bar{b}_\nu)$ is expanded according to (A0′);

(A6)
$$\left[y_\alpha(b_\nu, u_{\alpha,\nu}), y_\alpha(b_\mu, u_{\alpha,\mu})\right] = y_\alpha\left(0, \epsilon_\alpha\left(\bar{b}_\nu b_\mu - b_\nu \bar{b}_\mu\right)\right)$$
$$(\nu = 1, 2, m+1, m+2,\ 1 \le \mu \le 2m),$$

where again the right-hand side is expanded according to (A0′);

(A7)
$$\left[y_\alpha(b_\nu, u_{\alpha,\nu}), y_\alpha(0, b_\mu \xi)\right] = 1 \qquad (\nu = 1, 2, m+1, m+2,\ 1 \le \mu \le m).$$

We use the following relations in place of (B):

(B0)
$$\left[y_\alpha(b_\nu), y_\beta(b_\mu)\right] = \begin{cases} 1 & \text{for } \alpha, \beta \text{ short, } \alpha + \beta \notin \Phi, \\ y_{\alpha+\beta}\left(\epsilon_{\alpha\beta} b_\nu b_\mu\right) & \text{for } \alpha, \beta, \alpha+\beta \text{ short,} \\ y_{\alpha+\beta}\left(0, \epsilon_{\alpha\beta}\left(b_\nu \bar{b}_\mu - \bar{b}_\nu b_\mu\right)\right) & \\ & \text{for } \alpha, \beta \text{ short, } \alpha+\beta \text{ long} \end{cases}$$
$$(\nu, \mu \in \{1, 2, m+1, m+2\}),$$

$$\left[y_\alpha(0, b_\nu \xi), y_\beta(0, b_\mu \xi)\right] = 1 \qquad \text{for } \alpha, \beta \text{ long } (\nu, \mu \in \{1, 2, 3, 4\}),$$

$$\left[y_\alpha(0, b_\nu \xi), y_\beta(b_\mu, u_{\beta,\mu})\right] = 1$$
$$\text{for } \alpha, \beta \text{ long } (\nu = 1, 2, 3, 4,\ \mu = 1, 2, m+1, m+2),$$

$$\left[y_\alpha(b_\nu, u_{\alpha,\nu}), y_\beta(0, b_\mu \xi)\right] = 1$$
$$\text{for } \alpha, \beta \text{ long } (\nu = 1, 2, m+1, m+2,\ \mu = 1, 2, 3, 4),$$

$$\left[y_\alpha(b_\nu, u_{\alpha,\nu}), y_\beta(b_\mu, u_{\beta\mu})\right]$$
$$= \begin{cases} 1 & \text{for } \alpha, \beta \text{ long, } \tfrac{1}{2}(\alpha+\beta) \notin \Phi, \\ y_{\frac{1}{2}(\alpha+\beta)}\left(\epsilon_{\alpha\beta} b_\nu b_\mu\right) & \text{for } \alpha, \beta \text{ long, } \tfrac{1}{2}(\alpha+\beta) \text{ short} \end{cases}$$
$$(\nu, \mu \in \{1, 2, m+1, m+2\}),$$

$$\left[y_\alpha(b_\nu, u_{\alpha,\nu}), y_\beta(b_\mu)\right]$$
$$= \begin{cases} 1 \quad \text{for } \alpha \text{ long, } \beta \text{ short, } \alpha+\beta \notin \Phi, \\ y_{\alpha+\beta}\left(\epsilon_{\alpha\beta} \bar{u}_{\alpha,\nu} b_\mu\right) y_{\alpha+2\beta}\left(\eta_{\alpha\beta} b_\nu b_\mu, \delta_{\alpha\beta} u_{\alpha,\nu} b_\mu \bar{b}_\mu\right) \\ \quad \text{for } \alpha, \alpha+2\beta \text{ long, } \beta, \alpha+\beta \text{ short} \end{cases}$$
$$(\nu, \mu \in \{1, 2, m+1, m+2\}),$$

$$\left[y_\alpha(0, b_\nu \xi), y_\beta(b_\mu)\right] = \begin{cases} 1 & \text{for } \alpha \text{ long, } \beta \text{ short, } \alpha + \beta \notin \Phi, \\ y_{\alpha+\beta}\left(\epsilon_{\alpha\beta} \bar{b}_\nu \bar{\xi} b_\mu\right) y_{\alpha+2\beta}\left(0, \delta_{\alpha\beta} b_\nu \xi b_\mu \bar{b}_\mu\right) & \\ & \text{for } \alpha, \alpha + 2\beta \text{ long, } \beta, \alpha + \beta \text{ short} \end{cases}$$

$$(\nu = 1, 2, 3, 4,\ \mu = 1, 2, m+1, m+2)$$

(as before, the right-hand sides of (B0) are expanded using (A0), (A0′), and (A0″) where necessary).

To define the elements $h_\alpha(t)$, for short roots α we can use (4) and (5) (see Section 5.1) as before, while for long roots α we use the variation

$$w_\alpha(t, u) = y_\alpha(t, u) y_{-\alpha}(-\bar{t}\bar{u}^{-1}, \bar{u}^{-1}) y_\alpha(tu^{-1}\bar{u}, u) \tag{11}$$

and

$$h_\alpha(t) = w_\alpha(v, t\xi) w_\alpha(0, \xi)^{-1}, \tag{12}$$

where $t, u \in GF(q^2)$ with $u + \bar{u} = \epsilon_\alpha t\bar{t}$ as usual, and, in (12), v can be taken as any element of $GF(q^2)$ for which $\epsilon_\alpha v\bar{v} = \xi(t - \bar{t})$ [Ste81, 5.3, 5.8, and 5.6]; in particular, $v = 0$ if $t \in GF(q)$.

Given this, the remaining relations we need are the following:

(H0) $$h_\alpha = h_\alpha(\theta\bar{\theta}) \qquad (\alpha \in \Psi);$$

(H0′) $$h'_\alpha = h_\alpha(\theta) \qquad (\alpha \in \Psi)$$

(where the right-hand sides are expanded as discussed above);

(H1) $$\left[h_\alpha, h_\beta\right] = 1 \qquad (\alpha, \beta \in \Psi);$$

(H2) $$h_\beta y_\alpha(b_\nu) h_\beta^{-1} = y_\alpha\left((\theta\bar{\theta})^{2(\alpha,\beta)/(\beta,\beta)} b_\nu\right)$$

$$(\alpha, \beta \in \Psi,\ \alpha \text{ short}, 1 \le \nu \le 2m),$$

$$h_\beta y_\alpha(0, b_\nu \xi) h_\beta^{-1} = y_\alpha\left(0, (\theta\bar{\theta})^{4(\alpha,\beta)/(\beta,\beta)} b_\nu \xi\right)$$

$$(\alpha, \beta \in \Psi,\ \alpha \text{ long}, 1 \le \nu \le m),$$

$$h_\beta y_\alpha(b_\nu, u_{\alpha,\nu}) h_\beta^{-1} = y_\alpha\left((\theta\bar{\theta})^{2(\alpha,\beta)/(\beta,\beta)} b_\nu, (\theta\bar{\theta})^{4(\alpha,\beta)/(\beta,\beta)} u_{\alpha,\nu}\right)$$

$$(\alpha, \beta \in \Psi,\ \alpha \text{ long}, 1 \le \nu \le 2m),$$

where the right-hand sides are expanded according to (A0), (A0′), and (A0″) when necessary, but as before (because of our choices of b_ν and $u_{\alpha,\nu}$) expansion is needed only for a bounded number of values of ν.

The fact that the relations (H2) hold in the twisted group G can be seen essentially as in Section 6.1. For α long, our $x_\alpha(t,u)$ corresponds to $x_S(t,u) = x_r(t)x_{\bar{r}}(\bar{t})x_{r+\bar{r}}(u)$ in Carter's notation [Car72, 13.6.4(iv)], so its conjugate by $h_\beta(w)$ has the form $x_\alpha(f(w)t, f'(w)u)$ for some functions f and f'. As before, for (H2) we only need these functions for $w \in GF(q)$, where they are determined inside a smaller untwisted group and $f(w)$ has the same form as before. Finally, note that $f'(w) = f(w)^2$ in this situation (for $w, f(w), f'(w) \in GF(q)$), because $u + \bar{u} = \epsilon_\alpha t\bar{t}$ and the same relation must hold between $f'(w)u$ and $f(w)t$.

As before, the above relations imply the rest of the Steinberg relations for the twisted group G (again using Lemma 5.2 to show that (H2) produces all other relevant values of ν, μ from the ones indicated in the relations (A1) through (B0), and then proving essentially as in Section 4.2 that, in each rank 2 subgroup, the Sylow p-subgroup has the appropriate order, hence all relations (B) follow).

Finally, observe that most of the above relations have length $O(\log|G|)$ as before, but each relation (A6) has length $O(m \log p)$ (and not necessarily bounded length) because of the expansion (A0′), and there are more than m relations (A6), leading to a factor of m^2 in the length of the presentation: the length does not appear to be $O(\log|G|)$. Nevertheless the above presentation certainly has length $O(\log^2|G|)$, and this remains true when the center is killed as before (exactly as in Example 5.3, using the generators h'_α).

6.4. *The Ree Groups* ${}^2F_4(q)$

Finally, the remaining twisted groups of rank ≥ 2 are the groups ${}^2F_4(q)$ of characteristic 2. Here the "twisted root system" is not one of the usual root systems, but has 16 "roots," corresponding to the 16 vertices $1, \dots, 16$ of a regular 16-gon. These alternate long and short around the 16-gon, with odd indices short and even ones long. Also $q = 2^{2e+1}$ for some integer e (we may assume $e \geq 1$ here, treating Tits' group ${}^2F_4(2)'$ as sporadic). Let ϕ be the field automorphism such that $2\phi^2 = 1$ (i.e., $t^\phi = t^{2^e}$ for $t \in GF(q)$).

The presentation from Griess [Gri73] for this group is the following: the generators are $x_\alpha(t)$ for each of the 16 roots α, where $t \in GF(q)$. The relations (A) are of two sorts, depending on whether α is short (odd numbered) or long (even numbered). If α is long, then the corresponding root subgroup is elementary abelian of order q, i.e., isomorphic to the additive group of $GF(q)$, so its relations are as before. If α is short, then the corresponding root subgroup $\langle x_\alpha(t) \mid t \in GF(q) \rangle$ is isomorphic to a Sylow 2-subgroup (of order q^2) of a Suzuki group $Sz(q) = {}^2B_2(q)$. Its center consists of the elements $x_\alpha(t)^2$, and is isomorphic to the additive group of $GF(q)$, as is the central quotient group. Only a basis of the field $GF(q)$ over $GF(2)$ is needed for a presentation of this 2-group. The relations (A) for short α are

$$\text{(A)}\qquad \begin{aligned} x_\alpha(t)x_\alpha(u) &= x_\alpha(t+u)x_\alpha(v)^2, \qquad \text{where } v^{2\phi+1} = tu^{2\phi},\\ [x_\alpha(t), x_\alpha(u)] &= x_\alpha(v)^2, \qquad \text{where } v^{2\phi+1} = tu^{2\phi} + t^{2\phi}u,\\ \left[x_\alpha(t), x_\alpha(u)^2\right] &= 1,\\ x_\alpha(t)^2 x_\alpha(u)^2 &= x_\alpha(t+u)^2. \end{aligned}$$

(Note that an element v for which $v^{2\phi+1}$ has a desired value can easily be computed using a primitive root of $GF(q)$.)

The relations of type (B) are the following, together with those obtained from them by transforming the subscripts using the Weyl group of order 16 generated by $\alpha \to \alpha + 2$ and $\alpha \to -\alpha$ (modulo 16).

$$\text{(B)}\qquad \begin{aligned} [x_1(t), x_2(u)] &= 1,\\ [x_1(t), x_3(u)] &= x_2(tu),\\ [x_1(t), x_4(u)] &= x_3(v)^2, \qquad \text{where } v^{2\phi+1} = tu,\\ [x_1(t), x_5(u)] &= 1,\\ [x_1(t), x_6(u)] &= x_3(v)^2 x_4(t^{2\phi}u) x_5(w)^2,\\ &\qquad \text{where } v^{2\phi+1} = t^{2\phi+1}u \text{ and } w^{2\phi+1} = tu^{2\phi},\\ [x_1(t), x_7(u)] &= x_2(t^{\phi+1}u) x_3(t^{2\phi}u) x_5(tu^{2\phi})^3 x_6(tu^{2\phi+1}),\\ [x_1(t), x_8(u)] &= x_2(t^{2\phi+2}u) x_3(t^{2\phi+1}u)^2 x_4(t^{4\phi+2}u^{2\phi+1})\\ &\qquad \times x_5(t^{2\phi+1}u^{2\phi})^3 x_6(t^{2\phi+2}u^{2\phi+1}) x_7(tu),\\ [x_2(t), x_4(u)] &= 1,\\ [x_2(t), x_6(u)] &= 1,\\ [x_2(t), x_8(u)] &= x_4(t^{2\phi}u) x_6(tu^{2\phi}). \end{aligned}$$

Using a suitable basis for $GF(q)$ over $GF(2)$, this can be reduced to a presentation of length $O(\log^2|G|)$ by the method in Section 6.3; we omit the details.

7. GROUPS OF RANK 1

The rank 1 universal Chevalley groups are the groups $A_1(q) = SL(2,q)$; their central quotients $PSL(2,q)$ are the corresponding simple (if $q > 3$) Chevalley groups. Presentations for the groups $G = PSL(2,q)$ were found by Todd [Tod36] using $m + 2$ generators (recall that $q = p^m$); the length of his presentation is $\Theta(m \log p + m^2) \leq O(\log^2|G|)$. Assuming $(*)$, Todd's presentations are explicit (see below), proving Conjecture 3 for this case, thereby completing the proof of Theorem 1.6. ▌

Remark 7.1. It is easy to modify any presentation of $PSL(2,q)$ to one of $SL(2,q)$ at a cost of at most doubling the length.

We quote *Todd's presentation* of $PSL(2,p^m)$ for the case of*odd p*. Let θ be a primitive root of $GF(p^m)$; let $\theta^m = \sum_{i=0}^{m-1} a_i\theta^i$ be the irreducible polynomial over $GF(p)$ satisfied by θ $(0 \leq a_i \leq p-1)$. Then $\theta^{m+1} = \sum_{i=0}^{m-1} b_i\theta^i$, where $b_0 = a_{m-1}a_0$ and $b_i = a_{m-1}a_i + a_{i-1}$ $(i = 1,\dots,m-1)$.

Generators (Todd's notation): $U, R, S_i \quad (i=0,\dots,m-1)$.

Relations:
$$R^{(p^m-1)/2} = U^3 = (UR)^2 = (US_0)^2 = S_i^p = 1 \qquad (i = 0,\dots,m-1),$$
$$S_iS_j = S_jS_i \qquad (i,j = 0,\dots,m-1),$$
$$RS_i = S_{i+2}R \qquad (i = 0,\dots,m-3),$$
$$(S_1RU)^3 = 1,$$
$$RS_{m-2} = S_0^{a_0}S_1^{a_1}\cdots S_{m-1}^{a_{m-1}}R,$$
$$RS_{m-1} = S_0^{b_0}S_1^{b_1}\cdots S_{m-1}^{b_{m-1}}R.$$

When $p = 2$, Todd's presentation is of similar nature. A shorter presentation for this case was found by Sinkov [Sin39] (who also mentioned that something similar could be done to shorten Todd's presentation for odd p). *Sinkov's presentation* for $G = SL(2,2^m) = PSL(2,2^m)$ is the following:

Generators: x, y, z

Relations:
$$x^{2^m-1} = y^2 = z^3 = (xz)^2 = (yz)^2 = [x^i, y]^2 = 1 \qquad (i = 1,\dots,m-1),$$
$$x^m = yxy^{a_{m-1}}x \cdots y^{a_1}xy^{a_0},$$

where the a_i are defined as above ($a_i \in GF(2)$). The length of this presentation is $\Theta(m \log m) = \Theta(\log|G|\log\log|G|) \leq O(\log^2|G|)$. (Furthermore, this presentation can be shortened to length $\Theta(m) = \Theta(\log|G|)$, by adding generators x_i and relations $x_i = xx_{i-1}$ so we can use x_i in place of

x^i in the above commutator relations, giving each commutator relation bounded length.)

Remark 7.2. We remark that the case $q = p$ has attracted a lot of attention; a number of presentations of length $O(\log|G|)$ are listed in [CM72, Sect. 7.5]. Refining a presentation of Behr and Mennicke, Sunday [Sun72] found the following simple presentation (p is an odd prime):

$$PSL(2, p) = \left\langle x, y : x^p = y^2 = (xy)^3 = \left(x^4 y x^{(p+1)/2} y\right)^2 = 1 \right\rangle.$$

Remark 7.3. More recently Campbell, Robertson, and Williams [CRW90] have found shorter presentations than Todd's and Sinkov's for $PSL(2, p^m)$. Their presentations can be written so as to have length $\Theta(\log|G|)$ for all p and m (without needing to add generators as in Sinkov's presentation mentioned above).

Remark 7.4. We remark that Steinberg [Ste81] has also found presentations for the Lie type groups of rank 1 (including the twisted ones), but those presentations are of a different form than in the rank ≥ 2 case discussed in the previous sections (the commutator relations (B) do not arise since there is only one positive root). His presentations have the following general form.

Let U be a maximal unipotent subgroup (i.e., a Sylow p-subgroup, which is just a root subgroup since the rank is 1), let H be a Cartan subgroup so that UH is a Borel subgroup, let r be an involution conjugating U to the negative root subgroup, and let $N = H\langle r \rangle$. Then $H = \langle h \rangle$ is cyclic (isomorphic to a subgroup of known index 1, 2, or 3 in the multiplicative group of the field $GF(q)$, or $GF(q^2)$ in the unitary case) and $N = H\langle r \rangle$ has a very simple presentation (it is dihedral except in the unitary case when it has a presentation of the form $\langle h, r | h^k = 1, r^2 = 1, r^{-1}hr = h^{-q} \rangle$). Then a presentation for G can be obtained (cf. [Ste81, Sect. 4]) by starting with presentations for U and N, then giving the action of h on U, and finally giving *all* relations of the form $w = uvu'$ with $w \in \langle h \rangle r$, $u, u' \in U$, and $v \in r^{-1}Ur$ (there are $|U| - 1$ such relations, exactly one for each nontrivial $u \in U$ [Ste81, Lemma 4.3]).

In the case of $A_1(q) = PSL(2, q)$ these relations can be shortened to length $O(\log^2|G|)$ by building in conjugation by h: there are at most two orbits of $\langle h \rangle$ on the nontrivial elements of U, so all $q - 1$ of the relations $w = uvu'$ can be deduced from at most two of them, by conjugating by h. It is easy to see that this leads to a short presentation very similar to Todd's above for $PSL(2, q)$.

However, in the case of the twisted rank 1 groups (${}^2A_2(q)$, ${}^2B_2(q)$, and ${}^2G_2(q)$), $\langle h \rangle$ has at least q orbits on U, and it is not clear how to deduce all those relations from a bounded number of them. Thus, Conjecture 1 remains open for these three types of simple groups.

8. SHORT PRESENTATIONS FOR ALL FINITE GROUPS

In this section, we prove Theorem 1.4. It will be important to reach rapidly any element of a group from any set of generators. This is always possible by the next lemma.

Let G be a group, and A and T two subsets of G. A *straight line program*, computing A from T, is a sequence $S = (w_1, \dots, w_s)$ of elements of G such that $A \subset S$ (we use the same symbol S to denote the set $\{w_1, \dots, w_s\}$), and each member w_i of S is either a member of T, or a product $w_j w_k$ for some $j, k < i$, or w_j^{-1} for some $j < i$. The *length* of S is s. If t of the elements in the sequence S belong to T, then the *reduced length* of S is $s - t$ (we do not count accesses to the generators).

The *cost* of A relative to T, $\operatorname{cost}(A|T)$, is the shortest reduced length of straight line programs computing A from T. (For a single element $A = \{g\}$ we abbreviate $\operatorname{cost}(\{g\}|T)$ as $\operatorname{cost}(g|T)$. If T does not generate A, the cost is ∞.) We note that for any sets $A, B, C \subseteq G$ we have

$$\operatorname{cost}(A|C) \leq \operatorname{cost}(A|B) + \operatorname{cost}(B|C).$$

The following lemma appears in [BS84].

LEMMA 8.1 (Reachability Lemma). *Let G be a finite group, T a set of generators, and $g \in G$. Then* $\operatorname{cost}(g|T) < (1 + \log|G|)^2$.

In fact, the proof given in [BS84] yields the following stronger version.

LEMMA 8.2. *Let G be a finite group and T a set of generators. Then there exists a set $A \subset G$ such that*:

(i) $\operatorname{cost}(A|T) < \log^2|G|$; *and*

(ii) *for any $g \in G$*, $\operatorname{cost}(g|A) < 2\log|G|$.

Thus, after a preprocessing, which costs $< \log^2|G|$, every element can be reached at only logarithmic cost. Let $\operatorname{setup}(G)$ denote the maximum, over all sets T of generators, of the cost of this preprocessing. We thus have

$$\operatorname{setup}(G) < \log^2|G|. \tag{13}$$

Let further $\operatorname{reach}(G)$ be the maximum, over $g \in G$, of $\operatorname{cost}(g|A)$ after proper preprocessing. Then

$$\operatorname{reach}(G) < 2\log|G|. \tag{14}$$

We prove the following, more specific, form of Theorem 1.4.

THEOREM 8.3. *If each composition factor H_i of the finite group G has a presentation of length $O(\log^C|H_i|)$ for some $C \geq 2$, then G has a presentation of length $O(\log^{C+1}|G|)$.*

Proof. We shall use the terms "short" and "efficient" in a formal sense. A set A of generators is *efficient* if it satisfies the bound (14). A straight line program which computes an efficient set of generators is *short* if it satisfies the bound (13). A straight line program computing an element from an efficient set of generators is *short* if it satisfies the bound (14).

Let $G = G_0 \triangleright G_1 \triangleright \cdots \triangleright G_m = 1$ be a composition series in G (note that $m \leq \log|G|$) and let H_i denote the (simple) factor group G_{i-1}/G_i. Let $\mathrm{Gen}(H_i)$ and $\mathrm{Rel}(H_i)$ denote the sets of generators and relations, respectively, of a presentation of H_i having length $O(\log^C|H_i|)$.

Let $\mathrm{Gen}_0(H_i) \subseteq \mathrm{Gen}(H_i)$ denote a subset of cardinality $|\mathrm{Gen}_0(H_i)| \leq \log|H_i|$ such that $\mathrm{Gen}_0(H_i)$ generates H_i.

Let L_i be a lifting of $\mathrm{Gen}(H_i)$ to G_{i-1}, with $L_{0,i}$ denoting its subset which lifts $\mathrm{Gen}_0(H_i)$. Each relation $\rho = \rho(u_1, \dots) = 1$ from $\mathrm{Rel}(H_i)$ is thus lifted to a relation $\rho(w_1, \dots) = z_\rho$ for some $z_\rho \in G_i$ (where $w_1 \in L_i$ is the lifting of $u_1 \in \mathrm{Gen}(H_i)$, etc.).

Let $M_i = L_{0,i+1} \cup \cdots \cup L_{0,m}$. Clearly, M_i generates G_i.

Let S_i be a short straight line program in G_i computing an efficient set of generators of G_i from M_i. We may assume that S_i includes M_i. Set $S = \bigcup_i S_i$.

For each $w \in L_i$ let $D(w) = (\dots, w)$ be a short straight line program computing w from S_{i-1}. Let $D_i = \bigcup\{D(w) : w \in L_i\}$ and $D = \bigcup_i D_i$.

For each $\rho \in \mathrm{Rel}(H_i)$, let $P(\rho) = (\dots, z_\rho)$ be a short straight line program computing z_ρ from S_i. Let $P_i = \bigcup\{P(\rho) : \rho \in \mathrm{Rel}(H_i)\}$, and $P = \bigcup_i P_i$.

Let $w \in L_{0,i}$ and $z \in L_{0,j}$ for some $j > i$. Since G_i is normal in G_{i-1}, the conjugate $w^{-1}zw$ belongs to G_i. Let $Q(w,z) = (\dots, w^{-1}zw)$ be a short straight line program computing $w^{-1}zw$ from S_i. Let $Q_{i,j} = \bigcup\{Q(w,z) : w \in L_i \text{ and } z \in L_j\}$ and $Q = \bigcup_{i<j} Q_{i,j}$.

Finally, let $T = S \cup D \cup P \cup Q$. Of course, S itself generates G, but it is this highly redundant set of generators that will yield our short presentation of G.

We associate a symbol $x(t)$ with each $t \in T$; these symbols will be the generators in our presentation of G. There will be three kinds of relations in the presentation:

R1. If some $t \in T$ arose as $t = vw$ or $t = w^{-1}$ in the course of one of the straight line programs S_i, $D(w)$, $P(\rho)$, or $Q(w,z)$ referred to above, we include the relation $x(t) = x(v)x(w)$ or $x(t)x(w) = 1$, respectively.

R2. For each $\rho \in \mathrm{Rel}(H_i)$, lifted to $\rho(w_1, \dots) = z_\rho$ as above, we include the relation $\rho(x(w_1), \dots) = x(z_\rho)$.

R3. For $w \in L_{0,i}$ and $z \in L_{0,j}$, $i < j$, we include the relation $x(w^{-1}zw) = x(w)^{-1}x(z)x(w)$.

Let $\hat{G}$ be the group defined by these relations. We clearly have a homomorphism ϕ: $\hat{G} \to G$ onto G mapping $x(t)$ to $t \in T$. We claim *ϕ is one-to-one*. It suffices to show that $|\hat{G}| \le |G|$.

Let $\hat{G}_i$ be the subgroup of $\hat{G}$ generated by the set $\{x(t) : t \in M_i\}$ [identifying these symbols $x(t)$ with the corresponding elements in $\hat{G}$]. We first observe that $\hat{G}_0 = \hat{G}$ because the relations R1 make all the other generators redundant. Second, $\hat{G}_i$ is normal in $\hat{G}_{i-1}$ because of the relations R3. Finally, the factor group $\hat{G}_{i-1}/\hat{G}_i$ is generated by its elements $\{x(t)\hat{G}_i : t \in L_i\}$ (since L_i contains $L_{0,i}$), and these generators satisfy the relations corresponding to $\mathrm{Rel}(H_i)$. Therefore $|\hat{G}_{i-1}/\hat{G}_i| \le |H_i|$ and thus $|\hat{G}| \le |G|$.

It remains to estimate the length of this presentation.

R1 falls into four classes, corresponding to the straight line programs comprising S, D, P, and Q, resp. Ignoring a possible doubling of the length when a straight line program is converted into relations, the respective lengths are bounded by the quantities

$$C(S) \le \sum_i \mathrm{setup}(G_i) \le m \cdot \mathrm{setup}(\max),$$

$$C(D) \le \sum_i |\mathrm{Gen}(H_i)| \mathrm{reach}(G_i) \le \mathrm{reach}(\max) \sum_i |\mathrm{Gen}(H_i)|,$$

$$C(P) \le \sum_i |\mathrm{Rel}(H_i)| \mathrm{reach}(G_i) \le \mathrm{reach}(\max) \sum_i |\mathrm{Rel}(H_i)|,$$

$$C(Q) \le \sum_{i<j} \log|H_i| \log|H_j| \mathrm{reach}(G_i) \le \log^2|G| \mathrm{reach}(\max),$$

where $\mathrm{setup}(\max) = \max_i \mathrm{setup}(G_i)$, and $\mathrm{reach}(\max)$ is defined analogously.

The contribution of R2 is the sum over all i of the total lengths of the relations in $\mathrm{Rel}(H_i)$.

The contribution of R3 is $O(\log^2|G|)$, negligible compared to our bound on $C(Q)$.

Now observe that $|\mathrm{Gen}(H_i)|$ and $|\mathrm{Rel}(H_i)|$ are bounded by the total length of the relations in $\mathrm{Rel}(H_i)$, which is $O(\log^C|H_i|)$ by hypothesis. Using that bound, together with the bounds (13) and (14) on the "setup" and "reach" functions, and the inequalities $m \le \log|G|$ and $2 \le C$, we conclude that G has a presentation of length

$$O\left(m \log^2|G| + \log|G| \left(\sum_i \log^C|H_i| \right) + \log^3|G| \right) \le O\left(\log^{C+1}|G|\right). \quad \blacksquare$$

In particular, we have

COROLLARY 8.4. *If the finite group G has no composition factor which is a rank* 1 *twisted group of Lie type, then G admits a presentation of length* $O(\log^3|G|)$.

COROLLARY 8.5. *Every solvable group G has a presentation of length* $O(\log^3|G|)$.

We note that the *exponent* 3 *here is best possible,* even for nilpotent groups:

PROPOSITION 8.6. *Let* $\operatorname{minpres}(G)$ *denote the length of the shortest presentation of the group G. For a positive integer N, let* $\operatorname{minpres}(N) = \max\{\operatorname{minpres}(G) : |G| = N\}$. *Then, for $N = p^n$ a prime power,*

$$\operatorname{minpres}(N) > \left(c/\log^2 p\right)\left(\log^3 N/\log\log N\right),$$

where c is a positive absolute constant.

Proof. Higman [Hig60] gave the following lower bound for the number $A(p^n)$ of pairwise nonisomorphic groups of order p^n:

$$A(p^n) > p^{(2/27-\epsilon_n)n^3}, \tag{15}$$

where ϵ_n depends on n but not on p and $\lim_{n\to\infty}\epsilon_n = 0$. On the other hand, the number of groups admitting presentations of length k (including parentheses and other special characters) is $< k^k$. Thus, for $k = 4 \cdot \operatorname{minpres}(p^n)$ we have

$$k^k > A(p^n). \tag{16}$$

A comparison of (15) and (16) proves Proposition 8.6. ▮

The possibility of improving the reachability lemma remains open.

Problem 8.7. Is it possible to improve the $O(\log^2|G|)$ bound of the reachability lemma for solvable groups G?

ACKNOWLEDGMENTS

We are indebted to J. H. Conway and A. Lubotzky for helpful suggestions.

REFERENCES

[AL86] L. M. Adleman and H. W. Lenstra, Finding irreducible polynomials over finite fields, *in* “Proceedings of the ACM Symposium on Theory of Computing, Berkeley, 1986,” pp. 350–355.

[Ba91] L. Babai, Computational complexity in finite groups, *in* ''Proceedings of the International Congress of Mathematicians, Kyoto, 1990,'' pp. 1479–1489, Springer-Verlag, Tokyo, 1991.

[Ba92] L. Babai, Bounded round interactive proofs in finite groups, *SIAM J. Discrete Math.* **5** (1992), 88–111.

[Ba97] L. Babai, Randomization in group algorithms: conceptual questions, *in* ''Groups and Computation II'' (L. Finkelstein and W. M. Kantor, eds.), DIMACS Ser. *Discrete Math. Theor. Comp. Sci.*, **28** (1997), 1–16.

[BS84] L. Babai and E. Szemerédi, On the complexity of matrix group problems I, *in* ''Proceedings of the 25th IEEE Symposium on Foundations of Computer Science, Palm Beach, FL, 1984,'' pp. 229–240.

[Ber70] E. R. Berlekamp, Factoring polynomials over large finite fields, *Math. Comp.* **24** (1970), 713–735.

[CRW90] C. M. Campbell, E. F. Robertson, and P. D. Williams, On presentations of $PSL(2, p^n)$, *J. Austral. Math. Soc. Ser. A* **48** (1990), 333–346.

[CZ81] D. G. Cantor and H. Zassenhaus, A new algorithm for factoring polynomials over finite fields, *Math. Comp.* **26** (1981), 587–592.

[Car23] R. D. Carmichael, Abstract definitions of the symmetric and alternating groups and certain other permutation groups, *Quart. J. Math.* **49** (1923), 226–270.

[Car72] R. Carter, ''Simple Groups of Lie Type,'' Wiley, London, 1972.

[CM72] H. S. M. Coxeter and W. O. J. Moser, ''Generators and Relations for Discrete Groups,'' 3rd ed., Springer-Verlag, Berlin/New York, 1972.

[Cur65] C. W. Curtis, Central extensions of groups of Lie type, *J. Reine Angew. Math.* **220** (1965), 174–185.

[GJ79] M. Garey and D. S. Johnson, ''Computers and Intractability: A Guide to the Theory of NP-Completeness,'' Freeman, New York, 1979.

[Gor80] D. Gorenstein, ''Finite Groups,'' 2nd ed., Chelsea, New York, 1980.

[Gri73] R. L. Griess, Jr., Schur multipliers of finite simple groups of Lie type, *Trans. Amer. Math. Soc.* **183** (1973), 355–421.

[Hig60] G. Higman, Enumerating p-groups I, *Proc. London Math. Soc.* **10** (1960), 24–30.

[Hum72] J. E. Humphreys, ''Introduction to Lie Algebras and Representation Theory,'' Springer-Verlag, New York, 1972.

[KL90] P. Kleidman and M. Liebeck, ''The Subgroup Structure of the Finite Classical Groups,'' London Math. Soc. Lecture Note Series, Vol. 129, Cambridge Univ. Press, 1990.

[Rab80] M. O. Rabin, Probabilistic algorithms in finite fields, *SIAM J. Comput.* **9** (1980), 273–280.

[Sin39] A. Sinkov, A note on a paper by J. A. Todd, *Bull. Amer. Math. Soc.* **45** (1939), 762–765.

[Ste62] R. Steinberg, Génerateurs, relations et revêtements de groupes algébriques, *in* ''Colloque sur la theorie des groupes algébriques, Bruxelles, 1962,'' pp. 113–127.

[Ste67] R. Steinberg, ''Lectures on Chevalley groups'' (mimeographed notes), Yale Univ., 1967.

[Ste81] R. Steinberg, Generators, relations and coverings of algebraic groups, II, *J. Algebra* **71** (1981), 527–543.

[Sun72] J. G. Sunday, Presentations of the groups $SL(2, m)$ and $PSL(2, m)$, *Canad. J. Math.* **24** (1972), 1129–1131.

[Tod36] J. A. Todd, A second note on the linear fractional group, *J. London Math. Soc.* **11** (1936), 103–107.