J. Eur. Math. Soc. 13, 391–458

DOI 10.4171/JEMS/257

R. M. Guralnick · W. M. Kantor · M. Kassabov · A. Lubotzky

Presentations of finite simple groups: a computational approach

Received November 17, 2007 and in revised form December 17, 2009

Abstract. All finite simple groups of Lie type of rank n over a field of size q, with the possible exception of the Ree groups ${}^2G_2(q)$, have presentations with at most 49 relations and bit-length $O(\log n + \log q)$. Moreover, A_n and S_n have presentations with 3 generators, 7 relations and bit-length $O(\log n)$, while $\mathrm{SL}(n, q)$ has a presentation with 6 generators, 25 relations and bit-length $O(\log n + \log q)$.

Contents

R. M. Guralnick: Department of Mathematics, University of Southern California, Los Angeles, CA 90089-2532, USA; e-mail: guralnic@usc.edu

W. M. Kantor: Department of Mathematics, University of Oregon, Eugene, OR 97403, USA; e-mail: kantor@math.uoregon.edu

M. Kassabov: Department of Mathematics, Cornell University, Ithaca, NY 14853-4201, USA; e-mail: kassabov@math.cornell.edu

A. Lubotzky: Department of Mathematics, Hebrew University, Givat Ram, Jerusalem 91904, Israel; e-mail: alexlub@math.huji.ac.il

Mathematics Subject Classification (2010): Primary 20D06, 20F05; Secondary 20J06

1. Introduction

In [GKKL1] we provided short presentations for all alternating groups, and all finite simple groups of Lie type other than the Ree groups ${}^2G_2(3^{2e+1})$, using at most 1000 generators and relations. In [GKKL2] we proved the existence of *profinite presentations* for all finite quasisimple groups using fewer than 20 relations. (Recall that a nontrivial group is called *quasisimple* if it is both perfect and simple modulo its center. These are precisely the perfect central extensions of nonabelian simple groups.) The goal of the present paper is similar: we provide several presentations for the same simple groups, including ones using 2 generators and at most 51 relations. These have the potential advantage that they are simpler than those in [GKKL1], at least in the sense of requiring fewer relations; we hope that both types of presentations will turn out to be useful, for example in Computational Group Theory.

The fundamental difference between this paper and [GKKL1] is that here we achieve a smaller number of relations at the cost of relinquishing some control over the length of the presentations, although our first result does not deal with length at all:

Theorem A. *All finite quasisimple groups of Lie type, with the possible exception of the Ree groups* ${}^2G_2(3^{2e+1})$, *have presentations with* 2 *generators and* 51 *relations. All symmetric and alternating groups have presentations with* 2 *generators and* 8 *relations.*

In fact, a similar result holds for *all* finite simple groups, except perhaps ${}^2G_2(q)$ (the sporadic groups are surveyed in [Soi]). There is also a version for almost quasisimple groups (Corollary 9.4). Both the bounds of 20 profinite relations in [GKKL2] and 51 relations here are not optimal (cf. Remark 4 in Section 11). Possibly 4 is the correct upper bound for both standard and profinite presentations. Wilson [Wi] has even conjectured that 2 relations suffice for the universal covers of all finite simple groups.

Although we are giving up the requirement of small word length used in [GKKL1], we can still retain control over *bit-length*, a weaker and less familiar notion of length introduced in [BS] for Computer Science complexity considerations and used in [BGKLP, BCLO]. This is the total number of bits required to write the presentation, where all

Table 1. Summary

Group	gens	rels	Group	gens	rels
A_n, S_n	3	7	$\mathrm{SL}(n,q)$	6	25
$\mathrm{SL}(2,q)$	3	9	$\mathrm{SU}(n,q)$	5	32
$\mathrm{SU}(3,q)$	3	21	$\mathrm{Sp}(2n,q)$, q odd	8	47
$\mathrm{Sz}(q)$	4	29	$\mathrm{Spin}_{2n+1}(q)$	9	48
$\mathrm{SL}(3,q)$	4	14	$\mathrm{Spin}^+_{2n}(q)$	9	42
$\mathrm{Sp}(4,q)$	$4+(2,q)$	27	$\mathrm{Spin}^-_{2n}(q)$	9	48
$\mathrm{SU}(4,q)$	5	27	$\hat{E}_n(q)$	6	30
$G_2(q)$, ${}^3D_4(q)$	6	31	$F_4(q)$	8	46
${}^2F_4(q)$	6	49	${}^2\hat{E}_6(q)$	8	46

exponents are encoded as binary strings, the sum of whose lengths enters into the bit-length. (The length that had to be kept small in [GKKL1] involves the much larger sum of the actual exponents; cf. Section 2. Thus, "short" length implies "short" bit-length.)

Theorem B. *All finite quasisimple groups of Lie type of rank n over a field of size q, with the possible exception of the Ree groups ${}^2G_2(q)$, have presentations with at most* 9 *generators and* 49 *relations, and bit-length $O(\log n + \log q)$.*[1]

As in [GKKL1], we view A_n (and S_n) as groups of Lie type of rank $n-1$ over "the field of order 1" [Ti1]; as in [GKKL1], Theorem B seems counterintuitive, in view of the standard types of known presentations for simple groups (such as [St2]); and as in [GKKL1], the $O(\log n + \log q)$ bound on bit-length is optimal in terms of n and q (due to the number of bits required to write both n and q).

By [GKKL1, Lemma 2.1] (cf. Lemma 2.3 below), if we have any presentation of a finite simple group G with at most $r \leq 49$ relations, we obtain a presentation with 2 generators and at most $r + 2 \leq 51$ relations. Thus, Theorem A is an immediate consequence of Theorem B (using [St1]). Moreover, the proof of those lemmas shows that *any* pair of generators of G can be used for such a presentation (cf. Remark 4 in Section 11). Indeed, those lemmas are so general that they allow us to cheat somewhat: the resulting presentations are not even slightly explicit, and we have no information concerning their bit-lengths. "Almost all" pairs of elements of a finite simple group generate the group [Di, KaLu, LiSh]. Some pairs force the length to be large (cf. the Appendix); we do not know whether there are pairs for which the corresponding presentations have bit-length $O(\log n + \log q)$.

In view of the preceding paragraph, our goal will be to prove Theorem B. Much better bounds are obtained in various cases, some of which are summarized in Table 1. If G is a group of Lie type then $\hat{G}$ will denote the corresponding simply connected group; in

[1] Logarithms will be to the base 2.

general, the simple quotient group is obtained by adding at most one further relation, but two are needed for some orthogonal groups.

We highlight one entry of this table. For alternating and symmetric groups we use classical permutation group ideas that would have been familiar in the 19th century in order to go further than previous types of presentations [GKKL1, BCLO] in terms of the small number of relations:

Theorem C. *For each $n \geq 5$, A_n and S_n have presentations with* 3 *generators*, 7 *relations and bit-length $O(\log n)$, using a bounded number of exponents each of which is at most n.*

This also yields a presentation of S_n having 5 generators, 9 relations and bit-length $O(\log n)$, with two of the generators mapping onto $(1,2)$ and $(1,\dots,n)$ (using Remark 3.37). Moreover, in addition to the second part of Theorem A, as above Lemma 2.3 implies that *if a and b are* any *generators of $G = A_n$ or S_n, then there is a presentation of G using* 2 *generators that map onto a and b, with* 9 *relations*. However, as already noted, we do not know if it is possible to choose a and b in order to obtain a presentation with bit-length $O(\log n)$. (See the Appendix for related results.) In [GKKL1] we gave a presentation of S_n having 58 relations and length (not just bit-length) $O(\log n)$. Different bounded presentations of S_n in [BCLO] have bit-length $O(\log n)$ and, for all practical purposes, at most 28 relations.

In order to obtain all of the presentations in the preceding theorems, although [GKKL1] was a starting point we need significant new methods for unbounded rank n; these ideas may prove to be more practical for actual group computation than some of those in [GKKL1]. Moreover, while a few of the arguments used here are streamlined, often simpler, and occasionally improved versions of ones in [GKKL1], they are still rather involved. As in [GKKL1] *we do not use the classification of the finite simple groups in the proofs of the above theorems*. This classification is needed in the proof of Corollary D, though only for the list of sporadic simple groups and not for any of the groups in Theorem B.

For groups of bounded rank, our presentations can be made short in the sense used in [GKKL1], at the cost of adding a small number of additional generators and relations (so that [GKKL1, (3.3) and (4.16)] will apply; cf. (3.17) and (4.3) below). It is our treatment of unbounded rank that contains new ideas to decrease the number of relations in [GKKL1] at the expense of the length of the presentation. We provide more than one approach for this purpose: some classical groups are handled in more than one way in Sections 9 and 10. The unitary groups are dealt with separately in Section 8 using an idea of Phan [Ph] as improved in [BeS].

In Sections 3–10 we consider various types of simple groups in order to prove the above theorems. Most of our results are summarized in Table 1 and Theorem 9.1. Some of these required computer assistance with small groups [Br, Hav, CHRR2] in order to obtain the relatively small numbers of generators and relations in a few cases. In many cases we only give hints regarding the bit-length assertion in Theorem B.

One of our original motivations for work on presentations was the following

Corollary D (Holt's Conjecture [Ho] for simple groups). *There is a constant C such that* $\dim H^2(G, M) \leq C \dim M$ *for every finite simple group G, every prime p and every irreducible $\mathbb{F}_p[G]$-module M.*

This conjecture has already been proven twice, in [GKKL1, Theorem B′] and [GKKL2, Theorem B]. As in [GKKL1], the conjecture is an immediate consequence of Theorem B (using the elementary result [GKKL1, Lemma 7.1]), except for the Ree groups ${}^2G_2(q)$—which also had to be handled separately in [GKKL1]. The proof based on the present paper is simpler than the previous proofs, since Theorem B takes less effort than before. On the other hand, [GKKL2] obtains the constant $C = 18$, and shows that this is virtually equivalent to the statement that all finite simple groups have profinite presentations with 2 generators and 18 relations. Also [GKKL2, Theorem C] is a generalization of the preceding Corollary to all finite groups (where C becomes 19).

Other consequences of some of our results are proved in Theorems 3.49 and 3.50:

Corollary E. (i) *The automorphism group of the free group F_n, $n > 2$, has a presentation with* 5 *generators and* 18 *relations.*

(ii) *If $n \geq 6$ then* $\mathrm{SL}(n, \mathbb{Z})$ *has a presentation using* 4 *generators and* 16 *relations.*

Section 11 contains additional remarks concerning our results. For now we note one further unexpected direction:

Efficient presentations. If $\langle X \mid R\rangle$ is a presentation of a finite group G, then $|R| - |X|$ is at least the smallest number $d(M)$ of generators of the Schur multiplier M of G; and $\langle X \mid R\rangle$ is called an *efficient* presentation if $|R| - |X| = d(M)$ [CRKMW, CHRR1, CHRR2]. For 35 years the groups $\mathrm{PSL}(2, p)$ with $p \geq 5$ prime have contained the only infinite family of finite nonabelian simple groups known to have efficient presentations ([Sun, CR3]; cf. (3.19)). In fact, $\mathrm{SL}(2, \mathbb{Z}_m)$ is efficient for any odd integer $m > 1$ [CRW1, p. 70], as is any quotient by a subgroup of its center (compare [CR2, p. 19]).

On the other hand, [Har, Cor. 5.4] states that, if p is any prime not dividing the order of a finite group G, then $G \times P$ is efficient for all sufficiently large elementary abelian p-groups P. (In particular, every perfect finite group is the derived group of an efficient finite group.) These groups have very large numbers of generators and much larger numbers of relations. Therefore, it may be of some interest that Corollaries 3.8(i) and 3.13(ii) contain examples of families of groups having efficient presentations with alternating groups as composition factors and small numbers of relations. For example, we show that, *for any prime $p \equiv 11 \pmod{12}$, there is an efficient presentation of $A_{p+2} \times T$ with* 2 *generators and* 3 *relations, where T is the subgroup of index* 2 *in* $\mathrm{AGL}(1, p)$.

Table 2 in Section 3.2 contains presentations for various groups A_n and S_n when n has a special form, including ones that use fewer relations than any others presently in print. For some alternating or symmetric groups Sections 3.1 and 3.2 contain more than one presentation. Section 3.5 contains explicit presentations of S_n for all $n \geq 50$. For general n it would be desirable to have even fewer relations than in Theorem C, with the goal of approaching efficiency for alternating groups (compare [GKKL3], where we use profinite presentations for a similar purpose).

While [GKKL1] concerned (word) lengths, the present paper concerns bit-lengths. In the Appendix we conclude with simple lower bounds for lengths (but not bit-lengths) of presentations.

2. Preliminaries

Presentation lengths. In [GKKL1, Section 1.2] there is a long discussion of various notions of "lengths" of a presentation $\langle X \mid R\rangle$ and some of the relationships among them. Here we only summarize what is needed for the present purposes.

- *length* = *word length*: $|X| +$ sum of the lengths of the words in R within the free group F_X on X. Thus, length refers to strings in the alphabet $X \cup X^{-1}$. This is the notion of length used in [GKKL1], and seems the most natural notion from a purely mathematical point of view. We reserve the term *short presentation* for one having "small" length. Achieving this was one of the goals in [GKKL1], though not of the present paper.
- *bit-length*: the total number of bits required to write the presentation, used in [BGKLP] and [BCLO]. All exponents are encoded as binary strings, the sum of whose lengths enters into the bit-length.
- *expo-length*: the total number of exponents used in the presentation.

By comparing the present paper with [GKKL1] it becomes clear that small bit-length is much easier to achieve than small length. The properties required of the bit-length $bl(w)$ of a word w are as follows:

$$bl(x) = 1 \quad \text{if } x^{\pm 1} \in X; \qquad bl(w^n) \le bl(w) + \log|n| \quad \text{if } n \in \mathbb{Z}\backslash\{0, 1, -1\};$$
$$bl(ww') \le bl(w) + bl(w') \qquad \text{for any words } w, w'.$$

Notation: Functions will always act on the left, and we use the notation $g^h = h^{-1}gh$ and $[g, h] = g^{-1}h^{-1}gh = g^{-1}g^h$.

Presentations. We will use the elementary fact [GKKL1, Lemma 2.3] that a group J that has a presentation based on a known group, using presentations of subgroups of that group, has those subgroups automatically built into J:

Lemma 2.1. *Let* $\pi : F_{X\cup Y} \to J = \langle X, Y \mid R, S\rangle$ *and* $\lambda \colon F_X \to H = \langle X \mid R\rangle$ *be the natural surjections, where* H *is finite. Assume that* $\alpha \colon J \to J_0$ *is a homomorphism such that* $\alpha\langle\pi(X)\rangle \cong H$. *Then* $\langle\pi(X)\rangle \cong H$.

In the present paper we will use this freely, often without comments.

We will need the following elementary observation in order to handle some central extensions. This was stated in [GKKL1, Proposition 2.4] for length, but the proof applies just as well to bit-length.

Lemma 2.2. *Suppose that* G *has a presentation* $\langle X \mid R\rangle$ *in which* R *has total bit-length* L. *If* $\hat{G}$ *is a perfect group such that* $\hat{G}/Z = G$ *with* $Z \le Z(\hat{G})$ *of prime order* p, *then* $\hat{G}$ *has a presentation* $\langle \hat{X} \mid \hat{R}\rangle$ *such that* $|\hat{X}| = |X| + 1$, $|\hat{R}| = |X| + |R| + 1$, *the bit-length of* $\hat{R}$ *is less than* $4|X| + 2L + p|R|$, *and* $\hat{X}$ *contains a generator of* Z.

We also note a straightforward improvement of [GKKL1, Lemma 2.1]:

Lemma 2.3. *Let $G = \langle D \rangle$ be a finite group having a presentation $\langle X \mid R \rangle$; let $\pi \colon F_X \to G$ be the natural map. Then G also has a presentation $\langle D \mid R' \rangle$ such that $|R'| = |D| + |R| - |\pi(X) \cap D|$.*

Proof. We recall the simple idea used in the proof of [GKKL1, Lemma 2.1]. Write each $x \in X$ as a word $v_x(D)$ in D, and each $d \in D$ as a word $w_d(X)$ in X; and let $V(D) = \{v_x(D) \mid x \in X\}$. According to the proof of [GKKL1, Lemma 2.1], we then obtain another presentation for G:

$$G = \langle D \mid d = w_d(V(D)),\ r(V(D)) = 1,\ d \in D,\ r \in R \rangle.$$

For each $d \in \pi(X) \cap D$, one of the above relations can be taken to be $d = d$, and hence can be deleted. □

The following is another elementary observation used later:

Lemma 2.4. *Let G be a normal subgroup of a finite group A. Suppose that A/G has a presentation $\langle Y \mid S \rangle$ and that G has a presentation $\langle X \mid R \rangle$. Then A has a presentation*

$$\langle X,\ Y \mid R,\ y^{-1}xy = w_{x,y},\ s = w_s\ \forall x \in X,\ y \in Y,\ s \in S \rangle,$$

where $w_{x,y}$ is a word in X such that the inner automorphism of A induced by $y \in Y$ sends x to $w_{x,y}$; and when s is evaluated in A it coincides with the word w_s in X. This new presentation uses $|X| + |Y|$ generators and $|R| + |S| + |X|\,|Y|$ relations.

Proof. By construction, the presented group H surjects onto A. Since G is finite, $\langle X \rangle$ is a normal subgroup of H we can identify with G. Then $H/\langle X \rangle$ is a homomorphic image of A/G. It follows that $H/\langle X \rangle \cong A/G$, and hence that $|H| = |A|$, as required. □

Curtis–Steinberg–Tits presentation. This is a standard presentation for the simply connected cover of a group of Lie type; see [Cur], [St2], [Ti2, Theorem 13.32] and [GLS, Theorem 2.9.3]. We will generally refer to [GKKL1, Sections 5.1 and 5.2] for a discussion of the versions we will use.

3. Symmetric and alternating groups

We will use a presentation for alternating groups, due to Carmichael in 1923 [Car1, p. 255] (cf. [Car2, p. 169]), that is somewhat more symmetrical than a presentation for symmetric groups due to Burnside and Miller ([Bur, p. 464], [Mi, p. 366]) in 1911 and used in [GKKL1, (2.6)]. In addition to its symmetry, Carmichael's presentation requires less data (i.e., fewer relations):

$$A_{n+2} = \langle x_1, \dots, x_n \mid x_i^3 = (x_i x_j)^2 = 1 \text{ whenever } i \neq j \rangle, \tag{3.1}$$

based on the 3-cycles $(i, n+1, n+2)$, $1 \leq i \leq n$.

In Section 3.1 we make crucial use of the symmetry of (3.1), as follows. Let $T = \langle X \mid R\rangle$ be a group acting (almost) 2-transitively on $\{1,\ldots,n\}$, viewed as acting on $\{1,\ldots,n,n+1,n+2\}$. Introduce a new generator z corresponding to the 3-cycle $(1,n+1,n+2)$. We include additional relations in order to guarantee that $|z^T| = n$ in our presented group, as well as a very small number of relations of the form $z^3 = (zz^t)^2 = 1$ for suitable $t \in T$, in order to use (3.1).

This idea will be reworked in Sections 3.1 and 3.2 in order to handle various special degrees n. The most important examples for later use are presentations in Corollaries 3.8 and 3.13 for A_{p+2} and S_{p+2} for suitable primes p. We glue two such presentations in Section 3.4, using a general idea described in Section 3.3, in order to deal with symmetric and alternating groups of arbitrary degrees.

3.1. Using 2-homogeneous groups for special degrees

We begin with an integer $n \geq 3$, together with the following ingredients:

- a group T acting transitively on the unordered pairs of distinct points in $\{1,\ldots,n\}$ (i.e., T is 2-*homogeneous*)—we do not assume that T acts faithfully on $\{1,\ldots,n\}$, nor that T acts inside A_n, although the following lemma and its proof are somewhat simpler when T induces a subgroup of A_n;
- a presentation $\langle X \mid R\rangle$ of T;
- a subset X_1 of T such that $T_1 = \langle X_1\rangle$ is the stabilizer of 1, where each element of X_1 is viewed as a word in X;
- a word w in X that moves 1 and induces an element of A_n (when w is viewed inside T).

We are given a surjection $\bar{\ }\colon T \to \bar{T}$ with $\bar{T} \leq S_n$. If T is not 2-transitive, note that $\bar{T} \leq A_n$ as $\bar{T}$ has odd order (since an involution in $\bar{T}$ would allow some *ordered* 2-set to be moved to any other one). In particular, $\bar{T} \cap A_n$ is transitive. Finally:

- View $\bar{T}$ as a subgroup of $S_{n+2} = \mathrm{Sym}\{1,\ldots,n,n+1,n+2\}$ fixing $n+1$ and $n+2$. If $t \in T$ write $\mathrm{sign}(\bar{t}) = (-1)^{\epsilon(\bar{t})}$, $\epsilon(\bar{t}) \in \{0,1\}$.

The kernel of $\bar{\ }$ will be present, but will not have any influence on our presentations. All of our results are based on the following idea:

Lemma 3.2. *If* $J := \langle X, z \mid R, z^3 = 1, (zz^w)^2 = 1, z^t = z^{\mathrm{sign}(\bar{t})}$ *for* $t \in X_1\rangle$, *then* $J \cong A_{n+2} \times T$.

Proof. Define $\varphi\colon X \cup \{z\} \to A_{n+2} \times T$ by

$$\begin{aligned} \varphi(x) &= (\bar{x} \cdot (n+1,n+2)^{\epsilon(\bar{x})}, x) \qquad \text{for } x \in X,\\ \varphi(z) &= (z',1) \quad \text{with } z' := (1,n+1,n+2). \end{aligned} \tag{3.3}$$

Then it is easy to check that the image of φ satisfies the defining relations for J, and we obtain a homomorphism $\varphi\colon J \to A_{n+2} \times T$. We claim that φ *is a surjection.* For, clearly $\langle \varphi(X)\rangle$ is naturally isomorphic to T, and acts 2-homogeneously as $\bar{T}$ on $\{1,\ldots,n\}$ when restricted to the first component. Thus, $\langle \varphi(z)^{\langle \varphi(X)\rangle}\rangle = A_{n+2} \times 1$, where A_{n+2} contains

$\bar{x}(n+1, n+2)^{\epsilon(\bar{x})}$ for each $x \in X$. Then $\varphi(J)$ also contains $(1, x)$ for each $x \in X$, so that φ is surjective, as claimed.

Then there is also a surjection $\pi : J \to A_{n+2}$.

By Lemma 2.1, J has a subgroup we can identify with $T = \langle X \rangle$. We also view z as contained in J.

Since $\langle z \rangle^{T_1} = \langle z \rangle$ by our relations, we have $|\langle z \rangle^T| \le n$; but $|\pi(\langle z \rangle^T)| = n$ and so $|\langle z \rangle^T| = n$. Consequently, T acts on $\langle z \rangle^T$ as it does on $\{1, \dots, n\}$, the kernel of $\bar{}$ acts trivially on $\langle z \rangle^T$, and we can view $\bar{T}$ as acting on $\langle z \rangle^T$.

If $t \in T$ and $\bar{t} = 1$ then $\operatorname{sign}(\bar{g}\bar{t}\bar{g}^{-1}) = 1$ for $g \in T$, so that $z^{gtg^{-1}} = z$, $(z^g)^t = z^g$, and hence $\bar{T}$ acts on z^T as T does. Then $N_{\bar{T} \cap A_n}(\langle z \rangle)$ centralizes z by the last of our relations, so that $z^{\bar{T} \cap A_n} \cap \langle z \rangle = \{z\}$, $|z^{\bar{T} \cap A_n}| = n$, and $\bar{T} \cap A_n$ acts on $z^{\bar{T} \cap A_n}$ as it does on $\{1, \dots, n\}$.

Moreover, if $\bar{T}$ is not inside A_n then our sign condition in the presentation implies that $|z^T| = 2n$ and $\bar{T} \cap A_n$ has 2 orbits on z^T, namely, $z^{\bar{T} \cap A_n}$ and $(z^{-1})^{\bar{T} \cap A_n}$.

By 2-homogeneity, any unordered pair of distinct members of $\langle z \rangle^T$ is T-conjugate to $\{\langle z \rangle, \langle z \rangle^w\}$. If $\bar{T} \cap A_n$ is 2-homogeneous, then any unordered pair of distinct members of z^T is $\bar{T} \cap A_n$-conjugate to $\{z, z^w\}$. Since the relation $(zz^w)^2 = 1$ in the presentation implies that $(z^w z)^2 = 1$, it follows that z^T satisfies (3.1), so that $N := \langle z^T \rangle \cong A_{n+2}$.

If $\bar{T} \cap A_n$ is not 2-homogeneous then, by hypothesis, $\bar{T}$ is 2-transitive but $\bar{T} \cap A_n$ is not. We claim that *we still have* $N := \langle z^T \rangle \cong A_{n+2}$. For, any ordered pair of distinct members of $\langle z \rangle^T$ is $\bar{T} \cap A_n$–conjugate to $(\langle z \rangle, \langle z \rangle^w)$ or to one other pair, $(\langle z \rangle, \langle z \rangle^y)$, say, where $y \in \bar{T} \cap A_n$. Some $g \in \bar{T} \backslash A_n$ satisfies $(\langle z \rangle, \langle z \rangle^w)^g = (\langle z \rangle, \langle z \rangle^y)$. Clearly, $z, z^w, z^y \in z^{\bar{T} \cap A_n}$. Since $g \notin A_n$, it follows that both z^g and $(z^w)^g$ lie in the other $\bar{T} \cap A_n$–class $(z^{-1})^{\bar{T} \cap A_n}$ of z. Thus, $z^g = z^{-1}$ and $(z^w)^g = (z^y)^{-1}$, so that $(zz^y)^2 = ([z^{-1}(z^w)^{-1}]^2)^g = 1$ by our relations, and we again have $N \cong A_{n+2}$ by (3.1).

Clearly $N \trianglelefteq J$ and $J/N = J/\langle z^T \rangle \cong T$. Then $|J| = |A_{n+2} \times T|$, so that $J \cong A_{n+2} \times T$. □

Examples 3.4. (1) Let p be an odd prime, $n = p + 1$ and $T = \mathrm{SL}(2, p)$. Then T has a presentation with 2 generators and 2 relations [CR2] (cf. (3.19)), while T_1 can be generated by 2 elements. Thus, by the Lemma, $A_{p+3} \times \mathrm{SL}(2, p)$ *has a presentation with* 3 *generators and* 6 *relations* (cf. Examples 3.18(1) and 3.21(9)).

(2) Let T be

$$\mathrm{AGL}(1, p) = \{x \mapsto \alpha x + \beta \mid \alpha \in \mathbb{F}_p^*,\ \beta \in \mathbb{F}_p\} = P \rtimes T_0$$

acting on $\mathbb{F}_p = \{0, \dots, p-1\}$. Here $P = \langle a \rangle$ is cyclic of odd prime order p and $T_0 = \langle b \rangle$ is cyclic of order $p - 1$, where

$$a = (0, 1, \dots, p-1) \quad \text{and} \quad b \colon x \mapsto r^{-1}x \tag{3.5}$$

if $\mathbb{F}_p^* = \langle r \rangle$. By [Neu], if $s(r-1) \equiv -1 \pmod{p}$ then

$$T = \mathrm{AGL}(1, p) = \langle a, b \mid a^p = b^{p-1},\ (a^s)^b = a^{s-1} \rangle. \tag{3.6}$$

Lemma 3.2 produces *a presentation*

$$A_{p+2} \times T = \langle a, b, z \mid a^p = b^{p-1},\ (a^s)^b = a^{s-1},\ z^3 = (zz^a)^2 = 1,\ z^b = z^{-1} \rangle$$

with 3 *generators and* 5 *relations*, since $T_0 = \langle b \rangle$ has 1 generator,

(3) Our standard example of a 2-homogeneous group that is not 2-transitive is the subgroup $T := \mathrm{AGL}(1, p)^{(2)}$ of index 2 in $\mathrm{AGL}(1, p)$ for a prime $p \equiv 3 \pmod 4$, $p > 3$. This time $T = P \rtimes T_0$ with P cyclic of order p and T_0 cyclic of order $(p-1)/2$. By [Neu],

$$T = \mathrm{AGL}(1, p)^{(2)} = \langle a, b \mid a^p = b^{(p-1)/2},\ (a^s)^b = a^{s-1} \rangle,$$

where this time $s(r-1) \equiv -1 \pmod p$ and $\mathbb{F}_p^{*2} = \langle r \rangle$; a and b behave as in (3.5). Once again $A_{p+2} \times T$ *has a presentation*

$$A_{p+2} \times T = \langle a, b, z \mid a^p = b^{(p-1)/2},\ (a^s)^b = a^{s-1}, z^3 = (zz^a)^2 = 1,\ z^b = z \rangle$$

with 3 *generators and* 5 *relations.*

(4) When $p \equiv 1 \pmod 4$, we can still use the above presentation for $T := \mathrm{AGL}(1, p)^{(2)}$, even though T has 2 orbits on unordered pairs. The argument in the lemma produces *a presentation for* $A_{p+2} \times T$ *with* 3 *generators and* 6 *relations*, using an additional relation $(zz^{w'})^2 = 1$.

(5) For future reference we note that, for any prime $p \equiv 3 \pmod 4$ with $p > 3$,

$$\mathrm{AGL}(1, p)^{(2)} \times \mathbb{Z}_2 = \langle a, b \mid a^{2p} = b^{(p-1)/2},\ (a^s)^b = a^{s-2} \rangle,$$

where this time $s(r-1) \equiv -2 \pmod p$ and $\mathbb{F}_p^{*2} = \langle r \rangle$; since s and $s+p$ both satisfy the preceding congruence, we may assume that s is odd. In order to see that the presented group T is as claimed, note that T surjects onto $\mathrm{AGL}(1, p)^{(2)} \times \mathbb{Z}_2$ via $a \mapsto (a', z)$ and $b \mapsto (b', 1)$, with a', b' playing the roles of a, b in (3) and $|z| = 2$. Since $a^{2ps} = (a^{2ps})^b = (a^{2p})^{s-2}$, we have $a^{4p} = 1$. Then $(a^{2p})^s = a^{2p}$ since s is odd, so that $(a^{ps})^b = a^{p(s-2)} = a^{-ps}$, and hence $a^{ps} = (a^{ps})^{b^{(p-1)/2}} = a^{-ps}$ since $(p-1)/2$ is odd. Now $a^{2ps} = 1$, and hence $a^{2p} = 1$ since s is odd. Clearly, $\langle a \rangle \trianglelefteq T$. Then $a^p \in Z(T)$, and T is as claimed.

Corollary 3.7. *If p is prime then* A_{p+2} *has a presentation with* 3 *generators*, 6 *relations and bit-length* $O(\log p)$.

Proof. The presentation for $A_{p+2} \times \mathrm{AGL}(1, p)$ in Example 3.4(2) has 3 generators we will call $\mathbf{a}, \mathbf{b}, \mathbf{z}$; we view w as $\mathbf{a}$. (N.B.: Here and later we use bold-faced letters in order to distinguish from letters such as a and b that already have meanings in Examples 3.4.)

Using the isomorphism (3.3) in the lemma, in $A_{p+2} \times T$ we have $\mathbf{b} = \varphi(b(p+1, p+2), b)$ for a $(p-1)$-cycle $b \in T_0$ (the point 1 in that lemma is now the point $0 \in \mathbb{F}_p$ fixed by b). Then b^a moves 0 (and fixes $p+1$ and $p+2$), so that $\varphi(\mathbf{b}^{\mathbf{a}}\mathbf{z}) = (b^a(p+1, p+2), b^a)(z', 1) = (c, b^a)$ for a p-cycle c. Thus, $\varphi((\mathbf{b}^{\mathbf{a}}\mathbf{z})^p) = (1, b^a)$. Since T is the normal closure of b^a in T, imposing the additional relation $(\mathbf{b}^{\mathbf{a}}\mathbf{z})^p = 1$ gives a presentation for A_{p+2}. □

Note that this is not, however, a short presentation (cf. Section 2).

For a more restricted choice of p we can improve the preceding result. The following is crucial for Theorem C and hence for occurrences of alternating groups later in this paper.

Corollary 3.8. *For any prime* $p \equiv 11 \pmod{12}$,

(i) $A_{p+2} \times \mathrm{AGL}(1,p)^{(2)}$ *has a presentation with* 2 *generators*, 3 *relations and bit-length* $O(\log p)$;

(ii) A_{p+2} *has a presentation with* 2 *generators*, 4 *relations and bit-length* $O(\log p)$.

Proof. (i) Since $p \equiv 3 \pmod 4$, we can let $T = \mathrm{AGL}(1,p)^{(2)} \le A_p$, r and s be as in Example 3.4(3). We will show that $A_{p+2} \times \mathrm{AGL}(1,p)^{(2)}$ *is isomorphic to the group* J *defined by the presentation*

$$\langle a, g \mid a^p = b^{(p-1)/2},\ (a^s)^b = a^{s-1},\ (zz^a)^2 = 1\rangle,$$

where $b := g^3$ *and* $z := g^{(p-1)/2}$.

First note that $A_{p+2} \times T$ satisfies this presentation. Namely, once again let T act on $\{0, \dots, p-1, p+1, p+2\}$, fixing $p+1$ and $p+2$. Let g be the product of the 3-cycle $(0, p+1, p+2)$ and an element that generates T_0 and hence has two cycles of length $(p-1)/2$ on $\{1, \dots, p-1\}$. Since $p \equiv 2 \pmod 3$, g has order $3(p-1)/2$, so that $T = \langle a, g^3\rangle$ is as in Example 3.4(3). The remaining relations $z^3 = 1$ and $z^b = z$ in that example are clear.

Now consider the presented group J. By Example 3.4(3) and Lemma 2.1, J has a subgroup we can identify with $T = \langle a, b\rangle$. In particular, $b = g^3$ has order $(p-1)/2$, and hence $|g| = 3(p-1)/2$ since $p \equiv 2 \pmod 3$. Clearly, $(zz^a)^2 = 1$ where a moves 0. Since $T_0 = \langle b\rangle$, the remaining relations $z^3 = 1$ and $z^b = z$ in Lemma 3.2 are automatic: they hold in $\langle g\rangle$. This proves (i).

(ii) This is similar to the preceding corollary. This time b has two cycles of length $(p-1)/2$, and we will see that

$$A_{p+2} \cong \langle a, g \mid a^p = g^{3\kappa}, (a^s)^{g^3} = a^{s-1}, (g^\kappa (g^\kappa)^a)^2 = (g^3 (g^\kappa)^a (g^\kappa)^{a^{-1}})^{\kappa+1} = 1\rangle \tag{3.9}$$

with $\kappa := (p-1)/2$, $s(r-1) \equiv -1 \pmod p$ and $\mathbb{F}_p^{*2} = \langle r\rangle$. For, we view (i) as a presentation for $A_{p+2} \times \mathrm{AGL}(1,p)^{(2)}$ with generators $\mathbf{a}$ and $\mathbf{g}$, and we use $\mathbf{b} := \mathbf{g}^3$ and $\mathbf{z} := \mathbf{g}^\kappa$. By (3.3), as in the proof of Corollary 3.7 we have $\varphi(\mathbf{b}\mathbf{z}^{\mathbf{a}}\mathbf{z}^{\mathbf{a}^{-1}}) = (b,b)(z'^a z'^{a^{-1}}, 1) = (c_1c_2, b)$ for disjoint cycles c_1 and c_2 of length $\kappa+1$, since $a(0) = 1$ and $a^{-1}(0) = -1$ are in different b-cycles (as $p \equiv 3 \pmod 4$). Then

$$\varphi((\mathbf{b}\mathbf{z}^{\mathbf{a}}\mathbf{z}^{\mathbf{a}^{-1}})^{\kappa+1}) = (1, b), \tag{3.10}$$

and imposing the additional relation $(\mathbf{b}\mathbf{z}^{\mathbf{a}}\mathbf{z}^{\mathbf{a}^{-1}})^{\kappa+1} = 1$ on the presentation in (i) gives (3.9). □

In [GKKL3] there is a similar presentation for $2A_{p+2} \times T$.

Remark 3.11. *In the presentations in the preceding two corollaries, replace* 0 *by* p*. Then every cycle* $(k, k+1, \dots, l)$ (*with* $k - l$ *even*) *can be written as a word of bit-length* $O(\log p)$ *in the generators. Any even permutation with bounded support can also be expressed as a word of bit-length* $O(\log p)$ *in the generators. For all of these elements, the indicated words use a bounded number of exponents.*

Namely, all 3-cycles $(k, p+1, p+2) = (p, p+1, p+2)^{a^{-k}}$ have bit-length $O(\log p)$, so the same is true of any permutation of bounded support. In particular, if $x = (p, 1)(p+1, p+2)$ then $xa = (1, \dots, p-1)(p+1, p+2)$ has bit-length $O(\log p)$, hence so do $(xa)^{-l}a^l = (p, 1, \dots, l)(p+1, p+2)^l$ and $[(xa)^{-l}a^k]^{-1}(xa)^{-l}a^l = (k, \dots, l)$ whenever $k < l \le p-1$ with $l - k$ even. The remaining cycles arise, for example, as $(k, \dots, p-1)(p-1, p, p+1) = (k, \dots, p+1)$.

Symmetric groups. There are analogous results for symmetric groups. This time we assume that our group $T = \langle X \mid R \rangle$ acts as a 2-transitive permutation group $\bar{T}$ on $\{1, \dots, n\}$. Once again write $T_1 = \langle X_1 \rangle$ where each element of X_1 is viewed as a word in X, and let w be a word in X that moves 1 when w is viewed inside T. This time we assume that $\bar{T}$ *does not lie in* A_n; let T^+ denote the subgroup of index 2 in T that induces $\bar{T} \cap A_n$. The obvious examples are $\mathrm{AGL}(1, p)$ and $\mathrm{PGL}(2, p)$. We will use the following analogue of Lemma 3.2:

Lemma 3.12. *If* $J = \langle X, z \mid R, z^3 = 1, (zz^w)^2 = 1, [z, X_1] = 1 \rangle$*, then* J *is isomorphic to a subgroup of index* 2 *in* $S_{n+2} \times T$ *that projects onto each factor.* (*In particular,* J *surjects onto* S_{n+2}*, and this quotient affords a presentation of* S_{n+2} *using one more relation if* T *is the normal closure of one of its elements.*)

Proof. View $S_{n+2} \times T$ as acting in the natural manner on the disjoint union $\{1, \dots, n, n+1, n+2\} \dot\cup \{1, \dots, n\}$, and let H be its subgroup of index 2 that induces a subgroup of A_{2n+2} (recall that $\bar{T}$ is not inside A_n). Clearly, H projects onto each factor.

Once again we view $\bar{T}$ as a subgroup of $S_{n+2} = \mathrm{Sym}\{1, \dots, n, n+1, n+2\}$ fixing $n+1$ and $n+2$.

We map J into $S_{n+2} \times T$ using a simpler version of (3.3): $\varphi(x) = (\bar{x}, x)$ for $x \in X$, and $\varphi(z) = ((1, n+1, n+2), 1)$. Since $\varphi(x)$ acts inside A_{2n+2} we have $\varphi(J) \le H$. We claim that $\varphi(J) = H$. For, as in the proof of Lemma 3.2, $\varphi(J)$ contains $A_{n+2} \times 1$ and $1 \times T^+$. Since $\bar{T}$ is not in A_n, some $\varphi(x)$, $x \in X$, lies in $H \backslash (A_{n+2} \times T^+)$. Thus, $\varphi(J) = H$.

We identify $T = \langle X \rangle$ with a subgroup of J and z with an element of J. As before, the relation $[z, X_1] = 1$ implies that T acts on z^T as it does on $\{1, \dots, n\}$, and hence is 2-transitive. Consequently, by (3.1) the relations $z^3 = 1$ and $(zz^w)^2 = 1$ imply that $N := \langle z^T \rangle \cong A_{n+2}$. Since $J/N \cong T$, we have $|J| = |A_{n+2}|\,|T| = |H|$, so that $J \cong H$. □

The following is crucial for Theorem C:

Corollary 3.13. *Let* p *be a prime.*

(i) S_{p+2} *has a presentation with* 3 *generators,* 6 *relations and bit-length* $O(\log p)$.

(ii) *If* $p \equiv 2 \pmod 3$ *then the subgroup of index* 2 *in* $S_{p+2} \times \mathrm{AGL}(1, p)$ *that projects onto each factor has a presentation with* 2 *generators*, 3 *relations and bit-length* $O(\log p)$.

(iii) *If* $p \equiv 2 \pmod 3$ *then* S_{p+2} *has a presentation with* 2 *generators*, 4 *relations and bit-length* $O(\log p)$.

Proof. Part (i) follows from the preceding lemma together with the group $T = \mathrm{AGL}(1, p)$ in Example 3.4(2), while (ii) is proved exactly as in Corollary 3.8 by using that example. This time, in (iii) we obtain

$$\begin{aligned} S_{p+2} \cong \langle a, g \mid a^p = (g^3)^{p-1}, (a^s)^{g^3} = a^{s-1}, \\ (g^{p-1}(g^{p-1})^a)^2 = (g^6(g^{p-1})^a(g^{p-1})^{a^r})^{(p+1)/2} = 1 \rangle \end{aligned} \tag{3.14}$$

with $s(r-1) \equiv -1 \pmod p$ and $\mathbb{F}_p^* = \langle r \rangle$. For, once again we view (ii) as a presentation with generators $\mathbf{a}$ and $\mathbf{g}$. Let $\mathbf{b} := \mathbf{g}^3$ and $\mathbf{z} := \mathbf{g}^{p-1}$. Using the isomorphism φ in the lemma, we have $\varphi(\mathbf{b}^2\mathbf{z}^{\mathbf{a}}\mathbf{z}^{\mathbf{a}^r}) = (b^2, b^2)(z'^a z'^{a^r}, 1) = (c_1c_2, b^2)$ for disjoint cycles c_1 and c_2 of length $(p+1)/2$, since $a^{-1}(0) = -1$ and $a^{-r}(0) = -r$ are in different b^2-cycles. Thus,

$$\varphi((\mathbf{b}^2\mathbf{z}^{\mathbf{a}}\mathbf{z}^{\mathbf{a}^r})^{(p+1)/2}) = (1, b^{p+1}) = (1, b^2). \tag{3.15}$$

Since $T \cap A_p$ is the normal closure of b^2 in T, imposing the additional relation $(\mathbf{b}^2\mathbf{z}^{\mathbf{a}}\mathbf{z}^{\mathbf{a}^r})^{(p+1)/2} = 1$ produces (3.14). □

Remark 3.16. *For the presentation in Corollary* 3.13, *the assertions in Remark* 3.11 *hold once again for even permutations. They also hold for odd permutations if* $p \equiv 11$ (mod 12).

We handle even permutations as before. Odd permutations are more delicate: they require constructing a transposition of the required bit-length, and we are only able to achieve this when $p \equiv 11 \pmod{12}$. First we recall a group-theoretic version of "Horner's Rule" [GKKL1, (3.3)] (cf. [BKL, p. 512]) for elements v, f in any group and any positive integer n:

$$vv^fv^{f^2}\cdots v^{f^n} = (vf^{-1})^n vf^n. \tag{3.17}$$

Note that

$$b_2 := \mathbf{b}^{(p-1)/2} = (1, p-1)(2, p-2)\cdots((p-1)/2, p-(p-1)/2)$$

is an odd permutation since $p \equiv 3 \pmod 4$. We use several additional permutations:

$$\begin{aligned} &c(i, j) := \mathbf{z}^{\mathbf{a}^{-i}}(\mathbf{z}^{\mathbf{a}^{-j}})^{-1}\mathbf{z}^{\mathbf{a}^{-i}} = (i, j)(p+1, p+2) \quad \text{for } 1 \le i, j \le p, \\ &v_\bullet := c(1, p-1)c(2, p-2) = (1, p-1)(2, p-2), \\ &c_{(p-1)/2} := (c(1,2)\mathbf{a})^{(p-1)/2-2}c(1,2)\mathbf{a}^{-((p-1)/2-2)} = (1, 2, \ldots, (p-1)/2) \end{aligned}$$

since $(p-1)/2$ is odd (cf. (3.17)),

$$c_\bullet := c_{(p-1)/2}\, c_{(p-1)/2}^{-\mathbf{a}^{(p+1)/2}} = (1, 2, \ldots, (p-1)/2)(p-1, p-2, \ldots, p-(p-1)/2),$$

and

$$\begin{aligned} v &:= (c(1, p-1)c_\bullet^{-1})^{(p-1)/2}c(1, p-1)c_\bullet^{(p-1)/2} = (c(1, p-1)c_\bullet^{-1})^{(p-1)/2}c(1, p-1) \\ &\equiv (1, p-1)(2, p-2)\cdots((p-1)/2, p-(p-1)/2)(p+1, p+2) \text{ (cf. (3.17))}. \end{aligned}$$

Then vb_2 is the transposition $(p+1, p+2)$ (compare Section 3.5). Now conjugate this transposition to get all others.

We do not know if the final assertion in the remark holds when $p \equiv 1 \pmod 4$.

Examples 3.18. We have been varying Example 3.4(2). Now we consider Example 3.4(1).

(1) Again consider a prime $p > 3$. We will give several presentations for A_{p+3} both here and in Example 3.21(9). Let $\mathbb{F}_p^* = \langle j \rangle$ and $j\bar{j} \equiv 1 \pmod p$. Then

$$\begin{aligned} A_{p+3} = \langle x, y, z \mid\, & x^2 = (xy)^3,\ (xy^4xy^{(p+1)/2})^2y^px^{2[p/3]} = 1, \\ & z^3 = (zz^x)^2 = [y, z] = [h, z] = 1,\ (hz^{yx}z^{y^jx})^{(p+1)/2} = 1\rangle, \end{aligned}$$

where $h := y^{\bar{j}}(y^j)^x y^{\bar{j}} x^{-1}$. This uses the following presentation for $T := \mathrm{SL}(2, p)$, obtained in [CR2] using [Sun]:

$$\mathrm{SL}(2, p) = \langle x, y \mid x^2 = (xy)^3,\ (xy^4xy^{(p+1)/2})^2y^px^{2[p/3]} = 1\rangle, \tag{3.19}$$

where x and y arise from elements of order 4 and p, respectively (corresponding to the matrices t and u given later in (4.4)). Then $T_\infty = \langle X_\infty \rangle$ with $X_\infty := \{y, h\}$ in notation imitating that in Lemma 3.2. The final relation in the above presentation for A_{p+3} is obtained as in the proof of Corollary 3.8(ii). In Example 3.21(9) we will decrease the number of relations, but we will not need this for our main results.

Another presentation for A_{p+3} is

$$\begin{aligned} A_{p+3} = \langle u, h, t, z \mid\, & u^p = t^2 = 1,\ u^h = u^{j^2},\ h^t = h^{-1},\ t = uu^tu,\ ht = u^{\bar{j}}(u^j)^tu^{\bar{j}}, \\ & z^3 = (zz^t)^2 = [u, z] = [h, z] = 1,\ (hz^{ut}z^{u^jt})^{(p+1)/2} = 1\rangle. \end{aligned}$$

This uses Lemma 3.2 together with the presentation for $T := \mathrm{PSL}(2, p)$ given in [GKKL1, Theorem 3.6]. Similar presentations can be obtained using the presentations for $\mathrm{PSL}(2, p)$ in [Fr] or [To].

(2) Once again let $\mathbb{F}_p^* = \langle j \rangle$. Then

$$\begin{aligned} S_{p+3} = \langle u, h, t, z \mid\, & u^p = t^2 = 1,\ u^h = u^j,\ h^t = h^{-1},\ t = uu^tu, \\ & z^3 = (zz^t)^2 = [u, z] = [h, z] = 1,\ (hz^{ut})^{p+1} = 1\rangle. \end{aligned}$$

This uses Lemma 3.2 together with a presentation $\langle u, h, t \mid u^p = t^2 = 1,\ u^h = u^j,\ h^t = h^{-1},\ t = uu^tu\rangle$ for $T := \mathrm{PGL}(2, p)$ analogous to [GKKL1, Theorem 3.6]. The final relation is obtained as in the proof of Corollary 3.8(ii).

Table 2. $S_{n+2},\ A_{n+2}$: Presentations for some small n

G	$n+2$	T	$\|R\|$	ρ	$\|X_1\|$	$\|x_1\|$	gens	rels	in Ex. ♯
S_{11}	$9+2$	AGL(1, 9)	4	1	1	8	2	6	3.21(1)
A_{11}	$9+2$	PSL(2, 8)	2	2	1	7	2	5	3.21(2)
A_{11}	$9+2$	AGL$(1,9)^{(2)}$	4	2	1	4	2	7	3.21(3)
S_{12}	$10+2$	2PGL(2, 9)	2	1	2	16	2	5	3.21(4)
A_{12}	$10+2$	6PSL(2, 9)	2	1	2	8	2	5	3.21(5)
A_{23}	$21+2$	12PSL(3, 4)	2	1	2	5	2	5	3.21(6)
A_{23}	$\binom{7}{2}+2$	$6A_7$	2	2	2	5	2	6	3.21(7)
A_{24}	$22+2$	$12M_{22}$	2	1	2	7	2	5	3.21(8)
A_{24}	$2\cdot 11+2$	AGL$(1,11)^{(2)}\times\mathbb{Z}_2$	2	3	1	5	2	6	3.21(12)
A_{24}	$2\cdot 11+2$	AGL(1, 11)	2	4	1	5	2	7	3.21(13)
A_{47}	$\binom{10}{2}+2$	A_{10}	2	2	2	8	2	6	3.21(10)
A_{47}	$\binom{10}{2}+2$	$A_{10}\times$ SL(2, 7)	4	2	2	7	2	8	3.21(11)
A_{48}	$2\cdot 23+2$	AGL$(1,23)^{(2)}\times\mathbb{Z}_2$	2	3	1	11	2	6	3.21(12)
A_{48}	$2\cdot 23+2$	AGL(1, 23)	2	4	1	11	2	7	3.21(13)

3.2. Small n

In order to handle a few degrees $n < 50$ we will need further variations on the idea used in Corollaries 3.8 and 3.13. All of the general presentations below have bit-length $O(\log n)$, but this is not significant since our goal involves bounded n. We suspect that most readers will prefer to skip this section.

In Table 2 we summarize presentations for $G = S_{n+2}$ or A_{n+2} needed later. For this table and our variations on Corollaries 3.8 and 3.13 we use the following notation:

- The group T acts as a transitive group $\bar{T} \le S_n$ on $\{1, \dots, n\}$.
- T has exactly ρ orbits of unordered pairs of distinct points.
- $T = \langle X \mid R\rangle$.
- $T_1 = \langle X_1\rangle$, where T_1 is again the stabilizer of 1.
- $x_1 \in X_1 \cap X$ has order $|x_1| = k$ not divisible by 3.
- Each $x' \in X_1' := X_1\backslash\{x_1\}$ is given as a word $w_{x'}(x_1, X')$, where $X' := X\backslash\{x_1\}$.
- $\bar{T}$ is also viewed as a subgroup of Sym$\{1, \dots, n+2\}$ fixing $n+1$ and $n+2$.
- Let H denote the preimage in $S_{n+2} \times T$ of $A_{2n+2} \cap (S_{n+2} \times \bar{T})$ when $S_{n+2} \times \bar{T}$ is viewed as acting on the disjoint union $\{1, \dots, n, n+1, n+2\} \dot{\cup} \{1, \dots, n\}$.

Proposition 3.20. (i) *If $\bar{T} \cap A_n$ is the normal closure of one of its elements, then the group $H/(1 \times (\bar{T} \cap A_n))$ has a presentation with $|X|$ generators and $|R| + \rho + |X_1|$ relations.*

(ii) *If T induces a subgroup of A_n, then $A_{n+2} \times T$ has a presentation with $|X|$ generators and $|R| + \rho + |X_1| - 1$ relations.*

Sketch of proof. Let $w_1, \dots, w_\rho$ be words in X such that the ρ pairs $\{1, w_j^{-1}(1)\}$ are in different T-orbits. Each $r \in R$ is a word $r(x_1, X')$ in $X = \{x_1\} \cup X'$. We will show that

$$J := \langle X', g \mid r(g^3, X') = (g^k (g^k)^{w_j})^2 = [g^k, w_{x'}(g^3, X')] = 1\ \forall r \in R\ \forall j\ \forall x' \in X_1' \rangle$$

is isomorphic to H. Let $z' := (1, n+1, n+2)$ and let the integer ν satisfy $3\nu \equiv 1 \pmod{k}$, so that $g := x_1^\nu z'$ satisfies $g^3 = x_1$ and $g^k = z'^{\pm 1}$. Then J surjects onto H as in the proof of Lemma 3.12.

Now consider J, and let $x_1 := g^3$. As before, we can view T as the subgroup $\langle X', g^3 \rangle = \langle X', x_1 \rangle$ of J. Our relations imply that $z := g^k$ commutes with $X_1' \cup \{x_1\}$ and hence with T_1. Then $|z^T| = n$, and we can use (3.1) as before to prove both (i) and (ii). □

Examples 3.21(12) and 3.21(13) contain further variations on the idea behind the proposition.

Examples 3.21. (1) $n + 2 = 11$: $T = \mathrm{AGL}(1, 9)$ has the presentation $\langle a, b \mid a^3 = b^8 = 1, a^{b^2} = aa^{-b}, [a, a^b] = 1 \rangle$, $\rho = |X_1| = 1$ and $|x_1| = 8$, so that S_{11} *has a presentation with* 2 *generators and* $4 + 1 + 1$ *relations.* However, for use in Theorem C it is easier simply to use the presentation for S_{11} with 2 generators and 6 relations in [Ar, p. 54] (cf. [CoMo, p. 64]).

(2) $n + 2 = 11$: $T = \mathrm{PSL}(2, 8)$ has a presentation with 2 generators and 2 relations [CHRR1], $\rho = 1$, $|X_1| = 2$ and $|x_1| = 7$, so that Proposition 3.20 produces a *presentation of* A_{11} *with* 2 *generators and* $2 + 2 + 1$ *relations.* It has long been known that A_{11} has a presentation with 2 generators and 6 relations [CoMo, p. 67].

(3) $n + 2 = 11$: T has index 2 in $\mathrm{AGL}(1, 9)$, T has the presentation $\langle a, b \mid a^3 = b^4 = 1, a^{b^2} = a^{-1}, [a, a^b] = 1 \rangle$, $\rho = 2$, $|X_1| = 1$ and $|x_1| = 4$, so that A_{11} *has a presentation with* 2 *generators and* $4 + 2 + 1$ *relations.*

(4) $n + 2 = 12$: $T = 2\mathrm{PGL}(2, 9)$ has presentations with 2 generators and 2 relations provided by Havas [Hav], such as

$$2\mathrm{PGL}(2, 9) = \langle a, b \mid (b^{-1}a^{-1})^3 baba^{-1} = a^{-2}b^{-1}a^4ba^{-1}b^{-1}ab = 1 \rangle,$$

where $x_1 := a$ has order 16. (Havas found this using [CHRR1, Method 2], modified to handle groups that are not necessarily perfect.) This time $\rho = 1$ and $|X_1| = 2$, so that S_{12} *has a presentation with* 2 *generators and* $2 + 1 + 2$ *relations.* Once again, it has long been known that S_{12} has a presentation with 2 generators and 7 relations [Ar, p. 54] (cf. [CoMo, p. 64]).

(5) $n + 2 = 12$: $T = 6\mathrm{PSL}(2, 9) \cong 6A_6$ has a presentation $6A_6 = \langle a, b \mid ab^3(ba)^{-4} = (ab^2ab^{-2})^2ab^2 = 1 \rangle$ with 2 generators and 2 relations [Ro], $\rho = 1$, $|X_1| = 2$ and $|ba^4| = 8$ in T, so that A_{12} *has a presentation with* 2 *generators and* $2 + 1 + 2$ *relations.* Once again, A_{12} is known to have a presentation with 2 generators and 7 relations [CoMo, p. 67].

(6) $n+2=23$: $T=12\mathrm{PSL}(3,4)$ has a presentation with 2 generators and 2 relations [CHRR1], $\rho=1$, $|X_1|=2$ and $|x_1|=5$, so that Proposition 3.20 produces a *presentation of* A_{23} *with* 2 *generators and* $2+2+1$ *relations.*

(7) $n+2=23$: $T=6A_7$ has a presentation with 2 generators and 2 relations [CRKMW], $n=\binom{7}{2}$, $\rho=2$, $|X_1|=2$ and $|x_1|=5$, so that A_{23} *has a presentation with* 2 *generators and* $2+2+2$ *relations.*

(8) $n+2=24$: $T=12M_{22}$ has a presentation with 2 generators and 2 relations [CHRR1], $\rho=1$, $|X_1|=2$ and $|x_1|=7$, so that Proposition 3.20 produces a presentation of A_{24} *with* 2 *generators and* $2+2+1$ *relations.*

(9) *If* $p>3$ *is prime then* $A_{p+3}\times\mathrm{SL}(2,p)$ *has a presentation with* 2 *generators and* 4 *relations*: apply Proposition 3.20 using (3.19) and $|R|=2=|X_1|$, $\rho=1$, $x_1=y$. It follows that A_{p+3} *has a presentation with* 2 *generators and* 5 *relations.* We will need this below in (11), including the fact that g^3 has order p and fixes $p+2$.

(10) $n+2=47$: $T=A_{10}$ has the following presentation with 2 generators and 3 relations [Hav]:

$$\begin{aligned} A_{10}=\langle a,b \mid a^3b^{-1}ab^{-1}a^3ba^2b &= a^2b^{-1}a^5b^{-3}a^3 \\ &= a^{-2}bab^{-1}aba^3bab^{-1}aba^{-2}b^{-1}=1\rangle, \end{aligned}$$

with $|a|=15$, $|b|=12$ and $|ab|=8$. View T as acting on $\binom{10}{2}$ unordered pairs with $\rho=2$, $|X_1|=2$ and $x_1=ab$, so that Proposition 3.20 produces a *presentation of* A_{45+2} *with* 2 *generators and* $2+2+2$ *relations.*

At present there is no known presentation for $2A_{10}$ with 2 generators and 2 relations. It would lead to a presentation for A_{47} with 2 generators and 5 relations.

(11) *If* $p>3$ *is prime then* $A_{\binom{p+3}{2}+2}$ *has a presentation with* 2 *generators and* 8 *relations.* For, let $T=A_{p+3}\times\mathrm{SL}(2,p)$ act on $\binom{p+3}{2}$ unordered pairs of a set of size $p+3$, with $\mathrm{SL}(2,p)$ acting trivially. Apply Proposition 3.20 using (9), with $|X|=2$, $|R|=4$, $\rho=2$, $|X_1|=2$ and $x_1=g^3$.

There is a similar presentation for $A_{\binom{p+2}{2}+2}$.

(12) *For any prime* $p\equiv 11 \pmod{12}$, A_{2p+2} *has a presentation with* 2 *generators and* 6 *relations.* We will vary the argument in Proposition 3.20 (and Lemma 3.2), using the transitive subgroup $T:=\mathrm{AGL}(1,p)^{(2)}\times\langle t\rangle$ of the transitive group $\mathrm{AGL}(1,p)\times\langle t\rangle$ of degree $2p$, where t is an involution interchanging two blocks of size p (namely, $\{0,\dots,p-1\}$ and its image under t). Note that the stabilizer of a point is cyclic of odd order $(p-1)/2$. Moreover, T has $\rho=3$ orbits of unordered pairs of the $2p$-set, with orbit-representatives $\{0,t(0)\}$, $\{0,1\}$, and $\{0,t(1)\}$ (note that the stabilizer in T of $\{0,t(1)\}$ is trivial). Clearly T is not in A_{2p}; if $w\in T$ then $\mathrm{sign}(w)$ will refer to the action of w on these $2p$ points.

We view T as a subgroup of A_{2p+2} preserving the set consisting of two new points $2p+1$ and $2p+2$, with each member of the set $\mathrm{AGL}(1,p)^{(2)}t$ of odd permutations on the $2p$-set interchanging these points.

Let

$$J := \langle a, g \mid a^{2p} = b^{(p-1)/2},\ (a^s)^b = a^{s-2},\ (zz^{\operatorname{sign}(w_i)w_i})^2 = 1 \text{ for } i = 1, 2, 3\rangle,$$

where $s(r-1) \equiv -2 \pmod p$ with s odd, $\mathbb{F}_p^{*2} = \langle r\rangle$, $b := g^3$, $z := g^{(p-1)/2}$, and with suitable words $w_1, w_2, w_3 \in T$ (such that the pairs $\{z, z^{w_i}\}$ are in different T-orbits on z^T; e.g., $\{z, z^t\}$, $\{z, z^a\}$ and $\{z, z^{at}\}$). As in the proof of Proposition 3.20, using Example 3.4(5) we find that $A_{2p+2} \times T$ satisfies our presentation.

As usual, using Example 3.4(5) we can view T as the subgroup $\langle a, b\rangle$ of J. Exactly as in Lemma 3.2 (and Proposition 3.20), $|z| = 3$, $|z^T| = 2p$, and $(zz^g)^2 = (z^{-t}z^{-gt})^2 = 1$ for all $g \in \mathrm{AGL}(1, p)^{(2)}$ with $z^g \neq z$, while $(zz^{-gt})^2 = 1$ for all $g \in \mathrm{AGL}(1, p)^{(2)}$.

Then $N := \langle z^T\rangle \cong A_{2p+2}$ by (3.1), and hence $J = NT \cong A_{2p+2} \times T$. One further relation gives a presentation of A_{2p+2} with 2 generators and $2 + 3 + 2$ relations.

Explicitly,

$$\begin{aligned} A_{2p+2} = \langle a, g \mid a^{2p} = b^{(p-1)/2},\ &(a^s)^b = a^{s-2},\\ &(zz^{-t})^2 = (zz^a)^2 = (zz^{-at})^2 = 1,\ (bz^{a^{-1}}z^a z^{a^{-1}t} z^{at})^p = 1\rangle, \end{aligned}$$

with z, r and s as above and once again $b := g^3$, $z := g^{(p-1)/2}$ and $t := a^p$, where the last relation is obtained as in the proof of Corollary 3.7.

(13) *For any odd prime* $p \equiv 2 \pmod 3$, A_{2p+2} *has a presentation with* 2 *generators and* 7 *relations.* Unlike in the preceding example we now handle the case $p \equiv 1 \pmod 4$ using $T = \mathrm{AGL}(1, p)$. Let a, b, r and s be as in Example 3.4(2). Note that $c := ab^2$ fixes $n := (1 - r^2)^{-1}$, and that $\tilde{c} := b^{a^{-n}}$ satisfies $\tilde{c}^2 = c$.

Let T act transitively on the $2p$ cosets of $\langle c^2\rangle$. This action of T has 2 blocks: $\{a^i\langle c^2\rangle \mid 0 \leq i < p\}$ and $\{a^i\tilde{c}\langle c^2\rangle \mid 0 \leq i < p\}$. There are 4 orbits of unordered pairs of points, with representatives $\{\langle c^2\rangle, a\langle c^2\rangle\}$, $\{\langle c^2\rangle, a^r\langle c^2\rangle\}$, $\{\langle c^2\rangle, \tilde{c}\langle c^2\rangle\}$ and $\{\langle c^2\rangle, a\tilde{c}\langle c^2\rangle\}$. Let "1" $:= \langle c^2\rangle$.

We replace the presentation for T in Example 3.4(2) by

$$T = \langle c, b \mid (cb^{-2})^p = b^{p-1},\ ((cb^{-2})^s)^b = (cb^{-2})^{s-1}\rangle.$$

Once again we view T as a subgroup of A_{2p+2} preserving the set consisting of the additional points $2p + 1$ and $2p + 2$. Signs will again refer to the actions of elements of T on the $2p$-set.

Consider the group

$$J := \langle g, b \mid (cb^{-2})^p = b^{p-1},\ ((cb^{-2})^s)^b = (cb^{-2})^{s-1},\ (zz^{\operatorname{sign}(w_i)w_i})^2 = 1 \text{ for } i = 1, 2, 3, 4\rangle,$$

where $c := g^3$, $z := g^{(p-1)/2}$, and $\{1, w_i^{-1}(1)\}$, $1 \leq i \leq 4$, are representatives for the orbits of T on pairs of the $2p$ points. Then J surjects onto $A_{2p+2} \times T$, and we can view $T = \langle c, b\rangle \leq J$. In particular, $|g| = 3(p-1)/2$, and hence $z^3 = 1$.

Since c centralizes z, as usual we obtain $|z^T| = 2p$ and $\langle z^T\rangle \cong A_{2p+2}$ by (3.1), and then $J \cong A_{2p+2} \times T$. One further relation produces the desired presentation of A_{2p+2}.

Examples 3.22. We conclude with additional examples of Proposition 3.20, this time using a presentation involving A_n to deduce one involving A_{n+2}. Specifically, in Proposition 3.20(ii) assume that $T = \langle X \mid R\rangle$ is "essentially" A_n, with $\rho = 1$, $|X_1| = 2$ and a suitable x_1. Then $A_{n+2} \times T$ has a presentation with $|X|$ generators and $|R| + 1 + 2 - 1$ relations. This idea was already used in Example 3.21(11).

(1) *For any prime* $p \equiv 11 \pmod{12}$, A_{p+4} *has a presentation with* 2 *generators and* 6 *relations.* For, we can use Corollary 3.8(i) in order to obtain a presentation for $A_{p+4} \times (A_{p+2} \times \mathrm{AGL}(1, p)^{(2)})$ with 2 generators and $3 + 2$ relations (note that $x_1 := a$ has order p in that corollary).

(2) *For any prime* $p > 3$, A_{p+5} *has a presentation with* 2 *generators and* 7 *relations.* For, we can use Example 3.21(9) to obtain a presentation for $A_{p+5} \times (A_{p+3} \times \mathrm{SL}(2, p))$ with 2 generators and $4 + 2$ relations ($x_1 := x$ has order 4 in that example).

3.3. Gluing alternating and symmetric groups

We now turn to all alternating and symmetric groups, starting with a general gluing lemma:

Lemma 3.23. *Let* $G = \langle X \mid R\rangle$ *and* $\bar G = \langle \bar X \mid \bar R\rangle$ *be presentations of* S_n *and* S_m, *respectively, and let* $m, n > k \ge l + 2 \ge 4$. *Consider embeddings* $\pi : G \to S_{m+n-k}$ *and* $\bar\pi : \bar G \to S_{m+n-k}$ *into* $\mathrm{Sym}\{-m + k + 1, \dots, n\}$ *such that*

$$\pi(G) = \mathrm{Sym}(\{1, \dots, n\}) \quad \textit{and} \quad \bar\pi(\bar G) = \mathrm{Sym}(\{-m + 1 + k, \dots, k\}).$$

Suppose that $a, b, c, d \in G$ *and* $\bar a, \bar b, \bar c, \bar e \in \bar G$, *viewed as* words *in* X *or* $\bar X$, *respectively, are nontrivial permutations such that the following all hold*:

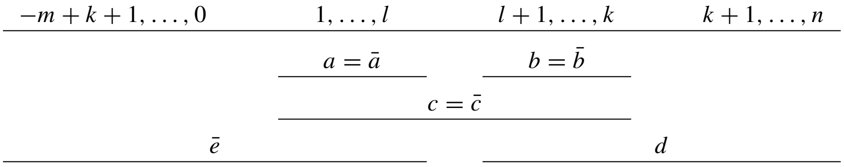
- $\pi(a) = \bar\pi(\bar a) \in \mathrm{Sym}(\{1, \dots, l\}) < \pi(G) \cap \bar\pi(\bar G)$,
- $\pi(b) = \bar\pi(\bar b) \in \mathrm{Sym}(l + 1, \dots, k) < \pi(G) \cap \bar\pi(\bar G)$,
- $\pi(c) = \bar\pi(\bar c) \in \mathrm{Sym}(\{1, \dots, k\}) < \pi(G) \cap \bar\pi(\bar G)$,
- $\pi(d) \in \mathrm{Sym}(\{l + 1, \dots, n\}) < \pi(G)$,
- $\bar\pi(\bar e) \in \mathrm{Sym}(\{-m + 1 + k, \dots, l\}) < \bar\pi(\bar G)$,
- $\langle \pi(a), \pi(c)\rangle = \mathrm{Sym}(\{1, \dots, k\}) < \pi(G) \cap \bar\pi(\bar G)$,
- $\langle \pi(b), \pi(d)\rangle = \mathrm{Sym}(\{l + 1, \dots, n\}) < \pi(G)$,
- $\langle \bar\pi(\bar a), \bar\pi(\bar e)\rangle = \mathrm{Sym}(\{-m + 1 + k, \dots, l\}) < \bar\pi(\bar G)$.

Then

$$J := \langle X, \bar X \mid R, \bar R,\ a = \bar a,\ c = \bar c,\ [d, \bar e] = 1\rangle \tag{3.24}$$

is isomorphic to $S_{m+n-k} = \mathrm{Sym}\{-m + k + 1, \dots, n\}$, *where* $\langle X\rangle = S_n$ *acts on* $\{1, \dots, n\}$ *and* $\langle \bar X\rangle = S_m$ *acts on* $\{-m + 1 + k, \dots, k\}$.

The following picture might be helpful.



Note that b and $\bar{b}$ do not appear in the presentation, but we need to know that such elements exist.

Proof. The restrictions on m, n, k and l are designed to guarantee that the desired permutations exist. There is an obvious surjection $J \to S_{m+n-k}$ (note that our 3 extra relations are satisfied). By Lemma 2.1, J has subgroups we identify with $G = \langle X \rangle = S_n$ and $\bar{G} = \langle \bar{X} \rangle = S_m$.

In particular, our relations state that $a = \bar{a}$ and $c = \bar{c}$ in J. Then the assumption $\pi(b) = \bar{\pi}(\bar{b})$ states that b and $\bar{b}$ represent the same element of $\langle a, c\rangle = \langle \bar{a}, \bar{c}\rangle$, so that the additional relation $b = \bar{b}$ is forced to hold in J. Then we also have the following relations in J:

- $[d, \bar{a}] = [d, a] = 1$ because $d, a \in G = S_n$ have disjoint supports,
- $[d, \bar{e}] = 1$ by the last relation in the presentation (3.24),
- $[b, \bar{a}] = [b, a] = 1$ because $b, a \in G$ have disjoint supports,
- $[b, \bar{e}] = [\bar{b}, \bar{e}] = 1$ because $\bar{b}, \bar{e} \in \bar{G}$ have disjoint supports.

Therefore

$$[\langle b, d\rangle, \langle \bar{a}, \bar{e}\rangle] = 1, \tag{3.25}$$

where $\langle b, d\rangle = \mathrm{Sym}(\{l+1, \dots, n\})$ and $\langle \bar{a}, \bar{e}\rangle = \mathrm{Sym}(\{-m+1+k, \dots, l\})$.

The symmetric groups G and $\bar{G}$ are generated, respectively, by the $n-1$ and $m-1$ transpositions $x_i := (i, i+1)$, $1 \le i < n$, and $x_i := \overline{(i, i+1)}$, $-m+1+k \le i < k$. The identification of the two copies of $S_k = \langle a, c\rangle = \langle \bar{a}, \bar{c}\rangle$ in (3.24) identifies the transpositions x_i, $1 \le i < k$, common to these generating sets, producing a generating set of J consisting of $m+n-k-1$ involutions. These involutions satisfy the relations in the Coxeter presentation [Moo]

$$S_{m+n-k} = \langle x_i, -m+1+k \le i < n \mid x_i^2 = (x_i x_{i+1})^3 = (x_i x_j)^2 = 1 \\ \text{for all possible } i, j \text{ with } j-i \ge 2\rangle:$$

any two x_i either both lie in G, or both lie in $\bar{G}$, or they commute by (3.25) since one is in $\langle b, d\rangle$ and the other is in $\langle \bar{a}, \bar{e}\rangle$. □

Remark 3.26. Although the preceding presentation may not even have bit-length $O(\log n)$, there is a great deal of flexibility in the choice of the elements a, b, c, d, $\bar{a}, \bar{b}$, $\bar{c}$, $\bar{e}$ for this and other purposes (cf. Remarks 3.11 and 3.16, Proposition 3.38 and Theorem 3.40). The last three relations in (3.24) are similar to ones appearing in the proof of [GKKL1, Theorem 3.17]; and they are used both there and here in essentially the same manner. However, the situation in that paper was more delicate, due to length considerations.

We have just glued presentations of symmetric groups using $|R|$ and $|\bar{R}|$ relations in order to obtain a presentation of a larger symmetric group using $|R| + |\bar{R}| + 3$ relations. The next lemma does the same for alternating groups. In particular, bounded presentations lead to further bounded presentations for these types of groups; together with small bit-length (cf. Remarks 3.11 and 3.16), this may suffice for some purposes. However, since we wish to use as few relations as we can, in Corollary 3.28 we will deduce $\bar{R}$ as a conjugate $\bar{R} = R^y$.

Lemma 3.27. *A presentation of A_{m+n-k} is obtained as in the preceding lemma by replacing symmetric groups by alternating groups throughout* (3.24) *and assuming that* $m, n > k \geq l + 3 \geq 6$.

Proof. Once again, the restrictions on m, n, k and l are designed to guarantee that the desired permutations exist. The previous picture can again be used. As in the proof of the preceding lemma, (3.24) implies that (3.25) holds.

We will use the presentation (3.1), this time with the union of the generating sets $x_i := (1, 2, i)$ for G and $x_j := \overline{(1, 2, j)}$ for $\bar{G}$, where $3 \leq i \leq n$ and $-m+1+k \leq j \leq k$ but $j \neq 1, 2$. As above, (3.24) implies that $(1, 2, i) = \overline{(1, 2, i)}$ if $3 \leq i \leq k$.

Then all required relations in (3.1) are immediate except for

$$\big((1, 2, i)\overline{(1, 2, j)}\big)^2 = 1 \text{ with } k < i \leq n \text{ and } -m + 1 + k \leq j \leq 0.$$

Recall that $\langle b, d\rangle = \mathrm{Alt}(\{l + 1, \dots, n\})$, where $6 \leq l + 3 \leq k < i \leq n$. Then some $g \in \langle b, d\rangle$ sends k to i and fixes 1 and 2. By (3.25), g commutes with $\overline{(1, 2, j)} \in \langle \bar{a}, \bar{e}\rangle$. Consequently,

$$\big[\big((1, 2, i)\overline{(1, 2, j)}\big)^2\big]^g = \big((1, 2, k)\overline{(1, 2, j)}\big)^2 = \big(\overline{(1, 2, k)}\,\overline{(1, 2, j)}\big)^2 = 1$$

(note that $k \neq j$ since $j \leq 0 < k$). Thus, (3.1) holds in all cases. □

Corollary 3.28. *If A_m has a presentation with M relations and if $m > k \geq 6$, then A_{2m-k} has a presentation with $M + 4$ relations. The same holds for the corresponding symmetric groups using the weaker assumption $m > k \geq 4$.*

Proof. Let $G = \langle X \mid R\rangle$ be a presentation for A_m with M relations. In Lemma 3.27 we use $m = n$ and $l = 3$, but this time we introduce an additional generator y corresponding to an even permutation sending $\{-m + 1 + k, \dots, k\}$ to $\{1, \dots, m\}$ and inducing the identity on $\{1, \dots, k\}$.

Consider the group

$$J := \langle X, y \mid R,\ a = a^y,\ c = c^y,\ [d, e^y] = 1\rangle, \tag{3.29}$$

with $a, b, c, d, \bar{a} := a^y, \bar{b} := b^y, \bar{c} := c^y, \bar{e} := e^y$ playing the same roles as in Lemma 3.27 (in particular, as in Lemma 3.23 they are *words* in $X \cup \bar{X}$ where $\bar{X} := X^y$). By Lemma 3.23 with $\bar{R} := R^y$, J has a subgroup $K := \langle X, X^y\rangle \cong A_{2m-k}$.

Finally, we add an extra relation to (3.29), expressing y as a word w in $X \cup X^y$, in order to ensure that our generator y is in K and that, as an element of $A_{2m-k} = \mathrm{Alt}(\{-m + k + 1, \dots, m\})$, the action of y is as described above.

The group S_{2m-k} is dealt with in the same manner. □

Remark 3.30. We have just glued two subgroups A_m in order to obtain a group A_{2m-k}, or two subgroups S_m in order to obtain a group S_{2m-k}, in each case with suitable restrictions on m and k. There is a variation on this process that glues two subgroups S_m in order to obtain a group A_{2m-k} (view S_m as lying in A_{m+2}, as occurred in Section 3.1).

3.4. All alternating and symmetric groups

Before proving Theorem C, we begin with a weaker result:

Proposition 3.31. *For all $n \geq 5$, A_n and S_n have presentations with* 3 *generators and* 10 *relations.*

Proof. If $n \leq 9$ then we have already obtained a presentation with fewer relations than required. By Ramanujan's version of Bertrand's Postulate [Ra], in all other cases we can write $n = 2p + 4 - k$ for a prime p and an integer k such that $m := p + 2 > k \geq 6$, and then use Corollaries 3.7, 3.13(i) and 3.28. (A related use of Bertrand's Postulate appears in [GKKL1, Theorem 3.9].) □

This proposition is weaker than Theorem C in two significant ways: the number of relations is larger than in that theorem, and bit-length is not mentioned. We deal with the second of these as follows:

Lemma 3.32. *In Corollary* 3.28 *and Proposition* 3.31, *assume that either*

(i) *the group is $A_n = A_{2m-k}$, or*
(ii) *the group is $S_n = S_{2m-k}$ and we used a prime $p \equiv 11 \pmod{12}$.*

Then it is possible to choose a, c, d, e such that a represents a 3*-cycle and the resulting presentation has bit-length $O(\log n)$ with a bounded number of exponents, each of which is at most n.*

Proof. In Lemmas 3.23 and 3.27 and Corollary 3.28 we can choose each of the elements a, c, d, e to represent a product of a cycle of the form $(i, \dots, j)$ with $i < j$ and a permutation having bounded support. (When we give explicit relations in Section 3.5(4), each of these permutations will be chosen to be a cycle.) We require a to represent a 3-cycle; and then in the symmetric group case (ii) we also need c and e to represent odd permutations (cf. the hypotheses of Lemma 3.23).

By Remarks 3.11 and 3.16, in both (i) and (ii) we can choose a, c, d, e to have bit-length $O(\log n)$ and with exponents as required. It remains to show that *the permutation represented by y can be chosen to meet the requirements in Corollary* 3.28: the crucial relation expressing y as word in $X \cup X^y$ must have bit-length $O(\log n)$, with exponents as required.

There is a reasonable amount of flexibility in the choice of y in Corollary 3.28; we will view y as a permutation and choose $y \in A_n$. In that corollary we were permuting the $n = 2p + 4 - k$ points (recall that $m = p + 2$)

$$-p-1+k, -p+k, \dots, -1, 0;\ 1, \dots, k;\ k+1, k+2, \dots, p+1, p+2, \tag{3.33}$$

where we have alternating or symmetric groups on the first and last $p + 2$ points, with an A_k or S_k on the overlap.

If $p - k$ is even, choose y to be the following product of $p + 2 - k$ transpositions:

$$y := (-p-1+k, p+2)(-p+k, p+1) \cdots (-1, k+2)(0, k+1). \tag{3.34}$$

We must write y as a word of the required bit-length in $X \cup X^y$. We use the following additional permutations:

- $x := (-1, k+2)(0, k+1) = [(1, k+2)(2, k+1)]^{(-1,1)(0,2)}$, which we have written using permutations from our two copies of A_{p+2};
- $u^{-1} := (1, \dots, k, k+1, \dots, p+2)(1, \dots, k, 0, -1, \dots, -p+k-1)$
  $= (1, \dots, k, k+1, \dots, p+2)(1, \dots, k, k+1, \dots, p+2)^y$.

By Remark 3.11, u and hence also s can be expressed as a word of bit-length $O(\log p)$ in $X \cup X^y$ using a bounded number of exponents; then, by (3.17), so can

$$y = xx^{u^2}x^{u^4} \cdots x^{u^{p-k}} = (xu^{-2})^{(p-k)/2}xu^{p-k}. \tag{3.35}$$

If $p-k$ is odd let $v := (-p-1+k, p+2)(-p+k, p+1) \cdots (-3, k+4)$ and use

$$\begin{aligned} y &:= v(-2, k+3)(-1, k+2, 0, k+1) \\ &= v[(2, k+3)(1, k+2, k, k+1)]^{(2,-2)(1,-1)(0,k)(1,2)}. \end{aligned} \tag{3.36}$$

Then v can be expressed as a word of bit-length $O(\log p)$ in $X \cup X^y$ using a calculation similar to (3.35), and the final term in the bottom line of (3.36) is a product of permutations from our two copies of A_{p+2}. Another application of Remark 3.11 completes the proof. □

Remark 3.37. *Every cycle* $(k, k+1, \dots, l)$ *in* A_n, *and every element with bounded support in* A_n, *has bit-length* $O(\log n)$ *in our generators*, by yet another application of Remark 3.11. Similar statements hold for symmetric groups in the situation of Remark 3.16.

With a bit more number theory, together with Table 2, we obtain an improvement of Proposition 3.31 that is needed for Theorem C:

Proposition 3.38. *If* $n \geq 5$ *then* S_n *and* A_n *have presentations with* 3 *generators*, 8 *relations and bit-length* $O(\log n)$. *Moreover*, *these presentations use a bounded number of exponents*, *each of which is at most* n.

Proof. We refine the argument in Proposition 3.31. First consider S_n. Here we need to write $n = 2p+4-k$ for a prime $p \equiv 2 \pmod 3$ such that $m := p+2 > k \geq 4$, so that we can use Corollaries 3.13 and 3.28, and then continue as in the proof of Proposition 3.31. In view of the requirements on bit-length and exponents, we also require that $p \equiv 11 \pmod{12}$ when n is not bounded, so that Lemma 3.32 will complete the proof for S_n.

The requirements on p are that $n/2 \leq p \leq n-3$ and, in general, $p \equiv 11 \pmod{12}$. The existence of such a prime p is guaranteed by Dirichlet's Theorem for large n. However, we need a more precise (and effective) result of this type. This is provided in [Mor] (updating [Bre, Er, Mol] with more precise estimates): if $n \geq 50$ then there is such a prime $p \equiv 11 \pmod{12}$. A straightforward examination when $n < 50$ leaves the cases $n \leq 7$ and $n = 11, 12$ or 13—since for small n we only need the requirement $p \equiv 2 \pmod 3$—and these were dealt with earlier.

For A_n we need to write $n = 2p+4-k$ for a prime $p \equiv 11 \pmod{12}$ such that $m := p+2 > k \geq 6$, then use Corollaries 3.8 and 3.28, and again finish as in the proof of Proposition 3.31 by using Lemma 3.32. This time $(n+2)/2 \leq p \leq n-3$. Once again, by [Mol, Mor], if $n \geq 50$ then there is such a prime p. Another straightforward examination leaves the cases $n \leq 13$ and $n = 21, 22, 23, 24, 25, 45, 46, 47, 48$ or 49 to deal with. The cases in which $n = p + 2$ or $p + 3$ for some prime p are handled in Corollary 3.7 and Example 3.4(1), and Table 2 deals with the remaining cases. □

For the next theorem we will use a presentation in the preceding proposition that is valid for most n. If $n \geq 50$ (or, more precisely, if n is not one of the exceptions mentioned in the above proof), then (3.29) together with one further relation $y = w$ is such a presentation:

$$A_n \text{ or } S_n = \langle X, y \mid R,\ a = a^y,\ c = c^y,\ [d, e^y] = 1,\ y = w\rangle, \tag{3.39}$$

with $\langle X \mid R\rangle$ in (3.9) or (3.14) for the cases A_n or S_n, respectively, a, c, d, e words in X as in Lemma 3.32, and a suitable word w in $X \cup X^y$ as in (3.34)–(3.36). (The properties required of w are described at the end of the proof of Corollary 3.28 and, in gory detail, in the proof of Lemma 3.32. Alternatively, see Section 3.5(6).)

We are now able to prove the main part of Theorem C:

Theorem 3.40. *If $n \geq 5$ then S_n and A_n have presentations with 3 generators, 7 relations and bit-length $O(\log n)$. Moreover, these presentations use a bounded number of exponents, each of which is at most n.*

Proof. Let $n = 2m - k$ with $m := p + 2 > k \geq 6$ for a prime p (cf. the preceding proposition; below we will discuss the existence of a suitable p).

Let $G := \langle X \mid R'\rangle$ be the group presented in Corollary 3.8(i) or 3.13(ii), so that $|X| = 2$, $|R'| = 3$ and one of the following holds:

$$\begin{array}{llll} A_n \textit{ case}: & T = \mathrm{AGL}(1, p)^{(2)}, & p \equiv 11 \pmod{12}, & G \cong A_m \times T; \\ S_n \textit{ case}: & T = \mathrm{AGL}(1, p), & p \equiv 2 \pmod 3, & G/(1 \times (T \cap A_m)) \cong S_m. \end{array}$$

(In the A_n case, T has index 2 in $\mathrm{AGL}(1, p)$; in the S_n case, G has index 2 in $S_m \times T$.) We also require that $p \equiv 11 \pmod{12}$ in the S_n case when $n \geq 50$, in order to obtain the desired bit-length.

Further background concerning G. Despite the fact that we have been dealing with groups such as T and G for a while, we need further properties and notation involving these groups.

The relations R' and the relations R in (3.39) are closely connected:

$R = R' \dot\cup \{h = 1\}$, where h is a word in X that represents the element $(1, t) \in S_m \times T$ on the right side of (3.10) or (3.15).

Thus, $h = 1$ is the relation used to kill the normal subgroup $1 \times (T \cap A_m)$ of G that is not inside our target group A_n or S_n, and $t = b$ or b^2 has order $(p - 1)/2$. Below we will need the fact that 3 *does not divide the order of t* (since $(3, p - 1) = 1$). This property was already crucial for obtaining the presentation $G = \langle X \mid R'\rangle$.

The presentation (3.39) can now be rewritten

$$A_n \text{ or } S_n = \langle X, y \mid R',\ h,\ a = a^y,\ c = c^y,\ [d, e^y] = 1,\ y = w\rangle, \tag{3.41}$$

with 3 generators, 8 relations and bit-length $O(\log n)$.

There was a great deal of freedom in our choice of the words a, b, c, d, e in Corollary 3.28 (cf. (3.29)). We will choose a to represent a 3-cycle in the alternating or symmetric group $G/(1 \times (T \cap A_m))$ (cf. Lemma 3.32).

Let $\varphi\colon F_X \to G \leq S_m \times T$ be the surjection in the proof of Corollary 3.8(i) or 3.13(ii) (cf. (3.3)). We use the following ingredients:

- $h \in F_X$ is, as above, the word on the left side of (3.10) or (3.15), so that $\varphi(h) = (1, t)$;
- $a \in F_X$ with $\varphi(a) = (a_1, *)$ for a 3-cycle $a_1 \in A_m$;
- $\tilde{a}, \hat{a} \in F_X$ such that $\varphi(\tilde{a}) = (a_1, 1)$ and $\varphi(\hat{a}) = (a_1, t)$.

This completes the additional background concerning G. *We claim that the* 7*-relator group*

$$J := \langle X, y \mid R', \tilde{a} = \hat{a}^y,\ c = c^y,\ [d, e^y] = 1,\ y = w\rangle \tag{3.42}$$

is isomorphic to the group in (3.41).

For, there is a natural surjection $\psi\colon F_{X\cup\{y\}} \to J$. It suffices to show that the image of ψ satisfies (3.41): $\psi(h) = 1$ and $\psi(\hat{a}) = \psi(\hat{a})^{\psi(y}$ for the word $\hat{a}$ in X that represents $a_1 \in A_m$ (i.e., in (3.41) we are replacing a by $\hat{a}$).

There is also a surjection $\pi\colon G \to \psi(\langle X\rangle)$ such that $\psi = \pi\varphi$ on F_X (since $\psi(\langle X\rangle)$ satisfies the defining relations of $G = \langle X \mid R'\rangle$). Since $(a_1, 1)$ has order dividing 3, so do $\pi((a_1, 1)) = \pi\varphi(\tilde{a}) = \psi(\tilde{a}) = \psi(\hat{a}^y) = \pi\varphi(\hat{a})^{\psi(y)} = \pi((a_1, t))^{\psi(y)},\ \pi((a_1, t))$ and hence also $\pi((1, t))$. We already noted that 3 does not divide the order of t, so that $\psi(h) = \pi\varphi(h) = \pi((1, t)) = 1$. Consequently, $\psi(\hat{a}) = \pi\varphi(\hat{a}) = \pi((a_1, 1))\pi((1, t)) = \pi((a_1, 1)) = \pi\varphi(\tilde{a}) = \psi(\tilde{a}) = \psi(\hat{a})^{\psi(y)}$. Thus, the image of ψ satisfies (3.41), as claimed.

Bit-length: As in the proof of Theorem 3.31, we may assume that the words a, c, d and e have bit-length $O(\log n)$. In Corollaries 3.8(i) and 3.13(ii) the generators of $1 \times (T \cap A_m)$ have bit-length $O(\log n)$ in X, hence $\tilde{a}$ and $\hat{a}$ can be chosen so that the same is true for these elements.

Finally, we need to discuss whether we have handled all of the groups A_n and S_n; or, what amounts to the same thing, for which n a prime p can be found satisfying all of the conditions we have imposed.

As in Proposition 3.38, (3.42) takes care of A_n except if $n \leq 13$ or $n = 21, 22, 23, 24, 25, 45, 46, 47, 48, 49$; and these are handled exactly as in that proposition.

For the S_n case we have imposed a further condition beyond what was used in Proposition 3.38: we need to write $n = 2p + 4 - k$ for a prime $p \equiv 2 \pmod 3$ such that $m = p + 2 > k \geq l + 2 \geq 5$, and $p \equiv 11 \pmod{12}$ if n is not bounded. (The conditions in Corollary 3.28 were that $m = p + 2 > k \geq l + 2 \geq 4$, but here we need to be able to find a 3-cycle a_1 in A_l.) Once again these requirements can be met for all n except if $n \leq 7$ or $n = 9, 10, 11$, and those cases can be handled as before. □

No presentation in this or the preceding section has length $O(\log n)$.

By Remarks 3.11 and 3.16, the exponents in (3.35) are all less than n. As already noted, these presentations have *bounded expo-length* (cf. Section 2). See Remark 7 in Section 11 for comments concerning the boundedness of expo-length for other families of almost simple groups. Since it is natural to prefer specific generators (such as the familiar ones $(1,2)$ and $(1,2,\dots,n)$), we note the following consequence of the preceding theorem and Lemma 2.3:

Corollary 3.43. *Let $G = A_n$ or S_n, $n \geq 5$.*

(i) *If a and b are any generators of G, then there is a presentation of G using 2 generators that map onto a and b, and 9 relations.*

(ii) *There is a presentation for G using 2 generators and 8 relations.*

Proof. Part (i) follows from Theorem 3.40 and Lemma 2.3 since $|\pi(X) \cap D| \geq 0$.

For (ii), note that we have provided a presentation $\langle X \mid R\rangle$ for G such that some element of X projects onto an element $a \in G$ that is either a 3-cycle (z in Lemma 3.2) or has a power that is a 3-cycle (such as g in Corollary 3.8(ii) or Proposition 3.20). Let b be any element of G such that $G = \langle a, b\rangle$. Now use $D = \{a, b\}$ in Lemma 2.3 (compare Section 11, Remark 4). □

Remark 3.44. The preceding presentations are not short in any sense. *In the case of the pair* $\{a, b\} = \{(1,2),\ (1,2,\dots,n)\}$, the bit-length is $O(n\log^2 n)$. For, by Remark 3.37 these generators a and b have bit-length $O(\log n)$ in the generators used in the Theorem. Also, since all cycles $(1,2,\dots,k)$, $2 \leq k \leq n$, have length $O(\log n)$ in $\{a,b\}$ (using (3.17)), induction shows that all elements of S_n have bit-length $O(n\log n)$ in $\{a,b\}$. Hence, by the proof of Lemma 2.3, the presentation has the stated length. This should be compared to Theorem A2 in the Appendix, which implies that any presentation using these generators a and b has length (*not* bit-length) at least $2n$ if $n > 2$.

3.5. An explicit presentation for S_n

The presentations in Sections 3.1 and 3.2 are not difficult to understand, and they visibly encode information concerning various alternating or symmetric groups. However, the presentations in Theorem 3.40 are not as explicit as one might wish. Therefore, we will provide a presentation of S_n for n odd (see Remark 3.45 for even n). Although this presentation is in no sense elegant or informative, it has the advantage of being explicit while using only 7 relations and having bit-length $O(\log n)$.

We will use a prime $p \equiv 11 \pmod{12}$ such that $n-3 \geq p \geq (n+2)/2$. (This places a mild restriction on n, as seen in the proof of Theorem 3.40. For $n \geq 50$ there is always such a prime.)

Let $k = 2p + 4 - n$, so that $p + 2 > k \geq 6$. Then $k \equiv n \equiv 1 \pmod 2$, so that $p - k$ is even.

The desired presentation is

$$\begin{aligned} S_n = \langle \mathbf{a}, \mathbf{g}, \mathbf{y} \mid {} & \mathbf{a}^p = (\mathbf{g}^3)^{p-1},\ (\mathbf{a}^s)^{\mathbf{g}^3} = \mathbf{a}^{s-1}, \\ & (\mathbf{g}^{p-1}(\mathbf{g}^{p-1})^{\mathbf{a}})^2 = 1,\ a = \hat{a}^{\mathbf{y}},\ c = c^{\mathbf{y}},\ [d, e^{\mathbf{y}}] = 1,\ \mathbf{y} = w\rangle, \end{aligned}$$

for words $a, c, d, e, \hat{a}, w$ defined below and integers r and s such that $s(r-1) \equiv -1 \pmod p$ and $\mathbb{F}_p^{*2} = \langle r\rangle$.

Notes: We write $\mathbf{a}$ and $\mathbf{b}$ in order to distinguish the uses of these letters in Sections 3.1 and 3.2 from those in Sections 3.3 and 3.4.

The permutations in S_n indicated below are not part of the presentation, but are provided in order to help keep track of the map $\langle \mathbf{a}, \mathbf{g}, \mathbf{y}\rangle \to S_n$ onto the symmetric group on the n points listed in (3.33). We view AGL$(1, p)$ as acting on $\{1, \dots, p\}$ with $\mathbf{a} \equiv (1, \dots, p) \in$ AGL$(1, p)$, but we use p in place of 1 in Lemma 3.2.

The notation used here should **not** be viewed mod p; in particular, 0 and p are different.

(1) $\mathbf{z} := \mathbf{g}^{p-1} \equiv (p, p+1, p+2)$,
$\mathbf{b} := \mathbf{g}^3$ (so that $\langle \mathbf{a}, \mathbf{b}\rangle$ is AGL$(1, p)$ as in (3.6)),
$\mathbf{y} \equiv (-p-1+k, p+2)(-p+k, p+1)\cdots(-1, k+2)(0, k+1)$.

(2) $\mathbf{z}(i) := \mathbf{z}^{\mathbf{a}^{-i}} \equiv (i, p+1, p+2)$ for $1 \le i \le p$,
$c(i, j) := \mathbf{z}(i)\mathbf{z}(j)^{-1}\mathbf{z}(i) \equiv (i, j)(p+1, p+2)$ for $1 \le i < j \le p$,
$c_i := (c(1, 2)\mathbf{a})^{i-2}c(1, 2)\mathbf{a}^{-(i-2)} \equiv (1, 2, \dots, i)$ whenever i is odd with $3 \le i \le p$ (cf. (3.17)).

(3) (Constructing a transposition as in Remark 3.16.)
$b_2 := \mathbf{b}^{(p-1)/2} \equiv (1, p-1)(2, p-2)\cdots((p-1)/2, p-(p-1)/2)$,
$c_\bullet := c_{(p-1)/2}c_{(p-1)/2}^{-\mathbf{a}^{(p+1)/2}} \equiv (1, 2, \dots, (p-1)/2)(p-1, p-2, \dots, p-(p-1)/2)$,
$v := (c(1, p-1)c_\bullet^{-1})^{(p-1)/2}c(1, p-1)$
$\equiv (1, p-1)(2, p-2)\cdots((p-1)/2, p-(p-1)/2)(p+1, p+2)$ (cf. (3.17)),
$t := vb_2c(1, 2) \equiv (1, 2)$.

(4) $a := \mathbf{z}(3)^{\mathbf{z}(1)\mathbf{z}(2)} \equiv (1, 2, 3)$,
$c := tc_k \equiv (2, \dots, k)$ (an odd permutation),
$d := c_3^{-1}\mathbf{az} \equiv (3, \dots, p+2)$,
$e := \mathbf{a}c_k^{-1}t\mathbf{z} \equiv (1, 2, k+1, \dots, p, p+1, p+2)$,
so that $e^{\mathbf{y}} \equiv (1, 2, 0, -1, \dots, -p+k-1)$ (odd permutations).

(5) $\hat{a} := a(\mathbf{b}^2z(1)z(-1))^{(p+1)/2}$ as in (3.15).

(6) $x := [c(1, k+2)c(2, k+1)]^{c(1,k+2)^{\mathbf{y}}c(2,k+1)^{\mathbf{y}}}$
$\equiv (-1, k+2)(0, k+1)$,
$u^{-1} := (\mathbf{az})(\mathbf{az})^{\mathbf{y}}$
$\equiv (1, \dots, k, k+1, k+2, \dots, p, p+1, p+2)(1, \dots, k, 0, -1, \dots, -p+k-1)$,
$w := (xu^{-2})^{(p-k)/2}xu^{p-k}$
$\equiv (-p-1+k, p+2)(-p+k, p+1)\cdots(-1, k+2)(0, k+1) \equiv \mathbf{y}$ (cf. (3.17)).

Remarks 3.45. 1. In the isomorphism given by (3.3), $\mathbf{z}$ maps to an element of $A_{p+2} \times 1$. Hence, the element a defined above also maps into $A_{p+2} \times 1$, so that the element $\tilde{\mathbf{a}}$ used in (3.42) is just our a. The remainder of the presentation given above is a straightforward translation from Section 3.4.

2. A presentation for the alternating groups is similar but slightly simpler: only even permutations are involved.

3. *Changes needed when n and k are even*:

(4′) $c := c_{k-1}t \equiv (1, \dots, k)$ (an odd permutation),
$e := \mathbf{a}tc_{k-1}^{-1}\mathbf{z} \equiv (1, k+1, \dots, p+2)$ (another odd permutation).

(6′) $w := t^{\mathbf{az}} t^{\mathbf{az}y} t^{\mathbf{az}} (xu^{-2})^{(p-k-1)/2} x u^{p-k-1}$
$\equiv (-p-1+k,\, p+2)(-p+k,\, p+1)\cdots(-1,\, k+2)(0,\, k+1) \equiv \mathbf{y}$
as before.

3.6. Weyl groups

It is easy to use Theorem 3.40 to obtain presentations for the Weyl groups of types B_n or D_n (cf. Lemma 3.47). Instead we begin with a subgroup W_n of those Weyl groups that is needed in Section 10.

Let $W_n := \mathbb{Z}_2^{n-1} \rtimes A_n$ be the subgroup of the monomial group of $\mathbb{R}^n$ such that $\mathbb{Z}_2^{n-1}$ consists of all ± 1 diagonal matrices of determinant 1, and the alternating group A_n permutes the standard orthonormal basis of $\mathbb{R}^n$. We will write elements of W as permutations of $\{1, \ldots, n, -1, \ldots, -n\}$.

Proposition 3.46. *If $n \geq 4$ then W_n has a presentation with* 4 *generators*, 11 *relations and bit-length $O(\log n)$. If $n = 4$ or* 5 *then W_n also has a presentation with* 3 *generators and* 7 *relations.*

Proof. Let $\langle X \mid R\rangle$ be the presentation for $A = A_n$ in Theorem 3.40 (or (3.1) or [CoMo, p. 137] when $n = 4$ or 5 in order to have a presentation with 2 generators and 3 relations). Let $\sigma = (3, 2, 1) \in A$, and let X_{12} consist of two words in X such that $\langle X_{12}\rangle$ is the stabilizer of the 2-set $\{1, 2\}$. We will show that W_n is isomorphic to the group J with the following presentation.

Generators: X, s (where s represents $(1, -1)(2, -2) = \mathrm{diag}(-1, -1, 1, \ldots, 1)$).

Relations:

(1) R.
(2) $s^2 = 1$.
(3) $[s, X_{12}] = 1$.
(4) $ss^{\sigma}s^{\sigma^2} = 1$.

There is an obvious surjection $\pi\colon J \to W_n$. We view $A = \langle X\rangle \leq J$. By (3), $\binom{n}{2} \geq |s^A| \geq |\pi(s^A)| = \binom{n}{2}$, so that s^A can be identified with the 2-sets in $I = \{1, \ldots, n\}$. Thus, there are well-defined elements $s_{ij} = s_{ji} \in s^A$ for all distinct $i, j \in I$.

By (4), $s_{1j}s_{jk}s_{k1} = 1$ whenever $1, j, k$ are distinct, so that A is generated by the elements s_{1i}. Since all s_{ij} are involutions, it follows that $s = s_{12}$ commutes with all s_{1j}, and hence also with all s_{jk}, so that $N := \langle s^A\rangle$ is elementary abelian of order $\leq 2^{n-1}$. Then $J = AN$ has order $\leq |W_n|$ and hence is W_n.

Now $|X| + 1$ and $|R| + 4$ are as stated. Bit-length is straightforward to check. □

We will need another similar result in the next section:

Lemma 3.47. *The Weyl group of type B_n has a presentation with* 4 *generators*, 11 *relations and bit-length $O(\log n)$.*

Proof. Let $\langle X \mid R\rangle$ be the presentation in Theorem 3.40, let X_1 consist of two words in X such that $\langle X_1\rangle$ is the stabilizer of the point 1, and let $x \in X$ move 1.

Then the Weyl group W is isomorphic to the group $J := \langle X, y \mid R,\ y^2 = [y^2, X_1] = [y, y^x] = 1\rangle$. For, W is a surjective image of J, $|y^J| \le n$ and $N := \langle y^J\rangle$ is an elementary abelian 2-group. Then $J/N \cong S_n$ and $|N| \le 2^n$, as required. □

Remark 3.48. Later we will need a similar group, $\mathbb{Z}_4^n \rtimes A_n$, with a similar presentation using $y^4 = 1$, with 4 generators, 11 relations and bit-length $O(\log n)$.

3.7. Aut(F_n)

One application of our results concerns the free group F_n:

Theorem 3.49. *If* $n > 2$ *then* $\mathrm{Aut}(F_n)$ *has a presentation with* 5 *generators*, 18 *relations and bit-length* $O(\log n)$; *and a presentation with* 19 *generators*, 65 *relations and length* $O(\log n)$.

We will use a presentation of a subgroup of index 2 in $\mathrm{Aut}(F_n)$ due to Gersten [Ger, Theorem 2.8]. Let $B = \{b_1, \ldots, b_n\}$ be a basis of F_n, $n > 2$, and let "bar" interchange $x \in B$ and x^{-1}. Gersten's presentation uses generators E_{ab}, $a, b \in B \cup \bar{B}$, $a \ne b, \bar{b}$, and the following relations:

(G1) $E_{ab}^{-1} = E_{a\bar{b}}$.
(G2) $[E_{ab}, E_{cd}] = 1$ if $a \notin \{c, d, \bar{d}\}$; $b \notin \{c, \bar{c}\}$.
(G3) $[E_{bc}, E_{ab}] = E_{ac}$ if $a \notin \{c, \bar{c}\}$.
(G4) $w_{ab} = w_{\bar{a}\bar{b}}$ where $w_{ab} = E_{ba}E_{\bar{a}b}E_{\bar{b}a}$.
(G5) $w_{ab}^4 = 1$.

Here E_{ab} is the Nielsen map $a \mapsto ab$ fixing all members of $(B \cup \bar{B})\backslash\{a, \bar{a}\}$.

Proof of Theorem 3.49. Let $\langle Y \mid S\rangle$ be the presentation of the Weyl group W in Lemma 3.47. We view W as acting on the set $\{i, \bar{\imath} \mid 1 \le i \le n\}$ in the obvious manner. Let σ and τ be elements of W, written as words in Y, that generate the stabilizer in W of both 1 and $\{2, \bar{2}\}$, where $\sigma(2) = 2$. We also view a few specific permutations as words in Y.

We will show that $\mathrm{Aut}(F_n)$ is the group J having the following presentation.

Generators: Y, g (here g is Gersten's E_{12}).

Relations:

(1) S.
(2) $g^\sigma = g,\ g^\tau = g^{-1}$.
(3) $[g, g^{(13)}] = 1$.
(4) $[g, g^{(1,3)(2,4)}] = 1$.
(5) $[g, g^{(1,\bar{1})(2,3)}] = 1$.
(6) $[g^{(3,2,1)}, g] = g^{(2,3)}$.
(7) $gg^{(\bar{2},\bar{1},2,1)}g^{(1,\bar{1})(2,\bar{2})} = (1, 2, \bar{1}, \bar{2})$.

The map sending $g \mapsto E_{12}$ and W to the obvious permutations of $B \cup \bar{B}$ maps J to $\mathrm{Aut}(F_n)$. It is surjective since $\mathrm{Aut}(F_n)$ is generated by Nielsen transformations.

As usual we can identify W with a subgroup $\langle Y\rangle$ of J. Relations (2) imply that the elements of g^W can be labeled E_{ab} as above. Since $\tau(2)=\bar{2}$, (G1) holds.

Relations (G3) follow from (1), (2) and (6).

Relations (G2) fall into 6 orbits under the action of W_n, with representatives

$$[E_{ab},E_{cd}]=[E_{ab},E_{cb}]=[E_{ab},E_{c\bar{b}}]=[E_{ab},E_{\bar{a}d}]=[E_{ab},E_{\bar{a}b}]=[E_{ab},E_{\bar{a}\bar{b}}]=1.$$

The first orbit follows from (4), while the second and third follow from (3) and (G1), and the fourth follows from (5). Then E_{12} commutes with E_{32}, $E_{\bar{1}3}$ and hence with $E_{\bar{1}2}$ using (6), which takes care of the fifth orbit. Finally, the sixth orbit follows from the fifth and (G1).

Relation (7) implies that $E_{12}E_{2\bar{1}}E_{\bar{1}2}=(1,2,\bar{1},\bar{2})=(\bar{1},\bar{2},1,2)=E_{\bar{1}\bar{2}}E_{\bar{2}1}E_{1\bar{2}}$, so that (G4) and (G5) hold.

By Gersten's presentation, $N:=\langle g^W\rangle$ is (isomorphic to) a subgroup of $\mathrm{Aut}(F_n)$, and $N\trianglelefteq J$. Since J/N is an image of W in which $(1,2,\bar{1},\bar{2})$ is trivial by (7), $|J/N|\le 2$. There is an obvious surjection $F_{Y\cup\{g\}}\to\mathbb{Z}_2$ sending $g\mapsto 1$ and $y\mapsto\det(y)\in\mathbb{Z}_2$ for $y\in Y\subset W$, where y is viewed as monomial matrix. Since all relations (1)–(7) lie in the kernel of this map, $|J/N|\le 2$. Also, $(1,\bar{1})$ is in $J\backslash N$, and its action on N is the same as that of the automorphism of F_n sending $b_1\mapsto b_1^{-1}$ and fixing all remaining members of B. Hence, $J\cong\mathrm{Aut}(F_n)$.

This is a presentation of $\mathrm{Aut}(F_n)$ with at most $4+1$ generators, $11+7$ relations and bit-length $O(\log n)$. For the second presentation in the theorem, in Lemma 3.47 replace Theorem 3.40 by [GKKL1, Theorem 3.17]. □

Another presentation for $\mathrm{Aut}(F_n)$, with slightly more relations, can be obtained by combining Lemma 3.47 with [AFV].

3.8. SL(n, $\mathbb{Z}$)

A simpler application of Theorem 3.40 is to groups over $\mathbb{Z}$:

Theorem 3.50. *For all $n\ge 6$, $\mathrm{SL}(n,\mathbb{Z})$ has a presentation with 4 generators and 16 relations.*

Proof. We will use the following presentation for $G=\mathrm{SL}(n,\mathbb{Z})$, essentially due to Nielsen and Magnus (cf. [Mil, p. 81]).

Generators: E_{ij}, $1\le i,j\le n, i\ne j$.

Relations:

(1) $[E_{ij},E_{km}]=1$ whenever $j\ne k$ and $i\ne m$.
(2) $[E_{ij},E_{jk}]=E_{ik}$ whenever i,j,k are distinct.
(3) $(E_{12}E_{21}^{-1}E_{12})^2=1$.

Let $\langle X\mid R\rangle$ be a presentation for $T=A_n$, which is embedded in $\mathrm{SL}(n,\mathbb{Z})$ as permutation matrices. Let X_{12} be a pair of generators for the stabilizer of 1 and 2, viewed as words in X (as are the various permutations used below). Let J be the group with the following presentation.

Generators: X, e (where e plays the role of E_{12}).

Relations:

(1) R.
(2) e commutes with X_{12}, $e^{(4,3,2)}$, $e^{(1,2,3)}$, $e^{(3,1,4)}$ and $e^{(1,3)(2,4)}$.
(3) $[e, e^{(3,2,1)}] = e^{(4,3,2)}$.
(4) $(e(e^{(1,2)(3,4)})^{-1}e)^2 = 1$.

The usual argument, using the 4-transitivity of T, shows that $J \cong \mathrm{SL}(n, \mathbb{Z}) \times T$. One further relation, killing an element of the form $(1, t)$ for a 3-cycle t, produces the desired presentation. This has $3 + 1$ generators and $6 + 9 + 1$ relations using the presentation in Theorem 3.40. □

In view of the Steinberg presentation, the groups $\mathrm{SL}(n, p)$ over prime fields $\mathbb{F}_p$ are obtained in a uniform manner by adding the relation $e^p = 1$ to the above presentation. There are similar but easier presentations for $n \le 5$. Other groups over $\mathbb{Z}$ can be dealt with similarly, using presentations in [War, Behr, St4].

4. Rank 1 groups

4.1. Steinberg presentation

Each rank 1 group G we consider has a Borel subgroup $B = U \rtimes \langle h\rangle$, with U a p-group. There is an involution t (mod $Z(G)$ in the case $\mathrm{SL}(2, q)$ with q odd) such that $h^t = h^{-1}$ (or $h^t = h^{-q}$ if G is unitary). The *Steinberg presentation* for G [St3, Sec. 4] consists of the following ingredients:

- a presentation for B,
- a presentation for $\langle h, t\rangle$,
- $|U| - 1$ relations of the form
$$u_0^t = u_1 h_0 t u_2, \tag{4.1}$$
with $u_0, u_1, u_2 \in U\backslash\{1\}$ and $h_0 \in \langle h\rangle$ (one relation for each choice of u_0).

4.2. Polynomial notation

Our groups will always come equipped with various elements having names such as u or h. For any polynomial $g(x) = \sum_{i=0}^{e} g_i x^i \in \mathbb{Z}[x]$, $0 \le g_i < p$, define powers as follows:

$$[[u^{g(x)}]]_h = (u^{g_0})(u^{g_1})^{h^1} \cdots (u^{g_e})^{h^e}, \tag{4.2}$$

so that

$$[[u^{g(x)}]]_h = u^{g_0}h^{-1}u^{g_1}h^{-1}u^{g_2} \cdots h^{-1}u^{g_e}h^e \tag{4.3}$$

by “Horner’s Rule” [GKKL1, (4.14)] (compare (3.17)).

As in [GKKL1, Sec. 4.3], we need to be careful about rearranging the terms in (4.2) when not all of the indicated conjugates of u commute.

4.3. SL$(2, q)$

In [CRW2] there is a presentation for PSL$(2, q)$ with at most 13 relations. We will provide a presentation having fewer relations, based on the matrices

$$u = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad t = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad h = \begin{pmatrix} \zeta^{-1} & 0 \\ 0 & \zeta \end{pmatrix}, \tag{4.4}$$

where $\mathbb{F}_q^* = \langle \zeta \rangle$.

Theorem 4.5. SL$(2, q)$ *and* PSL$(2, q)$ *have presentations with* 3 *generators*, 9 *relations and bit-length* $O(\log q)$. *When q is even,* PSL$(2, q)$ *has a presentation with* 3 *generators*, 5 *relations and bit-length* $O(\log q)$.

Proof. If $q \le 9$ then SL$(2, q)$ and PSL$(2, q)$ have presentations with 2 generators and at most 4 relations (e.g., [CoMo, pp. 137–138]). Assume that $q > 9$, and let $k, l \in \mathbb{Z}$ be such that $\zeta^{2k} = \zeta^{2l} + 1$ and $\mathbb{F}_q = \mathbb{F}_p[\zeta^{2k}]$ (as in [GKKL1, Section 3.5.1]).

Set $d = \gcd(k, l)$. Then $\mathbb{F}_q = \mathbb{F}_p[\zeta^{2d}]$.

Let $m(x) \in \mathbb{F}_p[x]$ be the minimal polynomial of ζ^{2d}. If $\gamma \in \mathbb{F}_q$, let $g_\gamma(x) \in \mathbb{F}_p[x]$ satisfy $g_\gamma(\zeta^{2d}) = \gamma$ and $\deg g_\gamma < \deg m$.

We will show that $G = $ SL$(2, q)$ (or PSL$(2, q)$; cf. relation (5) below) is isomorphic to the group J having the following presentation.

Generators: u, t, h.

Relations:

(1) $u^p = 1$.
(2) $u^{h^k} = uu^{h^l} = u^{h^l}u$.
(3) $[[u^{m(x)}]]_{h^d} = 1$ in the notation of (4.2).
(4) $u^h = [[u^{g_{\zeta^2}(x)}]]_{h^d}$.
(5) $[t^2, u] = 1$ (or $t^2 = 1$ in the case PSL$(2, q)$ with q odd).
(6) $h^t = h^{-1}$.
(7) $t = uu^t u$.
(8) $ht = [[u^{g_{\zeta^{-1}}(x)}]]_{h^d}\, [[u^{g_\zeta(x)}]]^t_{h^d}\, [[u^{g_{\zeta^{-1}}(x)}]]_{h^d}$.

Matrix calculations using (4.4) easily show that there is a surjection $J \to G$. We view u, t, h as elements of J. By (1), (2) and [GKKL1, Lemma 4.1] (compare [Bau, CR1, CRW2]), $U := \langle u^{\langle h^k, h^l \rangle} \rangle$ is elementary abelian; since $d = \gcd(k, l)$ we have $U = \langle u^{\langle h^d \rangle} \rangle$. By (1) and (3) we can identify U with the additive group of $\mathbb{F}_q$ in such a way that h^d acts as multiplication by ζ^{2d}. By (4), h acts on U as an automorphism of order $(q-1)/(2, q-1)$.

By (7) and (8), $J = \langle U, U^t \rangle$, and J is perfect since the action of h in (4) implies that $U = [U, h]$. Moreover, by (4) and (6), $z := h^{(q-1)/(2,q-1)}$ is inverted by t, centralizes U and U^t, and hence is an element of $Z(J)$ having order 1 or 2.

Thus, $\langle u, h \rangle / \langle z \rangle$ is isomorphic to a Borel subgroup of PSL$(2, q)$. By (6), $\langle h, t \rangle / \langle z \rangle$ is dihedral of order $2(q-1)/(2, q-1)$.

We already know that $\langle h\rangle$ acts on the nontrivial elements of U with $(2, q-1)$ orbits; an orbit representative is u, and also $[[u^{g_\zeta(x)}]]_{h^d}$ if q is odd. As in the proof of [GKKL1, Sec. 4.4.1], (7) and (8) provide the relations (4.1) required to let us deduce that $J/\langle z\rangle \cong \mathrm{PSL}(2, q)$.

Now J is a perfect central extension of $\mathrm{PSL}(2, q)$, and hence is $\mathrm{SL}(2, q)$ or $\mathrm{PSL}(2, q)$ (since U is abelian, $6\mathrm{PSL}(2, 9)$ and $3\mathrm{PSL}(2, 9)$ cannot occur). Finally, (5) distinguishes between these groups when q is odd. The bit-length of the presentation follows from (4.3).

Finally, if q is even there are significant simplifications. We may assume that $l=d=1$, so that $h^d = h$ acts on $U = \langle u^{\langle h^k, h^l\rangle}\rangle$; the induced automorphism has order $q-1$ by (3), and (4) can be deleted. The last relation in (2) can be deleted since we have an involution written as a product of two involutions. Relation (5) can be deleted since (1) and (7) imply that $t^2 = 1$. Lastly, (8) is not needed since $\langle h\rangle$ has only one orbit on the nontrivial elements of U, and since $h^2 = t^{h^{-1}}t \in J$ and $|h|$ odd imply that $h \in J$. Explicitly,

$$\begin{aligned}\mathrm{PSL}(2, 2^e) = \langle u, h, t \mid u^2 = 1,\ & u^{h^k} = uu^h,\\ & (u^{m_0})(u^{m_1})^{h^1} \cdots (u^{m_e})^{h^e} = 1,\ h^t = h^{-1},\ t = uu^tu\rangle,\end{aligned}$$

where $\zeta + 1 = \zeta^k$ with $\langle\zeta\rangle = \mathbb{F}_{2^e}^*$, and $\sum_{i=0}^{e} m_i x^i$ is the minimal polynomial of ζ over $\mathbb{F}_2$ (compare [CRW2, Theorem 3.4]). □

Remark 4.6. Every element of $\mathrm{SL}(2, q)$ has bit-length $O(\log q)$ in our generators. For, this is true of all elements of U by (4.3), while $\mathrm{SL}(2, q) = UU^tUU^tU$.

Remark 4.7. $\mathrm{SL}(2, q)$ and $\mathrm{PSL}(2, q)$ are generated by the elements in (4.4) even for those q not included in the preceding presentation. Later we will use these elements as if they were involved in the actual presentations given in [CoMo] or elsewhere.

Remark 4.8. If $d = 1$ then relation (4) can be removed since then (3) states that $h^d = h$ acts as multiplication by ζ^2.

In [GKKL1, Section 3.5.1] it was observed that we can choose ζ such that $k = 1$, $l = 1$ or $k = 2$. Thus, $d \le 2$ for some choice of ζ, k and l. If q is even then $d = 1$. If $q \equiv 3 \pmod 4$ and $k = 2$ we can change ζ to $-\zeta^2$ in order to obtain $k = 1$ and hence $d = 1$. We can also prove that there are choices for ζ, k, l that yield $d = 1$ when $q \equiv 5 \pmod 8$, but we do not know how to obtain such choices in general.

4.4. Unitary groups

We will obtain presentations for 3-dimensional unitary groups by taking the presentations in [GKKL1, Sec. 4.4.2] and deleting the portions that were needed to produce short presentations. We use matrices of the form

$$u = \begin{pmatrix} 1 & \alpha & \beta \\ 0 & 1 & -\bar\alpha \\ 0 & 0 & 1 \end{pmatrix},\ w = \begin{pmatrix} 1 & 0 & \gamma \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ t = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix},\ h = \begin{pmatrix} \bar\zeta^{-1} & 0 & 0 \\ 0 & \bar\zeta/\zeta & 0 \\ 0 & 0 & \zeta \end{pmatrix} \tag{4.9}$$

with $\mathbb{F}_{q^2}^* = \langle\zeta\rangle$, $\alpha, \beta, \gamma \in \mathbb{F}_{q^2}$ arbitrary such that $\beta + \bar\beta = -\alpha\bar\alpha \neq 0$, and $\gamma = -\bar\gamma \neq 0$.

Theorem 4.10. $\mathrm{SU}(3,q)$ *and* $\mathrm{PSU}(3,q)$ *have presentations with* 3 *generators*, 21 *and* 22 *relations*, *respectively*, *and bit-length* $O(\log q)$.

Proof. As in [GKKL1, Sec. 4.4.2], we will assume that $q \neq 2, 3, 5$, and use elements $a = \zeta^k, b = \zeta^l$, where $1 \leq k, l < q^2$, such that

$$a^{2q-1} + b^{2q-1} = 1 \quad \text{and} \quad a^{q+1} + b^{q+1} = 1,$$
$$\mathbb{F}_q = \mathbb{F}_p[a^{q+1}],$$
$$\mathbb{F}_{q^2} = \mathbb{F}_p[a^{2q-1}] \text{ if } q \text{ is odd} \quad \text{while} \quad \mathbb{F}_q = \mathbb{F}_p[a^{2q-1}] \text{ if } q \text{ is even.}$$

Then $u = u^{h^k} u^{h^l} w$ with $w \neq 1$ in the center of the group $\langle u^{\langle h \rangle} \rangle$ of order q^3.

Let $d = \gcd(k,l)$.

If $\gamma \in \mathbb{F}_{q^2}$ write $\gamma' := \gamma^{q+1}$ and $\gamma'' := \gamma^{2q-1}$, and also let $m_\gamma(x)$ denote its minimal polynomial over $\mathbb{F}_p$. If $\delta \in \mathbb{F}_p[\gamma]$, let $f_{\delta;\gamma}(x) \in \mathbb{F}_p[x]$ with $f_{\delta;\gamma}(\gamma) = \delta$ and $\deg f_{\delta;\gamma} < \deg m_\gamma$ (compare [GKKL1, Sec. 4.4.2]).

The required presentation is as follows:

Generators: u, h, t.

Relations:
(*Notation*: w is defined by $u = u^{h^k} u^{h^l} w$.)

(1) $w = w^{h^k} w^{h^l}$, and also $w = w^{h^l} w^{h^k}$ if q is odd.
(2) $w^p = 1$.
(3) $[[w^{m_{a'}(x)}]]_{h^k} = 1$.
(4) $[[w^{f_{\zeta';a'}(x)}]]_{h^k} = w^h$.
(5) $u = u^{h^l} u^{h^k} w_1$.
(6) $[u, w] = [u^{h^k}, w] = 1$.
(7) $u^p = w_2$.
(8) $[[u^{m_{a''}(x)}]]_{h^k} = w_3$.
(9′) $[[u^{f_{\zeta'';a''}(x)}]]_{h^k} = u^h w_4$ if q is odd.
(9″) $([[u^{f_{\alpha;a''}(x)}]]_{h^k})^h [[u^{f_{\beta;a''}(x)}]]_{h^k} = u^{h^2} w_5$ if q is even and ζ'' satisfies $\zeta''^2 = \alpha\zeta'' + \beta$ for $\alpha, \beta \in \mathbb{F}_q$.
(10) $[u, u^h] = w_6$ and $[u^{h^k}, u^h] = w_9$ if q is even.
(11) $t^2 = 1$.
(12) $h^t = h^{-q}$.
(13) $u_i^t = u_{i1} h_i t u_{i2}$ for $1 \leq i \leq 7$, relations due to Hulpke and Seress [HS].

Here, each w_i is a word in $w^{\langle h^d \rangle}$, each u_{ij} is a word in $u^{\langle h^d \rangle}$, and each h_i is a power of h; all are obtained from G, and all depend on the initial choice of u and ζ in (4.9) that are used to obtain the indicated polynomials.

Note that $w^{\langle h^d \rangle}$ and $u^{\langle h^d \rangle}$ involve h^d rather than h so that [GKKL1, Sec. 4.1] can be used together with (1), (5) and the definition of w. If q is even then (2) implies that both relations in (1) hold if one does.

In [GKKL1, Sec. 4.4.2] there were several powers of h that were handled using additional relations that stated that they all commute, gave their actions on u, and dealt with (12) for each of these powers of h. We have discarded these relations.

See [GKKL1, Sec. 4.4.2] for a proof that this is, indeed, a presentation for $\mathrm{SU}(3, q)$. (The relations (13), together with the previous relations, imply all relations (4.1) [HS].) As in [GKKL1], at most one further relation of bit-length $O(\log q)$ is needed to produce a presentation for $\mathrm{PSU}(3, q)$.

Finally, we leave $\mathrm{SU}(3, 2) = 3^{1+2} \rtimes Q_8$ to the reader, $\mathrm{SU}(3, 3)$ is in [CHRR1], and the case $\mathrm{SU}(3, 5)$ follows from the efficient presentation for $\mathrm{PSU}(3, 5)$ in [CHRR2] together with Lemma 2.2. □

Remark 4.11. Let $U := \langle u^{\langle h^d \rangle} \rangle$. Applying (4.3) to both $Z(U) = \langle w^{\langle h^d \rangle} \rangle$ and $U/Z(U)$ shows that all elements of U have bit-length $O(\log q)$, and hence the same holds for $\mathrm{SU}(3, q) = UU^tUU^tU$. More generally, if $c \in \mathrm{SU}(3, q)$ and $U^c \neq U$, then the 2-transitivity of $\mathrm{SU}(3, q)$ on the set of conjugates of U implies that $\mathrm{SU}(3, q) = UU^cUU^cU$, so that all elements of $\mathrm{SU}(3, q)$ have bit-length $O(\log q)$ in $\{u, h, c\}$.

4.5. Suzuki groups

The short presentation in [GKKL1, Section 4.4.3] uses 7 generators and 43 relations for $\mathrm{Sz}(q)$; the most crucial relation is due to Suzuki [Suz, p. 128]. Since we are trying to decrease the numbers of generators and relations, this presentation will not be used for ${}^2F_4(q)$ in Section 7.2. Instead, we use simple modifications to produce the following

Proposition 4.12. $\mathrm{Sz}(q)$ *has a presentation with* 4 *generators*, 29 *relations and bit-length* $O(\log q)$.

Sketch of proof. We will indicate the relations in [GKKL1, Section 4.4.3] that can be deleted. Among the 7 generators, 4 were $h_\star$ with $\star \in \{\zeta, \zeta^\theta, \zeta + 1, \zeta^\theta + 1\}$, where $\mathbb{F}_q^* = \langle \zeta \rangle$, $q = 2^{2k+1}$, and θ is the field automorphism $x \mapsto x^{2^{k+1}}$. All $h_\star$ are powers of $h := h_\zeta$. In [GKKL1, Section 4.4.3] we could not use the corresponding exponents, whereas here we can. Therefore, only h will be needed.

(1) Delete all 6 relations, since all $h_\star$ commute.

(4) Delete. For, (2), (3), (5) and (6) imply that $\langle w^{\langle h \rangle} \rangle \cong \mathbb{F}_q^+$, while (7) specifies the action of h in terms of ζ. Thus, the actions of all $h_\star$ on $\langle w^{\langle h \rangle} \rangle$ are known. In particular, (4) follows from the corresponding relation in $\mathbb{F}_q$.

(13) Delete: this states that h_{ζ^θ} acts on u as h_ζ^θ acts on u.

(15) One of these states that $h_\star^t = h_\star^{-1}$ for $\star = \zeta$. Delete the corresponding relations for the other 3 elements $\star$.

(16) One of these states that $h \in \langle u^{\langle h \rangle}, t \rangle$. Delete the other 3 relations, which state that all $h_\star \in \langle u^{\langle h \rangle}, t \rangle$.

There are now $7 - 3$ generators and $43 - 6 - 1 - 1 - 3 - 3$ relations. □

5. $\mathrm{SL}(3, q)$

The groups $\mathrm{SL}(3, q)$ will reappear more often in the rest of this paper than any other rank 2 groups: we will use $\mathrm{SL}(3, q)$ to obtain all higher-dimensional groups $\mathrm{SL}(n, q)$, and then most of the other higher-dimensional classical groups. Therefore we will be somewhat more explicit with these groups than with the other rank 2 groups (cf. Sections 7 and 8).

Theorem 5.1. SL$(3, q)$ *and* PSL$(3, q)$ *have presentations with* 4 *generators*, 14 *and* 15 *relations*, *respectively*, *and bit-length* $O(\log q)$.

Proof. Let SL$(2, q) \cong L = \langle X \mid R\rangle$, with $X = \{u, t, h\}$ and $\mathbb{F}_q^* = \langle\zeta\rangle$ as in (4.4) and the presentation used for Theorem 4.5 (cf. Remark 4.7).

We view the elements of G as matrices, with L consisting of the matrices $\left(\begin{smallmatrix} * & 0 \\ 0 & 1\end{smallmatrix}\right)$. Consider the group J having the following presentation.

Generators: X and c (corresponding to the permutation matrix $\left(\begin{smallmatrix} 0&1&0 \\ 0&0&1 \\ 1&0&0\end{smallmatrix}\right)$).

Relations:

(1) R.
(2) $c^t = t^2c^2$.
(3) $hh^ch^{c^2} = 1$.
(4) $u^{h^c} = u^{\mathrm{diag}(1,\zeta^{-1})}$, written as a word in X. (The matrix diag$(1, \zeta^{-1})$ is not in L, but it can be viewed as inducing an automorphism of L.)
(5) $[u, u^c] = u^{-tc^2}$ (i.e., $(u^{-1})^{tc^2}$).
(6) $[uu^c, u^{-tc^2}] = 1$.

We will show that $J \cong G$ in general, *but with* (6) *replaced by the* 2 *relations* $[u, u^{-tc^2}] = [u^c, u^{-tc^2}] = 1$ *when* $q = 2, 3, 7, 13, 16$ *or* 19; *and when* $q = 4$, *by those* 2 *and the additional* 3 *relations* $[u, u^{ch^{-1}}] = u^{cth^{-1}}$ *and* $[u, u^{cth^{-1}}] = [u^{ct}, u^{cth^{-1}}] = 1$. (N.B.: There are presentations for the smaller of these groups already in the literature. However, we prefer to have more uniform presentations for later use.)

Matrix calculations show that there is a surjection $J \to G$. There is a subgroup of J we identify with $L = \langle X\rangle$. We also view c as an element of J. We separate the argument into five steps.

1. *Computations in* $\langle t, c\rangle$. The relations $t^4 = 1$ and (2) imply that

$$tct = c^2, \quad t^{-1}c^2t^{-1} = c, \quad t^{-1}c^3t = t^{-1}c^2t^{-1}tct = cc^2, \tag{5.2}$$

and hence

$$t^c = c^{-1}tctt^{-1} = c^{-1}c^2t^{-1} = ct^{-1}, \quad t^{c^2} = (ct^{-1})^c = t^{-1}c. \tag{5.3}$$

It follows that

$$t^2(t^2)^{c^2}(t^2)^c = t^2 \cdot t^{-1}ct^{-1}c \cdot ct^{-1}ct^{-1} = tc(t^{-1}c^2t^{-1})ct^{-1} = tccct^{-1} = c^3. \tag{5.4}$$

2. *The actions of* h, h^c, h^{c^2} *and* c^3 *on* L. We have $h \in L$. We now show that the actions of h^c and h^{c^2} on L are the same as in the target group G.

By (5.2) and (3), $(h^c)^t = h^{t^{-1}c^2} = (h^{-1})^{c^2} = hh^c$. Consequently, conjugating each side of (4) by t, and using the interaction of t with diagonal matrices related to L, gives

$$\begin{aligned}(u^{h^c})^t &= (u^t)^{h^{ct}} = (u^t)^{hh^c} = (u^{th})^{h^c},\\ (u^{\mathrm{diag}(1,\zeta^{-1})})^t &= (u^t)^{\mathrm{diag}(1,\zeta^{-1})^t} = (u^t)^{\mathrm{diag}(\zeta^{-1},1)} = (u^t)^{h\,\mathrm{diag}(1,\zeta^{-1})} = (u^{th})^{\mathrm{diag}(1,\zeta^{-1})}.\end{aligned}$$

(Recall that $h = \mathrm{diag}(\zeta^{-1}, \zeta)$ in (4.4).) Thus, by (4), h^c acts on $L = \langle u, u^{th}\rangle$ as conjugation by $\mathrm{diag}(1, \zeta^{-1})$. In view of the action of h on L, (3) implies that h^{c^2} acts on L as conjugation by $\mathrm{diag}(\zeta^{-1}, 1)$.

If q is even then $t^2 = 1$ and hence $c^3 = 1$ by (5.4). Then c^3 centralizes J; we claim that *this also holds when q is odd.* Here $t^2 = h^{(q-1)/2}$ is in $Z(L)$. Now t^2, $(t^2)^c = (h^c)^{(q-1)/2}$ and $(t^2)^{c^2} = (h^{c^2})^{(q-1)/2}$ act on L as they should: as conjugation by 1, $\mathrm{diag}(1, \zeta^{-1})^{(q-1)/2} = \mathrm{diag}(1, -1)$ and $\mathrm{diag}(1, -1)$, respectively. Then (5.4) implies that $c^3 = t^2(t^2)^{c^2}(t^2)^c$ acts trivially on $L = \langle X\rangle$ and hence on $J = \langle X, c\rangle$, as claimed.

3. *The elements* $e_{ij}(\lambda)$. For all integers m and all $\lambda \in \mathbb{F}_q$, write

$$\begin{aligned}&e_{12}(\zeta^m) := u^{(h^c)^m},\ e_{12}(0) := 1, && e_{21}(\lambda) := e_{12}(-\lambda)^t,\\ &e_{23}(\lambda) := e_{12}(\lambda)^c, && e_{32}(\lambda) := e_{21}(\lambda)^c\\ &e_{31}(\lambda) := e_{12}(\lambda)^{c^2}, && e_{13}(\lambda) := e_{21}(\lambda)^{c^2}.\end{aligned}$$

Then $e_{12}(1) = u$, $e_{23}(1) = u^c$, and $e_{12}(\mathbb{F}_q)$ is an elementary abelian subgroup of L by (4). Then we also have $e_{21}(\mathbb{F}_q) < L$. By (5.3), $c = t^ct \in L^cL$, so that J is generated by the elements $e_{ij}(\lambda)$.

Clearly, $\langle c\rangle$ acts on the set of subgroups $e_{ij}(\mathbb{F}_q)$; in fact $\langle t, c\rangle$ acts as the symmetric group S_3 on subscripts (i.e., as the Weyl group of G). For example, by (5.2), $e_{23}(\mathbb{F}_q)^t = e_{12}(\mathbb{F}_q)^{ct} = e_{12}(\mathbb{F}_q)^{t^{-1}c^2} = e_{13}(\mathbb{F}_q)$; and (as we have seen) t^2 acts correctly on $L^{c^{-1}} = L^{c^2}$ (i.e., as it does in the target group G) and hence also on $e_{12}(\mathbb{F}_q)^{c^2} = e_{31}(\mathbb{F}_q)$. Similarly, t acts correctly on each $e_{ij}(\mathbb{F}_q)$.

4. *The Steinberg relations* (see [GKKL1, Section 5.1 or 5.2]). The relations $e_{ij}(\lambda)e_{ij}(\mu) = e_{ij}(\lambda + \mu)$ follow from the corresponding relations in L (with $\{i, j\} = \{1, 2\}$) by conjugating with t and c. We will deduce the remaining Steinberg relations from (5) and (6) by conjugating with t, c, h and h^c; we have seen that these act on the set of subgroups $e_{ij}(\mathbb{F}_q)$ as they do in G. We will use (5) and (6) to prove that, for all $\lambda, \mu \in \mathbb{F}_q$,

$$[e_{12}(\lambda), e_{23}(\mu)] = e_{13}(\lambda\mu), \tag{5.5}$$

$$[e_{12}(\lambda), e_{13}(\mu)] = 1, \tag{5.6}$$

$$[e_{13}(\lambda), e_{23}(\mu)] = 1. \tag{5.7}$$

By (6), $[e_{12}e_{23}, e_{13}] = 1$. Conjugating by $h^i(h^c)^j$ gives $[e_{12}(\zeta^{2i-j})e_{23}(\zeta^{2j-i}), e_{13}(\zeta^{i+j})] = 1$. Let $i = -j$ to see that e_{13} commutes with all $e_{12}(\lambda)e_{23}(\lambda^{-1})$, and hence with all $e_{23}(\mu^{-1})e_{12}(\mu)$, where we temporarily restrict the letters λ, μ to $\mathbb{F}_q^{*3}$. Conjugating the relation $[e_{12}, e_{23}] = e_{13}$ by $h^i(h^c)^j$ gives

$$[e_{12}(\zeta^{2i-j}), e_{23}(\zeta^{2j-i})] = e_{13}(\zeta^{i+j}). \tag{5.8}$$

This does not cover all relations (5.5) since $\det\left(\begin{smallmatrix} 2 & -1 \\ -1 & 2 \end{smallmatrix}\right) = 3$, but it does imply that $e_{13}(\lambda\mu) = e_{12}(-\lambda)e_{23}(-\mu)e_{12}(\lambda)e_{23}(\mu)$. If $\nu := -(\lambda + \mu) \in \mathbb{F}_q^{*3}$, then e_{13} commutes with $e_{13}(\lambda\mu)$ and hence with

$$\begin{aligned} &e_{23}(\lambda^{-1})e_{12}(\lambda) \cdot e_{12}(-\lambda)e_{23}(\mu^{-1})e_{12}(\lambda)e_{23}(-\mu^{-1}) \cdot e_{23}(\mu^{-1})e_{12}(\mu) \cdot e_{12}(\nu)e_{23}(\nu^{-1}) \\ &\quad = e_{23}(\lambda^{-1} + \mu^{-1})e_{12}(\lambda + \mu) \cdot e_{12}(\nu)e_{23}(\nu^{-1}) = e_{23}(\lambda^{-1} + \mu^{-1} - (\lambda + \mu)^{-1}). \end{aligned}$$

Let A be the additive group generated by all $\lambda^{-1} + \mu^{-1} - (\lambda + \mu)^{-1}$ with $\lambda, \mu, \lambda + \mu \in \mathbb{F}_q^{*3}$. If we replace λ, μ by $\theta^3\lambda, \theta^3\mu$ we see that $\mathbb{F}_q^3 A = A$.

Assume that $q \neq 2, 3, 4, 7, 13, 16, 19$.

Claim: $A \neq 0$. We must show that there exist $\lambda, \mu \in \mathbb{F}_q^{*3}$ such that $\lambda + \mu \in \mathbb{F}_q^{*3}$ and $\lambda^{-1} + \mu^{-1} - (\lambda + \mu)^{-1} \neq 0$, i.e., λ/μ is not a root of $z^2 + z + 1 = 0$. Thus, we need a solution to $x^3 + y^3 = 1$ with $x, y, x^6 + x^3 + 1 \neq 0$. This is obvious if $q > 3$ and $q \not\equiv 1 \pmod 3$. If $q \equiv 1 \pmod 3$ then the equation $x^3 + y^3 = 1$ has at least $q + 1 - 2\sqrt{q} - 3$ solutions in $\mathbb{F}_q$. (This follows from [We]. A more precise and elementary count is given in [Hal, p. 180].) Of these, we must exclude at most 6 solutions (x, y) with $xy = 0$ and $6 \cdot 3$ with $x^6 + x^3 + 1 = 0$, hence at most 24 solutions. It is now easy to check that we only need to exclude the stated values of q.

Thus, $\mathbb{F}_q^3 A = A \neq 0$. It follows that $A = \mathbb{F}_q$ since $\mathbb{F}_q$ is additively generated by $\mathbb{F}_q^3$ (recall that $q \neq 4$). Thus, e_{13} commutes with $e_{23}(\mathbb{F}_q)$, and similarly with $e_{12}(\mathbb{F}_q)$. Conjugating by all $h^i(h^c)^j$ yields (5.6) and (5.7).

For all $\mu \in F$, it follows that $[e_{12}(1), e_{23}(\mu)] = e_{13}(\mu)$ by using (5.7), (5.8), the standard identity

$$[x, ab] = [x, b][x, a]^b, \tag{5.9}$$

and the fact that $\mathbb{F}_q^3$ generates $\mathbb{F}_q$ additively. Conjugating by all $h^i(h^c)^j$ yields (5.5).

Finally, if q is 2, 3, 4, 7, 13, 16 or 19 then, proceeding as above, we again find that the assumed additional relations imply (5.5)–(5.7).

5. *Completion.* The Steinberg relations imply that J is a homomorphic image of $G = \mathrm{SL}(3, q)$, and hence $J \cong G$ [GLS, pp. 312–313]. By Theorem 4.5 and Remark 4.6, our presentation has the required bit-length. This presentation uses $|R| + 5$ relations if $q \neq 2, 3, 4, 7, 13, 16, 19$. In each of these excluded cases, we added 1 or 4 further relations; while by (3.19) and Theorem 4.5 there is a presentation for $\mathrm{SL}(2, q)$ using 2 generators and 2 or 3 relations if $q \neq 16$, or 3 generators and 5 relations when $q = 16$. Consequently, we still obtain the required numbers of generators and relations.

In order to obtain a presentation for $\mathrm{PSL}(3, q)$ when $m := (q - 1)/3$ is an integer, add the relation $h^m(h^{2m})^c = 1$. □

Remark 5.10. By Remark 4.6 and the Bruhat decomposition [GLS, Theorem 2.3.5], *each element of* $\mathrm{SL}(3, q)$ *has bit-length* $O(\log q)$ *in our generators.*

6. $\mathrm{SL}(n, q)$

We now turn to the general case of Theorem B for the groups $\mathrm{PSL}(n, q)$, using a variation on the approach in Section 5.

Theorem 6.1. *Let $n \geq 4$.*

(i) $\mathrm{SL}(n, q)$ *has a presentation with* 6 *generators and* 25 *relations.*
(ii) $\mathrm{SL}(4, q)$ *has a presentation with* 5 *generators and* 20 *relations.*
(iii) $\mathrm{SL}(n, q)$, $5 \leq n \leq 8$, *has a presentation with* 5 *generators and* 21 *relations.*

Each of these presentations has bit-length $O(\log n + \log q)$. At most one further relation of bit-length $O(\log n + \log q)$ is needed to obtain a presentation for $\mathrm{PSL}(n, q)$.

Proof. We use two presentations:

- The presentation $\langle X \mid R\rangle$ for $F = \mathrm{SL}(3, q)$ in Theorem 5.1. We view F as the matrices $\left(\begin{smallmatrix} * & 0 \\ 0 & I \end{smallmatrix}\right)$ in $G = \mathrm{SL}(n, q)$, and only write the upper left 3×3 block.
- The presentation $\langle Y \mid S\rangle$ for $T = A_n$ in Theorem 3.40, where T acts in the standard manner on $\{1, \ldots, n\}$; here X and Y are disjoint. We view T as permutation matrices.

We will also use the subgroup $L = \mathrm{SL}(2, q)$ of F consisting of matrices in the upper left 2×2 block, together with the following elements:

- $c = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $f = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \in F$,
- $a \in L$ such that $\langle a, a^f\rangle = L$,
- $(3, 2, 1)$, $(1, 3)(2, 4) \in T$,
- $\tau = (1, 2)(3, 4)$ and σ in T interchanging 1 and 2 and generating the set-stabilizer $T_{\{1,2\}}$ of $\{1, 2\}$ in T.

Bit-length: c, f and a have bit-length $O(\log q)$ using Remark 4.6. We may assume that σ is a cycle of length $n - 2$ or $n - 3$ on $\{3, \ldots, n\}$. The specified permutations can be written as words in Y of bit-length $O(\log n)$ (by Remark 3.37).

We will show that G is isomorphic to the group J having the following presentation.

Generators: X, Y.

Relations:

(1) R.
(2) S.
(3) $c = (3, 2, 1)$.
(4) $a^\sigma = a^f$.
(5) $a^\tau = a^f$.
(6) $(a^f)^\sigma = a$.
(7) $[a, a^{(1,3)(2,4)}] = 1$.
(8) $[a^f, a^{(1,3)(2,4)}] = 1$ (needed only when n is 4 or 5).

As usual, there is a surjection $\pi : J \to G$, and J has subgroups we will identify with $F = \langle X\rangle$ and $T = \langle Y\rangle$. Since τ has order 2, (5) implies that $(a^f)^\tau = a$. Hence, by (4)–(6), $\langle\sigma, \tau\rangle$ normalizes $\langle a, a^f\rangle = L$, inducing the same automorphism group as $\langle f\rangle$ does on L. In particular, elements of $\langle\sigma, \tau\rangle$ that fix 1 and 2 must centralize L, while elements interchanging 1 and 2 act as f.

It follows that $|L^T| \le \binom{n}{2}$; as usual, we use π to obtain equality. Then L^T can be identified with the set of all 2-sets of $\{1, \dots, n\}$. Its subset $L^{\langle c\rangle}$ (where $c \in T$ by (3)) consists of 3 subgroups corresponding to the 2-sets in $\{1, 2, 3\}$. Consequently, any two distinct members of L^T can be conjugated by a single element of T to one of the pairs $L, L^{(3,2,1)}$ or $L, L^{(1,3)(2,4)}$. Here $\langle L, L^{(3,2,1)}\rangle = \langle L, L^c\rangle = F$.

We claim that $[L, L^{(1,3)(2,4)}] = 1$. By (7) and (8), since $\langle a, a^f\rangle = L$ this is clear if n is 4 or 5. Assume that $n \ge 6$. By our comment about elements of $\langle\sigma, \tau\rangle$, we have $a^{(1,2)(5,6)} = a^f$ and $a^{(3,4)(5,6)} = a$. By (7),

$$\begin{aligned}
1 &= [a, a^{(1,3)(2,4)}]^{(1,2)(5,6)} = [a^f, (a^{(3,4)(5,6)})^{(1,3)(2,4)}] = [a^f, a^{(1,3)(2,4)}],\\
1 &= [a, a^{(1,3)(2,4)}]^{(1,2)(3,4)} = [a^f, (a^f)^{(1,3)(2,4)}], \qquad (6.2)\\
1 &= [a^f, a^{(1,3)(2,4)}]^{(1,2)(3,4)} = [a, (a^f)^{(1,3)(2,4)}],
\end{aligned}$$

where the first equations explain the comment in (8).

Thus, any two distinct members of L^T either generate a conjugate of $F = \mathrm{SL}(3, q)$ or commute. Consequently, $N := \langle L^T\rangle \cong G$ by the Steinberg presentation [GKKL1, Sections 5.1 or 5.2]. Moreover, $N \trianglelefteq J$, and J/N is a homomorphic image of $\langle Y\rangle \cong A_n$ in which (3, 2, 1) is sent to 1 (by (3)). Thus, $J/N = 1$.

The bit-length follows easily from those of $\langle X \mid R\rangle$ and $\langle Y \mid S\rangle$.

Fine-tuning: We still need to count the number of relations. If $n \ge 6$ then we used $4 + 3$ generators and $14 + 7 + 5$ relations by Theorems 3.40 and 5.1. However, *we can remove a generator and a relation as follows*. In Theorem 5.1 we used a generator "c" (corresponding to the permutation matrix acting as (3, 2, 1)); we will delete that generator and relation (3). Start with $T = \langle Y \mid S\rangle$. Let c' denote a word in Y representing the permutation (3, 2, 1) in T (cf. Remark 3.37). Replace c by c' throughout the presentation $\langle X \mid R\rangle$ used in Theorem 5.1, and insert all of the resulting relations into the above presentation for G. Then Theorem 5.1 shows that $\langle (X\backslash\{c\}) \cup \{c'\}\rangle$ is F; and when c' is viewed as an element of F it is precisely the permutation matrix c. We have now deleted the generator c from X, and we can also delete the above relation (3) since it already holds in F.

Additional fine-tuning when $4 \le n \le 8$: In these cases there is a presentation $\langle Y \mid S\rangle$ for $T = A_n$ with 2 generators and 3 relations [ThS, CRKMW, CHRR1]. Now we obtain $4 + 2$ generators and $14 + 3 + 5$ relations. Once again we can use c' in order to remove 1 generator and 1 relation from Theorem 5.1. This produces the numbers in (iii) when n is 6, 7 or 8. When n is 4 or 5, we have an extra relation (8). However, if $n = 4$ we can delete σ entirely (since $|T_{\{1,2\}}| = 2$), therefore also deleting 2 further relations (4) and (6) in order to obtain (ii); and if $n = 5$ we can choose $\sigma = (1, 2)(3, 5)$ of order 2 (since

Table 3. G, L_{α_i}

G	L_{α_1}	L_{α_2}
$\mathrm{Spin}_5(q) \cong \mathrm{Sp}(4,q)$	$\mathrm{SL}(2,q)$	$\mathrm{Spin}_3(q) \cong \mathrm{SL}(2,q)$
$\Omega(5,q) \cong \mathrm{PSp}(4,q)$	$\mathrm{SL}(2,q)$	$\Omega(3,q) \cong \mathrm{PSL}(2,q)$
$\mathrm{Spin}_6^-(q) \cong \mathrm{SU}(4,q)$	$\mathrm{SL}(2,q)$	$\mathrm{Spin}_4^-(q) \cong \mathrm{SL}(2,q^2)$
$\Omega^-(6,q)$	$\mathrm{SL}(2,q)$	$\Omega^-(4,q) \cong \mathrm{PSL}(2,q^2)$

$T_{\{1,2\}} \cong S_3$), and therefore delete (6) in order to obtain (iii). (N.B.: When $6 \le n \le 8$ this fine-tuning will be used in Section 9 for the exceptional groups $E_n(q)$.)

The group $\mathrm{PSL}(n,q)$: Finally, we need to add one further relation in order to obtain $\mathrm{PSL}(n,q)$. Let h_{ij} be the matrix with ζ^{-1} and ζ in positions i and j, respectively, and 1 elsewhere. Then $h_{12} \in L$ is one of the generators used in Theorem 4.5, and h_{1n} and h_{2n} have bit-length $O(\log n + \log q)$ using Remark 3.37. Let $m = (q-1)/(d, q-1)$. If n is odd, again use Remark 3.37 in order to obtain the $(n-2)$-cycle $(2,\dots,n-1)$, and then the additional relation $h_{1n}^m(h_{2n}^m(2,\dots,n-1))^{n-2} = 1$ produces $\mathrm{PSL}(n,q)$ with the required bit-length. Even n is handled similarly. □

7. Remaining rank 2 groups

In this section we will provide presentations required in Theorem B for most of the rank 2 groups of Lie type: $\mathrm{Sp}(4,q)$, $\Omega(5,q) \cong \mathrm{PSp}(4,q)$, $\mathrm{SU}(4,q)$, $\Omega^-(6,q)$, $G_2(q)$, ${}^3D_4(q)$ and ${}^2F_4(q)$. The groups $\mathrm{SU}(5,q)$ are handled in an entirely different manner in Section 8.

7.1. $\mathrm{Sp}(4,q)$*,* $\Omega(5,q)$*,* $\mathrm{SU}(4,q)$*,* $\Omega^-(6,q)$*,* $G_2(q)$ *and* ${}^3D_4(q)$

Here the Weyl group is dihedral of order $2m = 8$ or 12.

Theorem 7.1. (i) $\mathrm{Sp}(4,q)$ *and* $\mathrm{PSp}(4,q) \cong \Omega(5,q)$ *have presentations with* 5 *generators and* 27 *relations if q is odd, and with* 6 *generators and* 20 *relations if q is even.*

(ii) $\mathrm{SU}(4,q)$ *and* $\Omega^-(6,q)$ *have presentations with* 5 *generators and* 27 *relations. At most one further relation is needed to obtain a presentation for* $\mathrm{P}\Omega^-(6,q)$.

(iii) $G_2(q)$ *and* ${}^3D_4(q)$ *have presentations with* 6 *generators and* 31 *relations.*

Each of the presentations in (i)–(iii) *has bit-length* $O(\log q)$.

Proof. (i),(ii) The root system Φ of $G = \mathrm{Sp}(4,q)$, $\mathrm{PSp}(4,q)$, $\mathrm{SU}(4,q)$ or $\Omega^-(6,q)$ has 8 roots, half of them long and half short. Let $\Pi = \{\alpha_1, \alpha_2\}$ be a set of fundamental roots with α_1 long; the corresponding rank 1 groups L_{α_i} are in Table 3.

We use the presentation $\langle X_i \mid R_i \rangle$ for L_{α_i} in Theorem 4.5 (cf. Remark 4.7), with

$$X_i = \{u_{\alpha_i}, r_i, h_i\}, \quad i = 1, 2. \tag{7.2}$$

(Here we use r_i instead of t_i in order to approximate standard Lie notation. We assume that X_1 and X_2 are disjoint.) The action of h_i on u_{α_i} is given in R_i, and $U_{\alpha_i} = \langle u_{\alpha_i}^{\langle h_i \rangle} \rangle$ has order q (or q^2 for short α_2 in the cases related to $\Omega^-(6, q)$). The root groups U_α, $\alpha \in \Phi$, will be built into our presentation. Standard notation [Cart, p. 46] labels the root system as follows, with $w = r_1 r_2$:

$$\begin{aligned}\Phi &= \{w^{-n}(\alpha_1), w^{-n}(\alpha_2) \mid 0 \le n < 4\} \\ &= \{\pm\alpha_1, \pm\alpha_2, \pm(\alpha_1 + \alpha_2), \pm(\alpha_1 + 2\alpha_2)\}.\end{aligned} \tag{7.3}$$

If we exclude the case $G = \mathrm{Sp}(4, q)$ with q even, we can save 1 generator and 1 relation by using the fact that $U_{\alpha_1} = [U_{\alpha_1+\alpha_2}, U_{-\alpha_2}] = [U_{\alpha_2}^w, U_{\alpha_2}^{r_2}]$. (See [GLS, p. 47]; in the excluded case, $[U_{\alpha_2}^w, U_{\alpha_2}^{r_2}] = 1$.) In all cases we will show that G is isomorphic to the group J having the following presentation.

Generators: $(X_1 \backslash \{u_{\alpha_1}\}) \cup X_2$ except if $G = \mathrm{Sp}(4, q)$ with q even.
$X_1 \cup X_2$ if $G = \mathrm{Sp}(4, q)$ with q even.

Relations:
(*Notation*: $w := r_1 r_2$, and $u_{\alpha_1} := [u_{\alpha_2}^w, u_{\alpha_2}^{r_2}]$ except if $G = \mathrm{Sp}(4, q)$ with q even. In all cases R_1 can then be used for the elements u_{α_1}, r_1, h_1.)

(1) $R_1 \cup R_2$.
(2) $h_1^{r_2}, h_2^{r_1}$ and w^4 written as words in $\{h_1, h_2\}$ (obtained from G).
(3) $(u_{\alpha_i})^{h_j}$ written as a word in $u_{\alpha_i}^{\langle h_i \rangle}$ (obtained from G) whenever $\{i, j\} = \{1, 2\}$.
(4) $h_1 h_2 = h_2 h_1$.
(5) Let $H := \langle h_1, h_2 \rangle$. If $\alpha = w^{-n}(\alpha_i)$ with $1 \le i \le 2$ and $0 \le n < 4$ in (7.3), write $u_\alpha := (u_{\alpha_i})^{w^n}$.
   (a) $[u_{\alpha_1}, u_{\alpha_1+2\alpha_2}] = [u_{\alpha_1}, u_{\alpha_1+\alpha_2}] = 1$.
   (b) $[u_{\alpha_1}, u_{\alpha_2}]$ written as a product of a word in $u_{\alpha_1+2\alpha_2}^H$ and a word in $u_{\alpha_1+\alpha_2}^H$ (obtained from G).
   (c) If $G = \mathrm{Sp}(4, q)$ with q even: $[u_{\alpha_2}, u_{\alpha_1+\alpha_2}] = 1$.

The words mentioned in these relations are based on G. The words in (2) have the form $h_1^k h_2^l$, $0 \le k, l < q^2$; and those in (5) also have bit-length $O(\log q)$ by (4.3). Thus, this presentation has bit-length $O(\log q)$. If $G = \mathrm{Sp}(4, q)$ with q even, then using Theorem 4.5 we have $3 + 3$ generators and $5 + 5 + 3 + 2 + 1 + 4$ relations. In all other cases there are 5 generators and $9 + 9 + 3 + 2 + 1 + 3 = 27$ relations.

Remark 7.4. We digress to note that, in the proof of Theorem 9.1, we will need to have generators for one of the groups $L_{\alpha_1}, L_{\alpha_2}$ for use in a presentation of another subgroup of a target group. When we need L_{α_1} for this purpose, we will not be able to determine u_{α_1} from conjugates of u_{α_2} as in the above presentation. *In that event we will need to use an additional generator u_{α_1} and an additional relation $u_{\alpha_1} = [u_{\alpha_2}^w, u_{\alpha_2}^{r_2}]$ for the presentation of our rank* 2 *group*, *obtaining a presentation with* 6 *generators and* 28 *relations.*

We now continue our proof. Since we have chosen the groups L_{α_i} as in Table 3, it follows that there is a surjection $\pi : J \to G$. By (1) there is a subgroup $L_{\alpha_i} \cong \mathrm{SL}(2, q)$, $\mathrm{PSL}(2, q)$, $\mathrm{SL}(2, q^2)$ or $\mathrm{PSL}(2, q^2)$ of J that we can identify with $\langle X_i \rangle$.

By (1) and (2), $H \trianglelefteq N := \langle H, r_1, r_2\rangle$. Since $r_i^2 \in H$ by (1), we also have $W := N/H \cong D_8$.

Define $U_{\alpha_i} := \langle u_{\alpha_i}^{\langle h_i\rangle}\rangle$ for $i = 1, 2$.

By (1), (3) and (4), H normalizes each U_{α_i}, and hence $H = H^{w^n}$ normalizes each $U_{\alpha_i}^{w^n}$ for $0 \le n < 4$.

Since $U_{\alpha_i}^{r_i} = U_{-\alpha_i}$, the group N acts in the natural manner on the 8 root groups $U_\alpha = \langle u_\alpha^H\rangle$, $\alpha \in \Phi$, labeled as in (7.3).

The Chevalley commutator relations [GLS, p. 47] have the form

$$[y_\alpha, y_\beta] = \prod_\gamma v_\gamma \quad \text{with } 1 \ne y_\alpha \in U_\alpha,\ 1 \ne y_\beta \in U_\beta,\ v_\gamma \in U_\gamma, \tag{7.5}$$

for roots α and $\beta \ne \pm\alpha$, where γ runs through all positive integral combinations of α and β. Clearly, N acts by conjugation on the set of all such relations. There are 4 orbits of W on the unordered pairs $\{\alpha, \beta\}$, $\beta \ne \pm\alpha$. Instances of the corresponding relations are in (5); and if (5c) does not apply then the definition of u_{α_1} gives the fourth relation

(5c′) $[u_{\alpha_1+\alpha_2}, u_{-\alpha_2}] = u_{\alpha_1}$.

If $[y_\alpha, y_\beta] = 1$ with $\{y_\alpha, y_\beta\} = \{u_{\alpha_1}, u_{\alpha_1+2\alpha_2}\}$ or $\{u_{\alpha_1}, u_{\alpha_1+\alpha_2}\}$ in (5a), then $[U_\alpha, U_\beta] = [\langle y_\alpha^H\rangle, \langle y_\beta^H\rangle] = 1$. (Here y_α^H corresponds to the set of elements of a field $\mathbb{F}$ containing $\mathbb{F}^{*2}$, and hence generates $\mathbb{F}$ under addition, as in Section 5, Step 4.)

If $[y_\alpha, y_\beta] \ne 1$ for a pair $\{y_\alpha, y_\beta\}$ in (5b,c,c′) then, by [GKKL1, Lemma 5.4], we obtain most of the relations (7.5) by conjugating the ones in (5b,c,c′) by elements of H. Use of (5.9) yields all remaining relations, as in Section 5, Step 4.

At this point we have verified the Steinberg relations (cf. Section 2). Thus, J is a homomorphic image of the simply connected cover of G, which is $\mathrm{Sp}(4, q)$ or $\mathrm{Spin}_6^-(q) \cong \mathrm{SU}(4, q)$. By Table 3, $J \cong \mathrm{Sp}(4, q)$, $\Omega(5, q)$, $\mathrm{SU}(4, q)$ or $\Omega^-(6, q)$. If $J \cong \Omega^-(6, q)$ then we need at most one further relation with bit-length $O(\log q)$ in order to kill $Z(J)$ as in [GKKL1, p. 749].

(iii) These groups are handled in a manner similar to (i) and (ii), replacing the number 4 by 6 in (2) in order to obtain the Weyl group D_{12}. We sketch this very briefly since these groups do not arise in higher rank settings.

Once again we have fundamental subgroups $L_{\alpha_1} \cong \mathrm{SL}(2, q)$ and $L_{\alpha_2} \cong \mathrm{SL}(2, q)$ or $\mathrm{SL}(2, q^3)$, where the latter occurs with α_2 a short root for ${}^3D_4(q)$. We label the roots as in [Cart, p. 46].

We use versions of relations (1)–(5) given above, with w^6 replacing w^4 in (2). As before, $W := N/H$ acts in the natural manner on the root system Φ of type G_2, and on the associated root groups U_α—which are defined as before. This time W has 7 orbits on pairs $\{\alpha, \beta\}$, $\beta \ne \pm\alpha$, all of which are represented in (5). The long root groups $U_{\alpha_1}^{w^n}$, $0 \le n < 6$, satisfy the relations in Theorem 5.1 with $c := w^2 h_0$ for a word h_0 in $\{h_1, h_2\}$ obtained from G. (Relations (2) and (3) occur in N, while (4) follows from the known action of H on U_{α_1}.) This takes care of (7.5) when α and β are long. The remainder of the proof imitates (i) and (ii). There are $3 + 3$ generators and $9 + 9 + 3 + 2 + 1 + 7 = 31$ relations. □

Remark 7.6. *Bit-lengths*: *Every element of each of the above groups G has bit-length $O(\log q)$ in terms of the generators.* For, this is true of all elements of L_{α_i}, hence of their conjugates, hence of all root groups and all elements of H. Now the Bruhat decomposition [GLS, Theorem 2.3.5] implies the assertion.

Similar observations hold for ${}^2F_4(q)$ in the next section, and more generally for all groups of bounded rank, as in [GKKL1, Proposition 5.6] (compare Remark 5.10).

7.2. ${}^2F_4(q)$

Here $q = 2^{2e+1} > 2$. There is no root system in the classical sense, but there are 16 "root groups" U_i, $1 \le i \le 16$. There are rank 1 groups $L_1 = \mathrm{Sz}(q)$ and $L_2 = \mathrm{SL}(2, q)$, and we use the presentation $\langle X_i \mid R_i \rangle$ for L_i in Proposition 4.12 or Theorem 4.5. If $i = 2$ then (7.2) holds, and $U_2 := \langle u_2^{\langle h_2 \rangle} \rangle$ is elementary abelian of order q. On the other hand, X_1 has size 4 and contains elements u_1, r_1, h_1 behaving essentially as before, except that this time $U_1 = \langle u_1^{\langle h_1 \rangle} \rangle$ is nonabelian of order q^2 with $Z(U_1) = \langle (u_1^2)^{\langle h_1 \rangle} \rangle$ of order q.

In order to save one generator and one relation, we will use the commutator relation $[U_1, U_3] = U_2$ [GLS, p. 48].

Proposition 7.7. *${}^2F_4(q)$ has a presentation with 6 generators, 49 relations and bit-length $O(\log q)$.*

Proof. We use the following presentation.

Generators: $X_1 \cup X_2 \backslash \{u_2\}$.

Relations:
(*Notation*: $w := r_1 r_2$ and $u_2 := [u_1, u_1^w]$. Then R_2 can be used for the elements u_2, r_2, h_2. Let $u_{i+n} := u_i^{w^n}$ for $i = 1, 2$ and $1 \le n < 8$.)

(1) $R_1 \cup R_2$.
(2) $w^8 = 1$.
(3) $h_1^{r_2}$ and $h_2^{r_1}$ written as words in $\{h_1, h_2\}$ (obtained from G).
(4) $u_i^{h_j}$ written as a word in $u_i^{\langle h_i \rangle}$ (obtained from G) whenever $\{i, j\} = \{1, 2\}$.
(5) $h_1 h_2 = h_2 h_1$.
(6) $[u_i, u_j]$ written as a product of words (obtained from G) in $u_k^{\langle h_1, h_2 \rangle}$, $i < k < j$, for the pairs (i, j) with $i = 1$, $j \in \{2, 4, 5, 6, 7, 8\}$, or $i = 2$, $j \in \{4, 6, 8\}$.

There is a surjection from the presented group J onto ${}^2F_4(q)$. As usual there are subgroups L_i of J we can identify with $\langle X_i \rangle$, $i = 1, 2$. Also, $\langle r_1, r_2 \rangle$ is dihedral of order 16 by (1) and (2), and normalizes $H := \langle h_1, h_2 \rangle$ by (1) and (3). It then follows from (1), (4) and (5) that H normalizes each subgroup $U_{i+n} := U_i^{w^n}$ for $i = 1, 2$ and $0 \le n < 8$.

As in the case of the other rank 2 groups, the known action of $h_i^{w^n}$ on U_{i+n} obtained from (4) and (5) allows us to deduce from (6) an additional relation analogous to (7.5) for each pair of nontrivial cosets of the form $y_i \Phi(U_i)$, $y_j \Phi(U_j)$, for i, j as in (6) and $y_i \in U_i$, $y_j \in U_j$ (compare [GKKL1, Lemma 5.4]). By using (5.9), we see that these conjugates of the relations (6) imply all analogues of (7.5) for these i, j. It is now easy to see that we have all relations required for a presentation of G ([Gri, p. 412], [BGKLP, p. 105] and [GLS, p. 48] give the 10 formulas mimicked in (6) and in the definition $u_2 = [u_1, u_3]$).

Proposition 4.12 and Theorem 4.5 imply that we have 6 *generators and* $29 + 5 + 1 + 2 + 2 + 1 + 9 = 49$ *relations.* □

8. Unitary groups

Since the commutator relations for the odd-dimensional unitary groups are especially complicated (cf. [GLS, Theorem 2.4.5(c)]), we will deal with unitary groups separately. In fact, when combined with Theorems 3.40 and 4.10, presentations in [BeS] allow us to use surprisingly few generators and relations—significantly fewer than seem possible using the Curtis–Steinberg–Tits presentation.

8.1. Phan-style presentations

We will use the presentation for $G = \mathrm{SU}(n, q)$, $n \geq 4$, given in [BeS], based on one in [Ph]. In [BeS], subgroups $U_1, U_2 \cong \mathrm{SU}(2, q)$ of $\mathrm{SU}(3, q)$ are called a *standard pair* if U_1 and U_2 are the respective stabilizers in $\mathrm{SU}(3, q)$ of perpendicular nonsingular vectors.

Using an orthonormal basis, it is easy to see that G has subgroups $U_i \cong \mathrm{SU}(2, q)$, $1 \leq i \leq n-1$, and $U_{i,j}$, $1 \leq i < j \leq n-1$, satisfying the following conditions.

(P1) If $|j - i| > 1$ then $U_{i,j}$ is a central product of U_i and U_j.
(P2) For $1 \leq i < n-1$, $U_{i,i+1} \cong \mathrm{SU}(3, q)$, and U_i, U_{i+1} is a standard pair in $U_{i,i+1}$.
(P3) $G = \langle U_{i,j} \mid 1 \leq i < j \leq n-1\rangle$.

We will use the following analogue of the Curtis–Steinberg–Tits presentation.

Theorem 8.1 ([Ph, BeS]). *If* (P1)–(P3) *hold in a group G, then G is isomorphic to a factor group of* $\mathrm{SU}(n, q)$ *in each of the following situations.*

(a) $q > 3$ *and* $n \geq 4$.
(b) $q = 2$ *or* 3, $n \geq 5$ *and the following hold*:
  (1) $\langle U_{i,i+1}, U_{i+1,i+2}\rangle \cong \mathrm{SU}(4, q)$ *whenever* $1 \leq i \leq n-3$;
  (2) *if* $q = 2$ *then*
    (i) $[U_i, U_{j,j+1}] = 1$ *whenever* $1 \leq i \leq n-1$, $1 \leq j \leq n-2$ *and* $i \neq j-1, j, j+1, j+2$;
    (ii) $[U_{i,i+1}, U_{j,j+1}] = 1$ *whenever* $1 \leq i \leq n-2$, $1 \leq j \leq n-2$ *and* $i \neq j-2, j-1, j, j+1, j+2$.

In [BeS] it is remarked that (P1)–(P3) do not provide a presentation for $\mathrm{SU}(n, 2)$. It is also noted that a standard pair in $\mathrm{SU}(3, 2)$ does not generate that group.

8.2. Some specific presentations

When $q = 2$ or 3 we will use presentations of some small groups [Br, Hav] in order to decrease our numbers of generators and relations:

$$\begin{aligned}
\mathrm{SU}(4,3) = \langle x, y \mid x^5y^7 &= x^{-1}y^3x^{-1}y^{-1}x^{-1}y^{-3}x^{-1}y\\
&= yxy^{-1}xy^{-1}x^{-1}y^{-1}x^{-1}yxy^3x^{-1}\\
&= x^2y^{-1}x^{-1}y^{-1}x^{-2}y^{-1}x^{-2}y^{-1}x^2y^{-1}xyxy^{-2} = 1\rangle\\
\mathrm{SU}(6,2) = \langle x, y \mid x^2 = y^7 &= (xy^3)^{11} = [x,y]^2\\
&= [x,y^2]^3 = [x,y^3]^3 = (xy)^{33} = (xyxy^2xy^3xy^{-3})^{21} = 1\rangle\\
\mathrm{SU}(5,2) = \langle x, y \mid x^2 = y^5 &= (xy)^{11}\\
&= [x,y]^3 = [x,y^2]^3 = [x,yxy]^3 = [x,yxy^2]^3 = 1\rangle\\
\mathrm{SU}(4,2) = \langle x, y \mid x^5 &= (xy^2)^2 = x^2(y^{-1}x)^2yxy^{-2} = 1\rangle.
\end{aligned}$$

8.3. Bounded presentations

In this section we will prove the following

Theorem 8.2. *Let $n \geq 4$.*

(i) $\mathrm{SU}(n,q)$ *has a presentation with* 5 *generators and* 32 *relations.*
(ii) $\mathrm{SU}(n,q)$, $4 \leq n \leq 8$, *has a presentation with* 5 *generators and* 30 *relations.*

Each of these presentations has bit-length $O(\log n + \log q)$. At most one further relation of bit-length $O(\log n + \log q)$ is needed to obtain a presentation for $\mathrm{PSU}(n,q)$ *or* $\Omega^-(6,q)$.

Proof. We can ignore the small field cases appearing in the preceding section. Let

- $F := \mathrm{SU}(m,q)$, with $m = 3, 4$ or 6 and $m < n$,
- $T := A_n$, acting in the standard manner on $\{1, \ldots, n\}$.

We view both of these groups as lying in $G = \mathrm{SU}(n,q)$, using an orthonormal basis of the underlying vector space: F consists of the matrices $\left(\begin{smallmatrix} * & 0\\ 0 & I\end{smallmatrix}\right)$ with an $m \times m$ block in the upper left corner, and T consists of permutation matrices.

For each m we assume that we have the following additional ingredients:

- the presentation $\langle X \mid R\rangle$ for F in Theorem 4.10, except in the case of the pairs $(m,q) = (4,2)$, $(4,3)$ or $(6,2)$, in which case $\langle X \mid R\rangle$ is given in Section 8.2;
- the presentation $\langle Y \mid S\rangle$ for T in Theorem 3.40 (where X and Y are disjoint);
- $W := \mathrm{SU}(2,q) < F$, consisting of the matrices $\left(\begin{smallmatrix} * & 0\\ 0 & I\end{smallmatrix}\right)$ with a 2×2 block in the upper left corner;
- $a \in W$ and $f := \left(\begin{smallmatrix} 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & I\end{smallmatrix}\right) \in F$ such that $W = \langle a, a^f\rangle$, where a and f are viewed as words in X of bit-length $O(\log q)$ using Remark 4.11;
- $c_{(1,2,3)} \in F$ that acts as the 3-cycle $(1,2,3)$ on the orthonormal basis, viewed as a word in X of bit-length $O(\log q)$ (cf. Remark 4.11);
- permutations $(1,2,3)$, $\tau = (1,2)(3,4)$, $\sigma \in T$, with σ and τ interchanging 1 and 2 and generating the set-stabilizer S_{n-2} of $\{1,2\}$ in T, where these permutations are viewed as words in Y of bit-length $O(\log n)$ (using Remark 3.37).

Case $q > 3$: Here we let $m = 3$. We will show that G is isomorphic to the group J having the following presentation, which resembles the one in Theorem 6.1. In view of the preceding remarks, this presentation has the desired bit-length.

Generators: $X \cup Y$.

Relations:

(1) $R \cup S$.
(2) $c_{(1,2,3)} = (1, 2, 3)$.
(3) $a^\sigma = a^\tau = a^f$ and $(a^f)^\sigma = a$.
(4) $[a, a^{(1,3)(2,4)}] = 1$.
(5) $[a^f, a^{(1,3)(2,4)}] = 1$ (needed only when n is 4 or 5).

Since there is a surjection $\pi\colon J \to G$, there are subgroups of J we can identify with $F = \langle X\rangle$ and $T = \langle Y\rangle$.

Since τ has order 2, (3) implies that $W^{\langle\sigma,\tau\rangle} = \langle a, a^f\rangle^{\langle\sigma,\tau\rangle} = W$, so that $|W^T| \le \binom{n}{2}$. Using π we see that W^T consists of $\binom{n}{2}$ subgroups we can call $W_{i,j} = W_{j,i}$, $1 \le i < j \le n$. By (2), $W_{i,j} \le F$ for $1 \le i < j \le 3$, and $W_{1,2}$, $W_{2,3}$ is a standard pair in F since W and c are defined in terms of an orthonormal basis for the 3-space underlying F.

Relations (4) and (5) imply that $[W, W^{(13)(24)}] = 1$ exactly as in (6.2). If i, j, k, l are distinct, the transitivity of T then implies that $[W_{i,j}, W_{k,l}] = 1$ and $\langle W_{i,j}, W_{i,k}\rangle \cong \langle W_{1,2}, W_{1,3}\rangle = F$. Since $W_{1,3} \le \langle W_{1,2}, W_{2,3}\rangle$, we have $W_{i,k} \le \langle W_{i,j}, W_{j,k}\rangle$ for all distinct i, j, k, again by the transitivity of T.

Let $U_i := W_{i,i+1}$ and $U_{i,j} := \langle U_i, U_j\rangle$ for $i \ne j$. We have seen that U_i, U_{i+1} is a standard pair in $U_{i,i+1}$. Thus, these subgroups satisfy (P1)–(P2). By Theorem 8.1, $N := \langle U_{i,j} \mid 1 \le i < j \le n-1\rangle$ is a homomorphic image of G.

We claim that $W_{i,j} \le \langle W_{i,i+1}, \dots, W_{j-1,j}\rangle$ whenever $i < j$. This is clear if $i+1 = j$, so assume that $i + 1 < j$. By induction,

$$W_{i,j} \le \langle W_{i,j-1}, W_{j-1,j}\rangle \le \langle W_{i,i+1}, \dots, W_{j-2,j-1}, W_{j-1,j}\rangle,$$

which proves our claim.

Consequently, $W_{i,j} \le N$ for all $i < j$, so that $N = \langle W^T\rangle = \langle F^T\rangle \trianglelefteq J$. By (2), J/N is a quotient of A_n in which $(1, 2, 3)$ is mapped to 1. Thus, $J/N = 1$.

Total: $|X| + |Y| = 3 + 3$ generators and $|R| + |S| + 5 = 21 + 7 + 5 = 33$ relations if $n > 5$.

As in the first fine-tuning in the proof of Theorem 6.1, we can delete a generator and a relation as follows. Start with $T = \langle Y \mid S\rangle$. By Remark 4.11, the element $t \in F$ can be written as a word in $X' := \{u, h, c_{(1,2,3)}\}$ of bit-length $O(\log q)$. Use this to rewrite the presentation for F in Theorem 4.10 using the generating set X'. In the resulting presentation for F replace $c_{(1,2,3)}$ by $(1, 2, 3) \in T$, viewed as a word in Y, and insert all of the resulting relations into the above presentation for G. Then Theorem 4.10 shows that $\langle (X' \backslash \{c_{(1,2,3)}\}) \cup \{(1, 2, 3)\}\rangle$ is F and $(1, 2, 3)$ is precisely the permutation matrix $c_{(1,2,3)} \in F$. We have now deleted the generator t from X, and we can also delete the above relation (2) since it already holds in F.

This proves (i) when $q > 3$ and $n > 5$. (N.B.: Similar decreases can be obtained in the later cases $q \leq 3$, but these are not needed for (i).)

As in the fine-tuning in the proof of Theorem 6.1 when $4 \leq n \leq 8$, in these cases we use a presentation $\langle Y \mid S\rangle$ for A_n with 2 generators and 3 relations, producing $3+2$ generators and $21+3+6$ relations for G.

Case $q = 3$: This time we let $m = 4$ and use the presentation $\langle X \mid R\rangle$ for $\mathrm{SU}(4, 3)$ given in Section 8.2. We assume that we have

- the elements a, f, $c_{(1,2,3)}$, $(1, 2, 3)$, τ, σ listed above;
- $c_{(2,3,4)} \in F$ that acts as the indicated permutation on the orthonormal basis, viewed as a word in X;
- $(2, 3, 4)$, viewed as a word in Y of bit-length $O(\log n)$.

We will show that G is isomorphic to the group J having the following presentation.

Generators: $X \cup Y$.

Relations:

(1′) $R \cup S$.
(2′) $c_{(1,2,3)} = (1, 2, 3)$, $c_{(2,3,4)} = (2, 3, 4)$.
(3′) $a^\sigma = a^f$ and $(a^f)^\sigma = a$.

By (2′), $\tau = (1, 2)(3, 4) \in F$. Then $\tau, f \in F = \mathrm{SU}(4, 3)$ agree in their action on W. As before, (3′) implies that $W^{\langle\sigma,\tau\rangle} = W$ and hence that W^T consists of $\binom{n}{2}$ subgroups $W_{i,j} = W_{j,i}$, $1 \leq i < j \leq n$. By (2′), the 6 subgroups $W_{i,j}$, $1 \leq i < j \leq 4$, are in F, and $W_{1,2}$, $W_{2,3}$ is a standard pair in a subgroup $\mathrm{SU}(3, 3)$ of F.

This time $[W, W^{(13)(24)}] = 1$ already holds in F. Consequently, if i, j, k, l are distinct, then the transitivity of T implies that $[W_{i,j}, W_{k,l}] = 1$, $\langle W_{i,j}, W_{i,k}\rangle \cong \langle W_{1,2}, W_{1,3}\rangle \cong \mathrm{SU}(3, 3)$ and $\langle W_{i,j}, W_{i,k}, W_{i,l}\rangle \cong \langle W_{1,2}, W_{1,3}, W_{1,4}\rangle = F$.

Once again, the subgroups $U_i := W_{i,i+1}$ and $U_{i,j} := \langle U_i, U_j\rangle$ of J satisfy (P1)–(P2). They also behave as in Theorem 8.1(b1) since $\langle U_{i,i+1}, U_{i+1,i+2}\rangle \cong \langle U_{1,2}, U_{2,3}\rangle = \langle W_{1,2}, W_{2,3}, W_{3,4}\rangle = F$.

Once again, the subgroup N generated by all $U_{i,j}$ is isomorphic to G. As before, N is normal in $J = \langle X, Y\rangle$ and hence is J.

Total: $|X| + |Y| = 2 + 3$ generators and $|R| + |S| + 4 = 4 + 7 + 4$ relations.

Case $q = 2$: This time we let $m = 6$. Using the presentation for $F = \mathrm{SU}(6, 2)$ in Section 8.2 we may assume that we have

- generators a', b' for $V := \mathrm{SU}(3, 2)$, viewed as words in X;
- $c_{(1,2,3)}, c_{(1,2)(4,5)}, c_{(2,3,4,5,6)} \in F$ acting as the indicated permutations on the orthonormal basis and viewed as words in X;
- $(2, 3, 4, 5, 6)$, $\tau' = (1, 2)(4, 5)$, $\sigma' \in T$, where $\langle\sigma', \tau'\rangle$ is the set-stabilizer of $\{1, 2, 3\}$ in T and these permutations are viewed as words in Y of bit-length $O(\log n)$ (using Remark 3.37). We may assume that $(1, 2, 3)$ is a cycle of σ'.

We will show that G is isomorphic to the group J having the following presentation.

Generators: $X \cup Y$.

Relations:

(1″) $R \cup S$.
(2″) $c_{(1,2)(4,5)} = \tau'$, $c_{(2,3,4,5,6)} = (2, 3, 4, 5, 6)$.
(3″) $a'^{\sigma'} = a'^{c_{(1,2,3)}}$ and $b'^{\sigma'} = b'^{c_{(1,2,3)}}$.

By (2″), $\tau' \in F$ and hence τ' normalizes V. By (3″), it follows that $\langle a', b'\rangle = V$ is normalized by $\langle \sigma', \tau'\rangle$. Then V^T consists of $\binom{n}{3}$ subgroups $V_{i,j,k}$, $1 \le i < j < k \le n$. By (2″), the 20 subgroups $V_{i,j,k}$, $1 \le i < j < k \le 6$, are in $F = \mathrm{SU}(6, 2)$.

By the transitivity of T, if i, j, k, l, r, s are distinct then $\langle V_{i,j,k}, V_{i,j,l}\rangle \cong \langle V_{1,2,3}, V_{1,2,4}\rangle \cong \mathrm{SU}(4, 2)$ and $[V_{i,j,k}, V_{l,r,s}] = 1$.

Again by (2″), $F \cap T$ induces A_6 on an orthonormal basis of the 6-space for F. Let $W_{1,2}$ denote the subgroup $\mathrm{SU}(2, 2)$ on the 2-space spanned by the first two basis vectors. Within F we see that $W_{1,2}$ lies in both $V_{1,2,3}$ and $V_{1,2,4}$. Then $N_T(W_{1,2})$ contains $(1, 2)(3, 4) \in F$, $\mathrm{Alt}\{4, 5, \ldots, n\}$, $\mathrm{Alt}\{3, 5, \ldots, n\}$, and hence also the set-stabilizer of $\{1, 2\}$.

It follows that W^T consists of $\binom{n}{2}$ subgroups $W_{i,j} = W_{j,i}$, $1 \le i < j \le n$. By (2″), the 15 subgroups $W_{i,j}$, $1 \le i < j \le 6$, are in $F = \mathrm{SU}(6, 2)$, and $W_{1,2}$, $W_{2,3}$ is a standard pair in $V = \mathrm{SU}(3, 2) < F$. Moreover, $W_{i,j} \le V_{i,j,k}$ for all distinct i, j, k, since $W_{1,2} \le V_{1,2,3}$.

Let $U_i := W_{i,i+1}$, $U_{i,i+1} := V_{i,i+1,i+2}$ and $U_{i,j} := \langle U_i, U_j\rangle$ iff $|i - j| > 1$.

The subgroups $U_{i,j}$ satisfy (P1)–(P2). They also behave as in Theorem 8.1(b1) since $\langle U_{i,i+1}, U_{i+1,i+2}\rangle \cong \langle U_{1,2}, U_{2,3}\rangle = \mathrm{SU}(4, 2)$. The conditions in Theorem 8.1(b2) also hold since they hold for the subgroups U_1, $U_{4,5}$ and $U_{1,2}$, $U_{4,5}$ of F.

Hence, the subgroup N generated by all $U_{i,j}$ is isomorphic to G. As before, N is normal in $J = \langle X, Y\rangle$ and hence is J.

Total: $|X| + |Y| = 2 + 3$ generators and $|R| + |S| + 4 = 8 + 7 + 4$ relations.

This proves (i) and (ii) for all q. Each group in the final assertion is the quotient of a group in (i) or (ii) by a cyclic group, and hence is obtained as in the proof of Theorem 6.1 by adding at most one new relation of bit-length $O(\log n + \log q)$. □

When $n = 4$, this theorem should be compared to Theorem 7.1, which contains a presentation for $\mathrm{SU}(4, q)$ with 5 generators, 27 relations and bit-length $O(\log q)$.

9. General case

We now complete the proofs of Theorems A and B.

Theorem 9.1. *All simply connected groups of Lie type and rank $n \ge 3$ over $\mathbb{F}_q$, and their simple quotients, have presentations with at most* 9 *generators*, 49 *relations and bit-length $O(\log n + \log q)$. More precisely, with the stated bit-length,*

(i) $\mathrm{SL}(n + 1, q)$ *has a presentation with* 6 *generators and* 25 *relations, with at most one further relation needed to obtain a presentation for* $\mathrm{PSL}(n + 1, q)$;

(ii) $\mathrm{SU}(2n, q)$ *and* $\mathrm{SU}(2n+1, q)$ *have presentations with* 5 *generators and* 32 *relations, with at most one further relation needed to obtain a presentation for* $\mathrm{PSU}(2n, q)$ *or* $\mathrm{PSU}(2n + 1, q)$;

(iii) $\mathrm{Sp}(2n, q)$ *has a presentation with* 8 *generators and* 47 *relations if q is odd, and* 9 *generators and* 40 *relations if q is even, with at most one further relation needed to obtain a presentation for* $\mathrm{PSp}(2n, q) \cong \Omega(2n + 1, q)$;

(iv) $\mathrm{Spin}_{2n+1}(q)$ *with q odd,* $\mathrm{Spin}^-_{2n+2}(q)$ *and* $\Omega^-(2n + 2, q)$ *have presentations with* 9 *generators and* 48 *relations, with at most one further relation needed to obtain a presentation for* $\Omega(2n + 1, q)$ *or* $\mathrm{P}\Omega^-(2n + 2, q)$;

(v) $\mathrm{Spin}^+_{2n}(q)$ *has a presentation with* 9 *generators and* 42 *relations, with at most two further relations needed to obtain a presentation for* $\mathrm{P}\Omega^+(2n, q)$;

(vi) $\hat{E}_n(q)$ *has a presentation with* 6 *generators and* 30 *relations, with at most one further relation needed to obtain a presentation for* $E_n(q)$;

(vii) $F_4(q)$ *and* ${}^2\hat{E}_6(q)$ *have presentations with* 8 *generators and* 46 *relations, with at most one further relation needed to obtain a presentation for* ${}^2E_6(q)$.

Proof. Theorems 6.1 and 8.2 take care of (i) and (ii): we may assume that G is neither a special linear nor unitary group. We use a variation on the methods in [GKKL1, Section 6.2]. As usual, $\Pi = \{\alpha_1, \ldots, \alpha_n\}$ is the set of fundamental roots of G, and for each i there are root groups $U_{\pm\alpha_i}$.

Case 1: G is a classical group. We will assume that G is one of the groups in Table 4. As in [BGKLP, GKKL1] or the proof of Theorem 6.1, each of the remaining groups can be obtained by killing (part of) the center of one of these using at most two additional relations of bit-length $O(\log n + \log q)$ [GLS, pp. 312–313].

Number Π as in [GKKL1, Section 6.2]: the subsystem $\{\alpha_1, \ldots, \alpha_{n-1}\}$ is of type A_{n-1}, α_n is an end node root and is connected to only one root α_j in the Dynkin diagram (here $j = n - 1$ except for type D_n, where $j = n - 2$). Let

$$G_1 = \langle U_{\pm\alpha_i} \mid 1 \le i < n\rangle, G_2 = \langle U_{\pm\alpha_n}, U_{\pm\alpha_j}\rangle, L_2 = \langle U_{\pm\alpha_n}\rangle, L = \langle U_{\pm\alpha_j}\rangle. \quad (9.2)$$

Summarizing part of Table 4: G_1 has type A_{n-1} and G_2 is a rank 2 group—of type $A_1 \times A_1$ in the D_n case. (Recall that $\mathrm{Sp}(2n, q) \cong \mathrm{Spin}_{2n+1}(q)$ if q is even.)

Let L_1 be the subgroup of G_1 generated by the fundamental root groups that commute with L_2. Then L_1 is of type A_{n-2} unless G has type D_n, in which case L_1 is of type $A_1 \times A_{n-3}$.

We will use the following presentations:

- the presentation $\langle X_0 \mid R_0\rangle$ for $L = \mathrm{SL}(2, q)$ in Theorem 4.5 (cf. Remark 4.7);
- the presentation $\langle X \mid R\rangle$ for G_1 in Theorem 6.1;
- the presentation $\langle Y \mid S\rangle$ for G_2 in Theorem 7.1 or Remark 7.4 (or Theorem 4.5 for the group $\mathrm{SL}(2, q) \times \mathrm{SL}(2, q)$).

Remarks concerning these presentations.

○ Our use of $\langle X_0 \mid R_0\rangle$ will save several relations, but requires some care.

Table 4. G, G_i

G	G_1	G_2	L
$\mathrm{Sp}(2n,q)$, q odd	$\mathrm{SL}(n,q)$	$\mathrm{Sp}(4,q)$	short
$\mathrm{Spin}^+_{2n}(q)$	$\mathrm{SL}(n,q)$	$\mathrm{Spin}^+_4(q) \cong \mathrm{SL}(2,q) \times \mathrm{SL}(2,q)$	short
$\mathrm{Spin}_{2n+1}(q)$	$\mathrm{SL}(n,q)$	$\mathrm{Spin}_5(q) \cong \mathrm{Sp}(4,q)$	long
$\Omega(2n+1,q)$	$\mathrm{SL}(n,q)$	$\Omega(5,q)$	long
$\mathrm{Spin}^-_{2n+2}(q)$	$\mathrm{SL}(n,q)$	$\mathrm{Spin}^-_6(q) \cong \mathrm{SU}(4,q)$	long
$\Omega^-(2n+2,q)$	$\mathrm{SL}(n,q)$	$\Omega^-(6,q)$	long

- We assume that $X \cap Y = X_0$ and $R \cap S = R_0$: both of the presentations $\langle X \mid R\rangle$ and $\langle Y \mid S\rangle$ use a presentation for $\mathrm{SL}(2,q)$. In more detail: the presentation $\langle Y \mid S\rangle$ in Theorem 7.1 uses presentations for L and L_2 from Theorem 3.40, together with additional relations; and $\langle X_0 \mid R_0\rangle$ is used as part of a presentation for $\mathrm{SL}(3,q)$ in Theorem 5.1, which is then used as part of a presentation for G_1 in Theorem 6.1.
- Since we use $\langle X_0 \mid R_0\rangle$ for two different groups, we need the sets X_0 and R_0 explicitly, not implicitly as might have occurred in the presentation $\langle Y \mid S\rangle$ in Theorem 7.1. Thus, if L is a *long* SL_2 then we need to include an additional generator u_{α_1} and relation $u_{\alpha_1} = [u_{\alpha_2}^{w}, u_{\alpha_2}^{r_2}]$, as stated in Remark 7.4. This occurs for $\Omega(2n+1,q)$, $\mathrm{P}\Omega^-(2n+2,q)$ and their covers.

○ A smaller change is needed in the proof of Theorem 6.1. Instead of reading the roots $\alpha_1, \dots, \alpha_{n-1}$ from left to right, we read them from right to left. Therefore, in place of matrices such as $\left(\begin{smallmatrix} * & 0 \\ 0 & I \end{smallmatrix}\right)$ one should think of matrices $\left(\begin{smallmatrix} I & 0 \\ 0 & * \end{smallmatrix}\right)$. Clearly this has no meaningful effect on Theorem 6.1 or its proof.

Choose two generators for L_1 and two for L_2, all viewed as words in X or Y. We still need to verify that these choices can have the required bit-lengths.

Bit-lengths of generators for L_1 and L_2. A presentation for L_2 was already used in our presentation for G_2, so apply Remark 4.6 to any pair of generators for L_2.

While there are certainly pairs of generators of L_1, we need to find generators b, c of bit-length $O(\log n + \log q)$. First assume that $L_1 = \mathrm{SL}(n-1,q)$. As in Remark 7.6, we may assume that $n-1 \geq 4$. Let $c \in \mathrm{SL}(n-1,q)$ be an element of the monomial group that is $(1,2,\dots,n-1)$ (acting on the standard basis) if $n-1$ is odd, and is $(1,\dots,n-2)t'$ if $n-1$ is even, where $t' := \left(\begin{smallmatrix} I & 0 \\ 0 & t \end{smallmatrix}\right) \in L$ with t one of the generators in (4.4). Since the presentation for A_n in Theorem 3.40 was involved in the presentation appearing in Theorem 6.1, c has bit-length $O(\log n + \log q)$ in $X \cup Y$ using Remark 3.37.

Two of the generators of L are u and h (in (4.4))), where u and h^{c^2} commute. Then $b := uh^{c^2}$ has bit-length $O(\log n + \log q)$ as in Remark 7.6, and u and h^{c^2} are powers of b. Now $L_1 = \mathrm{SL}(n-1,q)$ is generated by the $\langle c\rangle$-conjugates of the root group $\langle u^{\langle h\rangle}\rangle$ of L, and hence by b and c, as required.

When G has type D_n the group L_1 has type $A_1 \times A_{n-3}$, and 2 generators of the desired bit-length can be found in the same way.

We will show that G is isomorphic to the group J having the following presentation which, in view of the preceding remarks, has the desired bit-length.

Generators: X, Y.

Relations:

(1) $R \cup S$.

(2) $[L_1, L_2] = 1$.
More precisely, impose 4 commutation relations using pairs of words in $X \cup Y$ that map onto the chosen generators of L_1 or L_2.

We claim that $J \cong G$. For, J surjects onto G by Table 4, and hence we may assume that $G_1 = \langle X \rangle$, $G_2 = \langle Y \rangle$, L, L_1 and L_2 are subgroups of J, where $L \leq G_1 \cap G_2$. Clearly J is generated by the fundamental root groups contained in G_1 or G_2. Any two of these root groups satisfy the Curtis–Steinberg–Tits relations (see the references in Section 2): either they are both in G_1 or in G_2, or they commute since $[L_1, L_2] = 1$. Thus, J is a homomorphic image of the simply connected cover of G. Table 4 tells us that there is only one possible image for each choice of G_2, which proves the claim.

This presentation uses $|X| + |Y| - |X \cap Y|$ generators and $|R| + |S| + 4 - |R \cap S|$ relations (4 relations ensure that $[L_1, L_2] = 1$). By Theorems 5.1, 6.1 and 7.1, together with Remark 7.4, these numbers are as follows.

G	Generators	Relations
$\mathrm{Sp}(2n, q)$, q odd	$6 + 5 - 3$	$25 + 27 + 4 - 9$
$\mathrm{Spin}_{2n+1}(q)$	$6 + 6 - 3$	$25 + 28 + 4 - 9$
$\mathrm{Spin}^-_{2n+2}(q)$	$6 + 6 - 3$	$25 + 28 + 4 - 9$
$\mathrm{Spin}^+_{2n}(q)$	$6 + (3 + 3) - 3$	$25 + (9 + 9 + 4) + 4 - 9$

For type D_n, G_2 uses two commuting copies of the presentation for $\mathrm{SL}(2, q)$ given in Theorem 4.5; one can slightly reduce this presentation for $\mathrm{SL}(2, q) \times \mathrm{SL}(2, q)$. For $\Omega(2m + 1, q) \cong \mathrm{Sp}(2m, q)$ with q even in (iii), the presentation for $\mathrm{Sp}(4, q)$ in Theorem 7.1(i) has 7 fewer relations than when q is odd.

This proves parts (iii)–(v) of the theorem.

The groups $\mathrm{Sp}(6, q)$ *and* $\mathrm{Spin}^-_8(q)$: In the cases $F_4(q)$ and ${}^2\hat{E}_6(q)$ we will use the $\langle X_0 \mid R_0 \rangle$ trick for each of the fundamental groups A_1, and this requires Remark 7.4. For $\mathrm{Sp}(6, q)$ and $\mathrm{Spin}^-_8(q)$ we also use the presentation in Theorem 5.1 instead of the one in Theorem 6.1, in each case obtaining $|X| + |Y| - |X \cap Y| = 4 + 6 - 3 = 7$ generators and $|R| + |S| + 4 - |R \cap S| = 14 + 28 + 4 - 9 = 37$ relations.

Case 2: G is an exceptional group. We modify the above argument slightly. Let G be the simply connected cover of the simple group. We use subgroups G_1, G_2, L_1 of G whose types are in the following table:

G	G_1	G_2	L_1
E_8	A_7	A_2	$A_2 \times A_4$
E_7	A_6	A_2	$A_2 \times A_3$
E_6	A_5	A_2	$A_2 \times A_2$
F_4	C_3	A_2	A_2
${}^2E_6$	${}^2D_4$	A_2	A_2

while L and L_2 have type A_1. Here $G_2 = \langle L, L_2 \rangle$ and L_2 corresponds to an end node of the Dynkin diagram.

We again use a presentation $\langle X \mid R \rangle$ for G_1 and a presentation $\langle Y \mid S \rangle$ for G_2, where $X \cap Y = X_0$ and $R \cap S = R_0$ for a presentation $\langle X_0 \mid R_0 \rangle$ for $L = \mathrm{SL}(2, q)$. Once again L_1 and L_2 are generated by 2 elements of bit-length $O(\log q)$ (as in Remarks 5.10 and 7.6). Precisely as above the previous relations (1) and (2) produce a presentation:

- *For* $G = \hat{E}_n(q)$: $|X|+|Y|-|X \cap Y| = 5+4-3$ generators and $|R|+|S|+4-|R \cap S| = 21 + 14 + 4 - 9 = 30$ relations by Theorems 6.1(iii) and 5.1;
- *For* $G = F_4(q)$ *or* ${}^2\hat{E}_6(q)$: $|X| + |Y| - |X \cap Y| = 7 + 4 - 3 = 8$ generators and $|R| + |S| + 4 - |R \cap S| = 37 + 14 + 4 - 9 = 46$ relations using Theorem 5.1 and the presentation just obtained for $G_1 = \mathrm{Sp}(6, q)$ or $\mathrm{Spin}_8^-(q)$ with 7 generators and 37 relations.

Since $\langle X \mid R \rangle$ and $\langle Y \mid S \rangle$ have bit-length $O(\log q)$, the same is true of our presentation for G. Once again, as in [GKKL1] it is easy to kill the center when G is $\hat{E}_6(q)$, ${}^2\hat{E}_6(q)$ or $\hat{E}_7(q)$ [GLS, p. 312], using at most one additional relation of bit-length $O(\log q)$. □

Remark 9.3. The exact same approach could have been used for the unitary groups $\mathrm{SU}(2m, q)$, starting with the presentation for $\mathrm{SU}(4, q)$ in Theorem 7.1(ii). The result would be *a presentation for* $\mathrm{SU}(2m, q)$ *using* 8 *generators and* 47 *relations*—more relations than in Section 8. Odd-dimensional unitary groups are more of a problem: even for $\mathrm{SU}(5, q)$ we do not know how to use the Steinberg presentation to obtain as few relations as in Section 8.

The same method also produces a presentation for $\mathrm{SU}(6, q)$ having 7 generators and 37 relations, yielding a presentation for ${}^2\hat{E}_6(q)$ having the same numbers of generators and relations as above.

Proof of Theorems A and B. As pointed out in Section 1, Theorem A follows from Theorem B and [GKKL1, Lemma 2.1] or Lemma 2.3 (cf. Corollary 3.43).

For most of the groups in Theorem B, the required presentation is contained in Theorems 3.40, 4.5, 4.10, 5.1, 6.1, 8.2 and 9.1, together with Propositions 4.12 and 7.7. For other quotients of quasisimple linear groups we need at most one further relation. For example, at the end of the proof of Theorem 6.1 we provided a generator for the center of $\mathrm{SL}(n, q)$, and hence we can factor out any of its powers.

By Lemma 2.2, starting with a presentation $\langle X \mid R \rangle$ for some perfect central extension of the simple group, each perfect central extension by a cyclic group has a presentation with $|X|+1$ generators, $|X|+|R|+1$ relations and the desired bit-length. This takes care of alternating groups using Theorem 3.40.

It remains to consider the sporadic perfect central extensions of the groups G already dealt with [GLS, pp. 312–313]: $6A_6 \cong 6\mathrm{PSL}(2, 9)$, $(4 \times 4)\mathrm{SL}(3, 4)$, $2\mathrm{Sp}(6, 2)$, $3\Omega(7, 3)$, $(2 \times 2)\mathrm{SU}(6, 2)$, $(3 \times 3)\mathrm{SU}(4, 3)$, $3G_2(3)$, $2G_2(4)$, $(2 \times 2)\mathrm{Sz}(8)$, $(2 \times 2)\Omega^+(8, 2)$, $2F_4(2)$, $(2 \times 2){}^2\hat{E}_6(2)$. (Note that $2\mathrm{SL}(4, 2) \cong 2A_8$ and $2\mathrm{SU}(4, 2) \cong \mathrm{Sp}(4, 3)$.) Some cases are already in the literature [CHRR1, CHRR2, CHRR3] or have been seen earlier. We need to improve other presentations so that our initial number of relations is significantly smaller than before. This amounts to using known presentations for groups such as SL(2, 2), SL(3, 2) or SL(3, 4) in place of our previous general presentations, and then proceeding as in earlier sections. It is straightforward to check that, even when we need to use Lemmas 2.2 or 2.3 more than once, in all remaining cases we obtain at most 9 generators and 49 relations. Of course, bit-length is not an issue here.

As an example, consider $G = {}^2\hat{E}_6(2)$. Use $G_1 = \mathrm{SU}(6, 2)$ and $G_2 = \mathrm{SL}(3, 2)$ in the proof of Theorem 9.1. Since these have presentations with 2 generators and 8 relations (Section 8.2) and 2 generators and 3 relations, respectively, the argument used in Theorem 9.1 produces a presentation for G with $2+2$ generators and $8+3+4$ relations, hence Lemma 2.3 produces another presentation with 2 generators and 16 relations. Next we obtain a presentation of any cover $2^2\hat{E}_6(2)$ with $2+1$ generators and $2+16+1$ relations by using Lemma 2.2, and then one with 2 generators and 20 relations by Lemma 2.3. Finally, we obtain a presentation for $(2 \times 2)^2\hat{E}_6(2)$ with $2+1$ generators and $2+20+1$ relations using Lemma 2.2, and then one with 2 generators and 24 relations by Lemma 2.3. It is now easy to obtain presentations for all covers of ${}^2E_6(2)$. □

Call a group G *almost quasisimple* if it has a normal quasisimple subgroup S with $C_G(S) = Z(S)$. Then Theorem A implies the following

Corollary 9.4. *All* 2*-generated almost quasisimple finite groups of Lie type, except possibly when the socle is* ${}^2G_2(q)$, *have presentations with* 2 *generators and at most* 60 *relations. If the group is not* 2*-generated then there is a presentation with* 3 *generators and at most* 60 *relations.*

Proof. Let G be such a group with quasisimple normal subgroup S. Then G/S embeds in Out(S). We claim that any subgroup of Out(S) has a presentation with at most 3 generators and 6 relations. If S is not of type D_4, by [GLS, Theorem 2.5.1] there is a normal series of Out(S) of length at most 3 with all quotients cyclic. If there are 2 terms then there is a presentation with 2 generators and 3 relations. If there are 3 terms then there is a presentation with 3 generators and 6 relations. If $S/Z(S) \cong \mathrm{P}\Omega^+(8, p^e)$, then $\mathrm{Out}(S) \cong S_4 \times \mathbb{Z}_e$ when $p > 2$ and $\mathrm{Out}(S) \cong S_3 \times \mathbb{Z}_e$ when $p = 2$. Since any subgroup of S_4 has a presentation with 2 generators and at most 3 relations, the claim holds in this case as well.

By [DaL], if $d(G)$ is the minimum number of generators of G then we have $d(G) = \min\{2, d(G/S)\}$. (N.B.: For groups of Lie type, the proof in [DaL] does not use the classification of the finite simple groups.) Thus, by Lemma 2.4, if S has a presentation with 2 generators and r relations, then G has a presentation having at most $2 + d(G)$ generators and $r + 6 + 2 \cdot 3$ relations, and hence another presentation with 2 generators and $(2 + d(G)) + (r + 12)$ relations by Lemma 2.3. As in the proof just provided for Theorems A and B, it is now straightforward to check all cases in Theorem A. □

10. Additional presentations for classical groups

We now provide an alternative to the preceding section for presentations of classical groups. We will use Section 3.6 and its notation concerning the group $W_n = \mathbb{Z}_2^{n-1} \rtimes A_n$, consisting of $n \times n$ real monomial matrices with respect to the standard orthonormal basis $\{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$.

10.1. Groups of type D_n

By Theorem 6.1, since $\mathrm{P}\Omega^+(6, q) \cong \mathrm{PSL}(4, q)$ we only need to consider the case $n \geq 4$. We will use Proposition 3.46 in order to imitate the argument in Theorem 6.1.

Theorem 10.1. *$\Omega^+(2n, q)$ has a presentation with*

(i) 8 *generators and* 31 *relations if $n \geq 4$,*
(ii) 7 *generators and* 27 *relations if $n = 4$ or* 5.

Each of these presentations has bit-length $O(\log n + \log q)$. At most one additional relation of bit-length $O(\log n + \log q)$ is needed to obtain a presentation for $\mathrm{P}\Omega^+(2n, q)$.

Proof. There is a hyperbolic basis $e_1, f_1, \dots, e_n, f_n$ of $V = \mathbb{F}_q^{2n}$ associated with $G = \Omega^+(2n, q)$. The group $W := W_n$ lies in G and permutes the pairs $\{e_i, f_i\}$, $1 \leq i \leq n$: if $n \geq 5$ then $W < G$ follows from the fact that W is perfect, and for $n = 4$ we can see this by restricting from the group $\Omega^+(10, q)$.

Each element of W can be viewed using two different vector spaces: $\mathbb{R}^n$ and V. In the action on $\mathbb{R}^n$, we write elements in terms of the standard orthonormal basis as permutations of $\{1, \dots, n, -1, \dots, -n\}$, as in Section 3.6. The resulting diagonal matrices in W are the elements of W leaving each pair $\{e_i, f_i\}$ invariant in its action on V. Since these two views are potentially confusing (especially when q is even), we will usually write elements in both manners.

We digress in order to observe that, *when q is odd, W does not lift to an isomorphic copy inside the simply connected cover $\hat{G}$ of G.* For, Steinberg's criterion [St3, Corollary 7.6] states that an involution in $\Omega^+(2n, q)$ lifts to an element of order 4 in the spin group if and only if the dimension of its -1 eigenspace on V is $\equiv 2 \pmod 4$. Apply this to $\mathrm{diag}(-1, -1, 1, \dots, 1) = (e_1, f_1)(e_2, f_2) \in W$ in order to obtain an element of order 4 in $\hat{G}$, which proves our claim.

We view $\mathrm{SL}(3, q)$ as the subgroup of G preserving the subspaces $\langle e_1, e_2, e_3\rangle$ and $\langle f_1, f_2, f_3\rangle$ while fixing the remaining basis vectors of V. We use the following presentations:

- the presentation $\langle X \mid R\rangle$ for $F = \mathrm{SL}(3, q)$ in Theorem 5.1,
- the presentation $\langle Y \mid S\rangle$ for W in Proposition 3.46 (with X and Y disjoint).

We also use the following elements:

- $c = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $f = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \in F$;
- $a \in F$ such that $L := \langle a, a^f\rangle \cong \mathrm{SL}(2, q)$ consists of all matrices $\left(\begin{smallmatrix} * & 0 \\ 0 & 1 \end{smallmatrix}\right)$ in F;
- $(3, 2, 1) = (e_3, e_2, e_1)(f_3, f_2, f_1)$, $(1, 3)(2, 4) = (e_1, e_3)(e_2, e_4)(f_1, f_3)(f_2, f_4) \in W$, and $s = \mathrm{diag}(-1, -1, 1, 1, \dots, 1) = (e_1, f_1)(e_2, f_2) \in W$ representing an element of Y (cf. Proposition 3.46);
- $\tau = (1, 2)(3, 4) = (e_1, e_2)(e_3, e_4)(f_1, f_2)(f_3, f_4)$ and σ that generate the stabilizer in W of $\langle v_1 - v_2\rangle$ (within $\mathbb{R}^n$) and send $v_1 - v_2$ to $v_2 - v_1$.

Bit-length: c, f and a have bit-length $O(\log q)$ using Remark 5.10. We may assume that σ is the product of s and a cycle of odd length $n-2$ or $n-3$ on $\{3,\dots,n\}$, so that the stated elements of W all have bit-length $O(\log n)$ in Y (by Remark 3.37).

We will show that G is isomorphic to the group J having the following presentation.

Generators: X, Y.

Relations:

(1) R.
(2) S.
(3) $c=(3,2,1)$.
(4) $a^\sigma = a^f$, $a^\tau = a^f$.
(5) $(a^f)^\sigma = a$.
(6) $[a, a^{(1,3)(2,4)}]=1$.
(7) $[a^f, a^{(1,3)(2,4)}]=1$ if $n=4$ or 5.
(8) $[a, a^{s^{(3,2,1)}}]=1$.
(9) $[a^f, a^{s^{(3,2,1)}}]=1$ if $n=4$.

First we verify these relations in G. The vectors $\pm(v_1-v_2)$ and $\pm(v_1+v_2)$ are in the root system Φ for G. Since these pairs are perpendicular, the associated root groups determine commuting subgroups $\langle X_{\pm(v_1-v_2)}\rangle$ and $\langle X_{\pm(v_1+v_2)}\rangle \cong \mathrm{SL}(2,q)$, where $\langle X_{\pm(v_1-v_2)}\rangle^{s^{(3,2,1)}} = \langle X_{\pm(v_1+v_2)}\rangle$ since $s^{(3,2,1)} = \mathrm{diag}(1,-1,-1,1,\dots,1)$. This implies relations (8) and (9). For a similar reason $\langle X_{\pm(v_1-v_2)}\rangle$ and $\langle X_{\pm(v_3-v_4)}\rangle$ commute, which implies relations (6) and (7). Thus, there is a surjection $J\to G$.

Now consider J. As usual, we may assume that $F=\langle X\rangle$ and $W=\langle Y\rangle$ lie in J. Using (4), (5) and $|\tau|=2$, we see that $\langle\sigma,\tau\rangle$ normalizes L. Then L^W can be identified with the set of $n(n-1)$ pairs $\{\pm\alpha\}$ of vectors $\alpha=\pm v_i\pm v_j\in\mathbb{R}^n$, $i\neq j$, in the root system Φ. The groups in L^W produce additional root groups X_α, $\alpha\in\Phi$.

Any unordered pair of distinct, nonopposite roots can be moved by W to one of the following pairs within $\mathbb{R}^n$:

$$\text{(a)}\ v_1-v_2,\ v_1+v_2,\qquad \text{(b)}\ v_1-v_2,\ \pm(v_1-v_3),\qquad \text{(c)}\ v_1-v_2,\ v_3-v_4.$$

Then the corresponding root groups can be moved in the same manner.

Let $N:=\langle L^W\rangle = \langle X_\alpha \mid \alpha\in\Phi\rangle = \langle F^W\rangle \trianglelefteq J$.

We need to verify the Steinberg relations for the root groups X_α. The pairs (b) already lie in $F=\mathrm{SL}(3,q)$, so the desired relations are immediate. It remains to consider the pairs (a) and (c).

As in (6.2), (6) and (7) imply that the root groups determined by (c) commute.

Before considering (a), we note that (4) and (5) imply that every element of W that interchanges v_1-v_2 and v_2-v_1 also interchanges a and a^f. Two such elements are s and, when $n\ge 5$, also $(1,2)(4,5)$. Since $t:=s^{(3,2,1)}=\mathrm{diag}(1,-1,-1,1,\dots,1)$ commutes with s, and $t^{(1,2)(4,5)}=\mathrm{diag}(-1,1,-1,1,\dots,1)=st$, by (8) we have

$$\begin{aligned}1&=[a^s,a^{ts}]=[a^f,(a^f)^t],\\ 1&=[a^{(1,2)(4,5)},a^{t^{(1,2)(4,5)}}]=[a^f,(a^{(1,2)(4,5)s})^t]=[a^f,a^t].\end{aligned}$$

By (9), the second of these relations also holds when $n = 4$. Then also $1 = [a^{ft}, a]$. Now the root groups $X_{v_1-v_2} < L_{v_1-v_2} = L = \langle a, a^f \rangle$ and $X_{v_1+v_2} < L_{v_1+v_2} = L^t = \langle a^t, a^{ft} \rangle$ commute, as required for (a).

Thus, N is a perfect central extension of G. We have seen that W prevents it from being the simply connected cover, so that $N \cong G$. Relation (3) pulls (3, 2, 1) into N, so that $J/N = 1$.

If $n = 4$ we can delete σ in (4) and (5). If $n = 5$ we may assume that $|\sigma| = 2$ and hence delete (5). Thus, in all cases relations (3)–(9) contribute 6 relations.

By Proposition 3.46, this presentation for $\Omega^+(2n, q)$ has the stated numbers of generators and relations. Finally, at most one further relation of bit-length $O(\log n + \log q)$ is needed to kill the center. □

As in the proof of Theorem 6.1, we can remove a generator and a relation by using $c := (3, 2, 1) \in W$ in the presentation for $\mathrm{SL}(3, q)$ in Theorem 5.1.

10.2. Groups of type B_n or C_n

Next we will glue $\mathbb{Z}_2^{n-1} \rtimes A_n$ or $\mathbb{Z}_4^n \rtimes A_n$ and a group of type B_2 or C_2.

Theorem 10.2. *Let $n \geq 4$.*

(i) $\mathrm{Sp}(2n, q)$ *and* $\Omega(2n + 1, q)$ *with q odd*, $\mathrm{SU}(2n, q)$ *and* $\Omega^-(2n + 2, q)$ *have presentations with* 9 *generators and* 50 *relations*;

(ii) $\mathrm{Sp}(2n, q)$ *has a presentation with* 10 *generators and* 43 *relations if q is even.*

Each of these presentations has bit-length $O(\log n + \log q)$. At most one additional relation of bit-length $O(\log n + \log q)$ is needed to obtain a presentation for $\mathrm{PSp}(2n, q)$, $\mathrm{PSU}(2n, q)$ *or* $\mathrm{P}\Omega^-(2n + 2, q)$.

Proof. Let $G = \mathrm{Sp}(2n, q)$, $\mathrm{SU}(2n, q)$, $\Omega(2n + 1, q)$ or $\Omega^-(2n + 2, q)$. Its root system $\Phi \subset \mathbb{R}^n$ of type C_n or B_n consists of the vectors $\pm v_i \pm v_j$ for $1 \leq i < j \leq n$, and all $\pm 2v_i$ or $\pm v_i$, respectively. We may assume that a fundamental system is

$$\Pi = \{\alpha_1,\ \alpha_j = v_{j+1} - v_j \mid 2 \leq j \leq n - 1\},$$

where $\alpha_1 = v_1$ or $2v_1$.

We will use the subgroup $L_{12} \cong \mathrm{Sp}(4, q)$, $\mathrm{SU}(4, q)$, $\Omega(5, q)$ or $\Omega^-(6, q)$ corresponding to the root subsystem Φ_{12} generated by α_1 and α_2, and the rank 1 subgroups L_1 and L_2 of L_{12} determined by $\pm\alpha_1$ and $\pm\alpha_2$, respectively. We will also need the subgroup $L_{23} \cong \mathrm{SL}(3, q)$ (or $\mathrm{SL}(3, q^2)$ in the unitary case) corresponding to the root subsystem generated by α_2 and α_3.

There is a hyperbolic basis $e_1, f_1, \ldots, e_n, f_n$ of the natural module V for G (with additional basis vectors v, or v and v', perpendicular to all of these in the orthogonal cases). We claim that we may assume that $W = \mathbb{Z}_2^{n-1} \rtimes A_n$ or $\mathbb{Z}_4^n \rtimes A_n$ is a subgroup of G that permutes the 1-spaces $\langle e_1 \rangle, \langle f_1 \rangle, \ldots, \langle e_n \rangle, \langle f_n \rangle$, fixing any additional basis vectors v, v', with $O_2(W)$ fixing each pair $\{\langle e_i \rangle, \langle f_i \rangle\}$. Elements of A_n preserve each of the sets $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_n\}$. The specific W depends on the underlying form. If $(e_i, f_i) = (f_i, e_i) = 1$ for all i, then $W = \mathbb{Z}_2^{n-1} \rtimes A_n < G$

acts on $\{e_1, f_1, \dots, e_n, f_n\}$, with $O_2(W)$ fixing each pair $\{e_i, f_i\}$. If G is symplectic and q is odd, then $W = \mathbb{Z}_4^n \rtimes A_n < G$ and $O_2(W)$ consists of transformations that are products of isometries $e_i \mapsto f_i \mapsto -e_i \mapsto -f_i$, each of which fixes all other basis vectors. This proves our claim. Morever, in either case, there is a natural map $\bar{}: W \to \overline{W} \le \mathbb{Z}_2^n \rtimes A_n$ giving the permutation action of W on $\{\langle e_1\rangle, \langle f_1\rangle, \dots, \langle e_n\rangle, \langle f_n\rangle\}$; its kernel acts as 1 on $\mathbb{R}^n$, and $\bar{W}$ acts as isometries of $\mathbb{R}^n$.

We may assume that the groups L_i and L_{12} have the following supports:

G	Support of L_1	Support of L_{12}	Support of L_2
$\mathrm{Sp}(2m, q)$	$\langle e_1, f_1\rangle$	$\langle e_1, e_2, f_1, f_2\rangle$	$\langle e_1, e_2, f_1, f_2\rangle$
$\mathrm{SU}(2m, q)$	$\langle e_1, f_1\rangle$	$\langle e_1, e_2, f_1, f_2\rangle$	$\langle e_1, e_2, f_1, f_2\rangle$
$\Omega(2m+1, q)$	$\langle e_1, f_1, v\rangle$	$\langle e_1, e_2, f_1, f_2, v\rangle$	$\langle e_1, e_2, f_1, f_2\rangle$
$\Omega^-(2m+2, q)$	$\langle e_1, f_1, v, v'\rangle$	$\langle e_1, e_2, f_1, f_2, v, v'\rangle$	$\langle e_1, e_2, f_1, f_2\rangle$

Consequently, the support of $L_2^{(1,3)(2,4)}$ is $\langle e_3, e_4, f_3, f_4\rangle$, so that

$$[L_{12}, L_2^{(1,3)(2,4)}] = 1. \tag{10.3}$$

Similarly,

$$[L_1, L_2^{(3,2,1)}] = 1. \tag{10.4}$$

We use the following presentations:

- the presentation $\langle X \mid R\rangle$ for $L_{12} \cong \mathrm{Sp}(4, q)$, $\mathrm{SU}(4, q)$, $\Omega(5, q)$ or $\Omega^-(6, q)$ in Theorem 7.1 when $G = \mathrm{Sp}(2n, q)$, $\mathrm{SU}(2n, q)$, $\Omega(2n+1, q)$ or $\Omega^-(2n+2, q)$, respectively;
- the presentation $\langle Y \mid S\rangle$ for W in Proposition 3.46 or Remark 3.48 (where X and Y are disjoint).

We use the following elements, writing elements of $\overline{W}$ as isometries of $\mathbb{R}^n$:

- $c = (3, 2, 1) = (e_3, e_2, e_1)(f_3, f_2, f_1) \in W$;
- $s \in W$ representing an element of Y, where $\bar{s} = \mathrm{diag}(-1, -1, 1, \dots, 1)$ on $\mathbb{R}^n$ (cf. Proposition 3.46 or Remark 3.48);
- $(2, 3, 4) = (e_2, e_3, e_4)(f_2, f_3, f_4)$, $(1, 3)(2, 4) = (e_1, e_3)(e_2, e_4)(f_1, f_3)(f_2, f_4) \in W$;
- $\sigma, \tau \in W$ generating the set-stabilizer $W_{\{\pm(v_2 - v_1)\}}$ of the pair $\{\pm(v_2 - v_1)\}$;
- $u, r := t, h \in L_2$, with h normalizing the root groups $X_{\pm\alpha_2}$ and $u \in X_{\alpha_2}$, as in (4.4);
- $a \in L_2$ such that $L_2 = \langle a, a^s\rangle$;
- $b \in L_{12}$ such that $L_{12} = \langle b, b^s\rangle$;
- a pair of generators for L_1.

Bit-length: By Remarks 4.6, 7.6 and 3.37, the above elements of L_1, L_2 and L_{12} have bit-length $O(\log q)$, and the above elements of W have bit-length $O(\log n)$.

The required elements a and b exist, and can be chosen to be generators of nonsplit maximal tori.

As in Section 10.1, in the orthogonal cases with q odd [St3, Corollary 7.6] implies that the subgroup W of G does not lift to an isomorphic subgroup of the simply connected cover of G.

Write $q' := q$ except in the unitary case, where $q' := q^2$.

Consider the group J having the following presentation.

Generators: X, Y.

Relations:

(1) R.
(2) S.
(3) d^{σ} and d^{τ} written as words in X (obtained from G), for each $d \in \{b, b^s\}$.
(4) $[(2,3,4), L_1] = 1$.
(5) $c^r = r^2c^2$.
(6) $hh^ch^{c^2} = 1$.
(7) $u^{h^c} = u^{\mathrm{diag}(1,\zeta^{-1})}$ written as a word in X.
(8) $[u, u^c] = u^{-rc^2}$.
(9) $[uu^c, u^{-rc^2}] = 1$.
(10) $[b, a^{(1,3)(2,4)}] = 1$.

As in the proof of Theorem 5.1, we will show that $J \cong G$, provided that (9) is replaced by the two relations $[u, u^{-rc^2}] = [u^c, u^{-rc^2}] = 1$ when $q' = 2, 4, 7, 13, 16$ or 19, together with 3 further relations given in the proof of Theorem 5.1 when $q' = 4$.

There is a surjection $J \to G$. For example, relations (2)–(6) in Theorem 5.1 are the present relations (5)–(9) for $\langle L_2, c\rangle \cong \mathrm{SL}(3, q')$, while the present relations (10) and (11) follow from (10.3) and (10.4).

As usual, we may assume that $L_{12} = \langle X\rangle$ and $W = \langle Y\rangle$ lie in J, and hence the groups L_1 and L_2 are also in J. As we just noted, we have also built $L_{23} := \langle L_2, c\rangle \cong \mathrm{SL}(3, q')$ as a subgroup of J.

Relation (3) implies $W_{\{\pm(v_2-v_1)\}}$ normalizes $\langle b, b^s\rangle = L_{12}$. Then $|L_{12}^W| \le n(n-1)/2$, and as usual this implies that equality holds. Moreover, by (3), $W_{\{\pm(v_2-v_1)\}}$ acts on L_{12} as it does in G, and hence normalizes L_2, so that L_{12}^W and L_2^W both can be identified with the set of all pairs $\{\pm(v_i - v_j)\}$, $i \ne j$. For the same reason, $W_{\{\pm v_1, \pm(v_2-v_1)\}}$ acts on L_{12} as it does in G, and hence normalizes L_1. Then so does $\langle W_{\{\pm v_1, \pm(v_2-v_1)\}}, (2,3,4)\rangle = W_{\{\pm v_1\}}$ by (4). It follows that $|L_1^W| = n$.

Consequently, starting with the root subgroups X_{α_1} and X_{α_2} of L_{12}, in J we obtain the "correct" set $X_{\alpha_1}^W \cup X_{\alpha_2}^W = \{X_\alpha \mid \alpha \in \Phi\}$ of root subgroups.

Let $N := \langle X_{\alpha_1}^W \cup X_{\alpha_2}^W\rangle = \langle L_{12}^W\rangle \trianglelefteq J$. We will verify the Steinberg relations for N. Many of the required relations already hold for L_{12} or L_{23} (which is why we built L_{23} into the presentation).

Any unordered pair α, β of distinct, nonopposite roots can be moved by W to one of the following:

(a) $\alpha_1, \pm\alpha_2$, (b) $\alpha_1, {\alpha_1}^{(1,2)(3,4)}$, (c) α_1, α_4, (d) $\alpha_2, \pm\alpha_3$, (e) α_2, α_4.

Only pairs (c) and (e) are not inside L_{12} or L_{23}.

Since each element of $W_{\{\pm(v_2-v_1)\}}$ acts correctly on L_{12}, $s' := s^{(1,3)(2,4)}$ commutes with L_{12} (compare (10.3)). By (10),

$$\begin{aligned}
1 &= [b, a^{(1,3)(2,4)}]^s = [b^s, a^{s'(1,3)(2,4)}] = [b^s, a^{(1,3)(2,4)}],\\
1 &= [b, a^{(1,3)(2,4)}]^{s'} = [b, (a^s)^{(1,3)(2,4)}],\\
1 &= [b, a^{(1,3)(2,4)}]^{s's} = [b^{s's}, (a^{s's})^{(1,3)(2,4)}] = [b^s, (a^s)^{(1,3)(2,4)}].
\end{aligned}$$

Now $[L_{12}, L_2^{(1,3)(2,4)}] = [\langle b, b^s\rangle, \langle a, a^s\rangle^{(1,3)(2,4)}] = 1$, which takes care of the pairs (c) and (e) (and explains the use of α_4 instead of α_3 in (c)).

Clearly $c \in \langle L_2, c\rangle = L_{23} = \langle L_2, L_2^c\rangle \leq N$ since $c \in W$ by definition. Thus, J/N is a homomorphic image of W in which c is mapped to 1, so that $J = N$. Now J is a central extension of G, but we have seen that W prevents it from being the simply connected cover in the orthogonal cases. Hence, $J \cong G$.

If q' is not one of the excluded numbers, use Theorem 7.1 together with Proposition 3.46 or Remark 3.48 to see that there are $6 + 4$ generators and $27 + 11 + 12 = 50$ relations in (i), and (ii) is similar. In the excluded cases, as in Theorem 5.1 there are presentations for $\mathrm{SL}(2, q')$ with small enough numbers of generators and relations to counterbalance the few additional relations we have imposed. □

Remark 10.5. We already had presentations for these groups in Section 9 having at most 9 generators and 48 relations.

There are similar presentations if n is 3 or 4.

11. Concluding remarks

1. Short and bounded presentations are goals of one aspect of Computational Group Theory ([Sims, pp. 290-291], [HEO, p. 184]). Such presentations have various applications, such as in [LG, KaS] for gluing together presentations in a normal series in order to obtain a presentation for a given matrix group. The presentations in the present paper are not short in the sense of length used in [Sims, HEO, GKKL1], and it is not clear how small bit-length might influence the speed of computer processing of a word of large length in an abstract group. However, small bit-length is expected to speed up processing in the case of matrix groups, due to a fast exponentiation algorithm in [LGO, Sec. 10].

2. The presentations for S_n and A_n in Section 3 that are related to prime numbers appear to be practical. The ones in Section 3.4 for general n have one unusual and awkward relation: $y = w$, expressing y as a word in $X \cup X^y$; see (3.42) and the description of this relation in the proof of Corollary 3.28. Experimentation appears to be needed in order to find a “nice” additional relation of this sort. That is, the presentation given in Section 3.5 is among the easiest to describe of the presentations obtained using our methods, but it may not be the best in practice.

3. In Section 3.7 we dealt with $\mathrm{Aut}(F_n)$. The following are intriguing conjectures:

(1) All braid groups are boundedly presented.
(2) All mapping class groups are boundedly presented.

4. In Corollary 3.43, we used Lemma 2.3 in order to obtain presentations of alternating and symmetric groups using 2 generators and 8 relations. The same lemma can be used in order to *decrease the number of relations by one* for the remaining groups in Theorem A. The easiest way to see this is to use the “3/2-generation” of all finite simple groups [GK], according to which any one of our generators a is a member of a generating pair $\{a, b\}$;

and then proceed as in Corollary 3.43. However, this uses an unnecessary amount of machinery, since each of our generating sets contains a member a for which a suitable b can be found without too much difficulty.

5. Many of our presentations undoubtedly can be improved. For example, the number of relations in Proposition 3.46 probably can be improved somewhat by using the ideas in Sections 3.1 and 3.4 in place of Theorem 3.40.

Phan-type presentations for orthogonal groups [BGHS, GrHNS] should help decrease the numbers of relations in Sections 7, 9 and 10.

In the proofs of Theorems 6.1 and 8.2(v) we fine-tuned our presentations for some small-dimensional linear groups in order to decrease the number of relations used for exceptional groups in Section 9. There are further small-rank cases where we could have proceeded in a similar manner, using known presentations for A_n, $n \leq 10$, with 2 generators and 3 relations [CHRR1, CHRR2, Hav] (cf. Example 3.21(10)). Also see the end of the proof of Theorems A and B in Section 9.

All of the presentations in Sections 5–10 use presentations for $\mathrm{SL}(2,q)$. When q is even these only need 5 relations (Theorem 4.5) instead of the 9 relations involved in almost all of our counts (e.g., if $q > 16$ then the presentation for $\mathrm{SL}(3,q)$ in Theorem 5.1 needs only 10 relations). Similar decreases occur if q is prime.

6. We have observed that length is a more stringent requirement than bit-length. Nevertheless, bit-length is a nontrivial restriction: we cannot prove Theorem A while retaining control over bit-length. Another example is as follows:

> *In parts* (ii)–(vii) *of Theorem* 9.1 *we can decrease the numbers of relations by* 2, *at the cost of not being able to control bit-length for large rank groups.*

Namely, we choose generators of L_1 and L_2 different from those in the proof of Theorem 9.1. If $m \geq 2$ then $\mathrm{SL}(m,q)$ is always generated by a transvection (elementary matrix) and an element of order $(q^m-1)/(q-1)$ (cf. [Ka]). Applying this with $m=2$ and $m=n-1$ produces generators a_i, b_i of L_i with a_1, b_2 transvections and a_2, b_1 of orders relatively prime to q. Write $a = a_1a_2$ and $b = b_1b_2$. There are integers k_i, l_i such that $a^{k_i} = a_i$ and $b^{l_i} = b_i$ inside G, since $(|a_1|, |a_2|) = (|b_1|, |b_2|) = 1$. Then *the* 4 *relations used to guarantee that* $[L_1, L_2] = 1$ *can be replaced by the* 2 *relations* $[a^{k_1}, b^{l_2}] = [a^{k_2}, b^{l_1}] = 1$.

However, we do not know whether an element a_2 of order $(q^{n-1}-1)/(q-1)$ can be found having bit-length $O(\log n + \log q)$ in our generators if n is arbitrarily large (cf. Remark 7.6 for bounded rank groups).

7. As observed in Section 3 we have constructed presentations for alternating and symmetric groups with bounded expo-length. The remaining presentations in this paper do not have this property, unless we only consider groups over *bounded* degree extensions of the prime field. An obstacle to our obtaining presentations with bounded expo-length is that we do not know sufficiently nice presentations of $\mathbb{F}_q$ when this field has large degree over the prime field $\mathbb{F}_p$.

In Section 4.3 we started with a presentation of $\mathbb{F}_q$ (as an algebra over $\mathbb{F}_p$) of the form $\mathbb{F}_q = \mathbb{F}_p[x, y]/(m(x), y - g_{\zeta^2}(x))$, where x, y map onto ζ^{2d} and ζ^2, respectively, and used it in order to obtain a presentation of $\mathrm{SL}(2, q)$.

Roughly speaking, short (with length $O(\log q)$) presentations of $\mathbb{F}_q$ as an algebra over $\mathbb{F}_p$ yield presentations of $\mathrm{SL}(2, q)$ with short bit-length. However in order to obtain a presentation of $\mathrm{SL}(2, q)$ with bounded expo-length using the same method, we would need a presentation of $\mathbb{F}_q$ in which every relation involves only a bounded number of monomials. This is a computational question concerning "sparse" constructions of finite fields about which little appears to be known [GaN]. The same remarks also apply to the unitary and Suzuki groups.

Appendix: Lengths

This paper has dealt with bit-lengths of presentations. We conclude with some elementary results involving lengths (rather than bit-lengths) of presentations that were alluded to in Section 1.

Notation: For a word w in the free group $F = F_{\{x,y\}}$, let $f(w)$ denote the sum of the exponents of y. Then $f : F \to \mathbb{Z}$ is a homomorphism. The symbols x and y will denote free generators of F or permutations, depending on context.

Theorem A1. *Let $G = \langle x, y\rangle$ be a transitive subgroup of S_n, where the support of x has size m. Then any presentation of G based on x and y has length greater than n/m.*

Proof. Since G is transitive, every cycle of y (viewed as a set of points) has nonempty intersection with the support $\mathrm{supp}(x)$ of x. In particular, y has at most m cycles.

Regard each cycle c of y as a circular permutation. Then c splits into $|c \cap \mathrm{supp}(x)|$ (possibly empty) arcs using the members of $c \cap \mathrm{supp}(x)$, with each arc disjoint from $\mathrm{supp}(x)$. The sum of the lengths of these arcs is $|c \setminus \mathrm{supp}(x)| = |c| - |c \cap \mathrm{supp}(x)|$. Summing over all c we see that the sum of the arc lengths for all cycles of y is $n - m$. Hence, there is some arc of length $\geq (n - m)/m$, and so some arc of even length $n' \geq (n/m) - 2$. Relabeling the permuted points, we may assume that $\{1, \ldots, n'\}$ is part of a cycle of y in the stated order, and is disjoint from $\mathrm{supp}(x)$.

If $w \in F$ has length at most $n'/2$, we claim that $w(n'/2) = y^{f(w)}(n'/2)$ when w is evaluated in G. For, if $0 \leq |i| \leq |j| \leq n'/2$ then y^i sends $n'/2$ into the complement of $\mathrm{supp}(x)$, and hence $xy^i(n'/2) = y^i(n'/2)$: the x's occurring in w do not have any affect on the calculation of $w(n'/2)$. Thus, if $w = 1$ in G then $y^{f(w)}$ fixes $n'/2$, so that $f(w) = 0$ since $|f(w)| \leq n'/2$.

In fact, if $w \in R \subseteq F$ has length $\leq n'$ then once again $f(w) = 0$. For, write $w = w_1 w_2^{-1}$ for words w_1 and w_2 of length $\leq n'/2$. We have seen that $w_i(n'/2) = y^{f(w_i)}(n'/2)$ for $i = 1, 2$. Since $w(n'/2) = n'/2$, it follows that $y^{f(w_1)}(n'/2) = y^{f(w_2)}(n'/2)$. Then $y^{f(w_1)-f(w_2)}(n'/2) = n'/2$ with $|f(w_1)|, |f(w_2)| \leq n'/2$. Thus, $f(w) = f(w_1) - f(w_2) = 0$, as claimed.

Consequently, all elements of R of length at most n' are contained in $\ker f$.

We may assume that G is not cyclic. If G has a presentation of length at most n/m, it follows that each relator has length at most $(n/m) - 2 \le n'$. By the preceding paragraph, $F/\ker f$ is a surjective image of G. However, $F/\ker f$ is infinite since it has an abelianization with y of infinite order: the surjection sending $F \to \langle y \rangle$ via $x \mapsto 1$ and $y \mapsto y$ has $\ker f$ in the kernel. This contradiction completes the proof. □

The preceding argument also works when there are more than two generators and the union of the supports of all but one has size m. However, this argument cannot be used for bit-length in place of length, and the result is false in that setting [BCLO].

The theorem implies that there is no $O(\log n)$-length presentation of S_n based on the elementary generators $(1, 2)$ and $(1, \ldots, n)$ (and then the proof is even easier). In this case we can go further:

Theorem A2. *Any presentation of S_n, $n > 2$, based on $x = (1, 2)$ and $y = (1, \ldots, n)$ has a relation of length at least $2n - 2$.*

Proof. Since there is no surjection $S_n \to \mathbb{Z}_n$, n does not divide $f(r)$ for some relation r. *We will show that each such relation has length* $\ge 2n - 2$.

Rearrange (i.e., conjugate) r into the form

$$\begin{aligned} 1 &= xy^{b_1}xy^{b_2}x \cdots y^{b_k} \\ &= x \cdot y^{b_1}xy^{-b_1} \cdot y^{b_1+b_2}xy^{-b_1-b_2} \cdots y^{b_1+\cdots+b_{k-1}}xy^{-b_1-\cdots-b_{k-1}} \cdot y^{b_1+\cdots+b_k} \end{aligned}$$

for integers b_i not divisible by n. Rearranging does not increase the length of our relator.

This expresses y^b, $b := f(r)$, as a product of $\langle y \rangle$-conjugates of x; here $y^b \ne 1$ by our choice of r. Considering supports shows that $2k \ge n$, but we claim more: $k \ge n - 1$. We have y^b written as a product of certain transpositions $(i, i + 1)$, where we view $1, \ldots, n$ mod n. Form a graph with vertices $1, \ldots, n$ and the above transpositions $(i, i + 1)$ as (undirected) edges. Suppose that some transposition $(j, j+1)$ is not an edge. Our product gets us from j to $y^b(j) = j + b$ by a path, and also from $j + 1$ to $y^b(j + 1) = j + 1 + b$ by a path. These two paths use all $n - 1$ transpositions $(s, s + 1)$ with $s \ne j$ (possibly using some of these transpositions more than once), as claimed.

Between any two of the k occurrences of x in r there is at least one occurrence of y (i.e., each $b_i \ne 0$). Hence, each of the generators x and y occurs at least k times in r, so that r has length at least $2k \ge 2n - 2$. □

This result is optimal in view of the presentation

$$S_n = \langle x, y \mid x^2 = y^n = (xy)^{n-1} = (xx^y)^3 = (xx^{y^i})^2 = 1,\ 2 \le i \le n/2 \rangle$$

(cf. [Moo, p. 358]), in which the unique relation $(xy)^{n-1} = 1$ with sum of y-exponents not divisible by n has length exactly $2n - 2$.

This argument can be pushed slightly further:

Theorem A3. *Assume that $G = \langle x, y \rangle$ does not have a surjection onto $\mathbb{Z}_{|y|}$.*

(i) *Suppose that $G \le S_n$, where the support of x has size m and the support of each nontrivial power of y has size $\ge n'$. Then any presentation of G based on x and y has a relation of length $\ge 2n'/m$.*

(ii) *Suppose that $G \le \mathrm{GL}(n, K)$ for a field K, where the support of x has dimension m and the support of each nontrivial power of y has dimension $\ge n'$. Then any presentation of G based on x and y has a relation of length $\ge 2n'/m$.*

Proof. We will prove (i); the proof of (ii) is essentially the same. Let $G = \langle x, y \mid R\rangle$. As before, $|y|$ does not divide $f(r)$ for some relation $r \in R$. *We will show that each such relation has length $\ge 2n'/m$.*

Rearrange r into the form

$$\begin{aligned}1 &= x^{a_1}y^{b_1}x^{a_2}y^{b_2}x^{a_3}\cdots y^{b_k}\\ &= x^{a_1}\cdot y^{b_1}x^{a_2}y^{-b_1}\cdot y^{b_1+b_2}x^{a_3}y^{-b_1-b_2}\cdots y^{b_1+\cdots+b_{k-1}}x^{a_k}y^{-b_1-\cdots-b_{k-1}}\cdot y^{b_1+\cdots+b_k}\end{aligned}$$

for nonzero integers a_i, b_i. Our choice of r implies that $y^{b_1+\cdots+b_k} = y^{f(r)} \neq 1$, and we have expressed $y^{f(r)}$ as a product of conjugates of k nontrivial powers of x. By hypothesis, the support of $y^{f(r)}$ has size $\ge n'$. By considering support sizes we see that $n' \le mk$.

Between any two occurrences of x in r there is at least one occurrence of y. Hence, each of the generators x and y occurs at least $k \ge n'/m$ times in r. □

Somewhat as above, this argument works when there are more than two generators, one of which behaves as y does while the support of each of the others has size at most m. Note that (i) is actually a special case of (ii).

Unlike in Theorem A1, there is no transitivity assumption concerning G; but in that theorem transitivity is the only restriction on y. Theorems A1 and A3 apply, for example, to $G = A_n$ with x a 3-cycle and y an n-cycle or the product of two $n/2$-cycles, or $y = (1, 2)(3, \dots, n)$; and to $G = S_n$, $n = 2s + 1$, with x a transposition and y the product of disjoint cycles of length s and $s + 1$.

Historical addendum

In November, 2008, we came across the following paper using Google: H. Saß, Eine abstrakte Definition gewisser alternierender Gruppen, Math. Z. **128**, 109–113 (1972). This paper does not seem to have been mentioned in any subsequent research involving presentations of groups such as PSL(2, p) or A_n. For any prime $p > 2$, Saß proved that

$$\begin{aligned}A_{p+2} \cong \langle T, C, B \mid\ & C^p = T^3 = B^{(p-1)/2} = (TC)^{p+2} = 1,\\ & (TC^{-1}TC)^2 = (TC^{-\alpha}TC^{\alpha})^2 = B^{-1}CBC^{-\alpha^2} = 1,\\ & B^{-1}TBT^{-1} = (C^{-1}TC^{1-\alpha}TC^{\alpha-1}T^2CB)^p = 1\rangle,\end{aligned}$$

where $\mathbb{F}_p^* = \langle\alpha\rangle$. (Moreover, if $p \equiv 3 \pmod 4$ then the relation $(TC^{-\alpha}TC^{\alpha})^2 = 1$ can be deleted.) Saß started with a presentation for PSL(2, p) [Fr], and then used a variant of (3.1) also due to Carmichael [Car1, p. 262]. Note that the above relations include a presentation $\langle B, C \mid C^p = B^{(p-1)/2} = 1,\ C^B = C^{\alpha^2}\rangle$ for the group AGL(1, $p)^{(2)}$ that was crucial for us, but Saß did not use this group.

Nevertheless, if Carmichael's relation $(TC)^{p+2} = 1$ is deleted and if Neumann's presentation [Neu] for AGL(1, $p)^{(2)}$ is used instead of the preceding one, then Saß's presentation becomes Examples 3.4(2), (4). In the notation used several times in Sections 3.1–3.2, Saß's relator $(C^{-1}TC^{1-\alpha}TC^{\alpha-1}T^2CB)^p$ turns out to be $(1, b)$ (this resembles (3.10)).

Thus, Saß essentially obtained one of our initial results, except for the unfortunate relation $(TC)^n = 1$ $(n = p+2)$ appearing in different notation in the presentation [Car1, p. 262] for A_n when n is odd (repeated in [Car2, p. 185, Ex. 1] and [CoMo, p. 67]). In fact, Carmichael's argument shows that his version of the relation $(TC)^n = 1$ can be deleted for all odd n.

Acknowledgments. We are grateful to George Havas for providing presentations in Example 3.21(4),(10) and Section 8.2, and to John Bray for providing presentations in Section 8.2; both provided more presentations than are actually included here, since some ultimately were not needed. We also thank Edmund Robertson for assistance with [Ro], and Bob Griess for pointing out Steinberg's criterion [St3, Corollary 7.6] used in Section 10.1.

The authors were partially supported by NSF grants DMS 0140578, DMS 0242983, DMS 0354731, DMS 0600244, DMS 0635607, DMS 0653873, DMS 0753640 and DMS 0900932. The third author was also supported by an AMS Centennial Fellowship.

References

[AFV] Armstrong, H., Forrest, B., Vogtmann, K.: A presentation for Aut(F_n). J. Group Theory **11**, 267–276 (2008) Zbl 1139.20028 MR 2396963

[Ar] Artin, E.: Theorie der Zöpfe. Abh. Math. Sem. Univ. Hamburg **4**, 47–72 (1925) JFM 51.0450.01

[BGKLP] Babai, L., Goodman, A. J., Kantor, W. M., Luks, E. M., Pálfy, P. P.: Short presentations for finite groups. J. Algebra **194**, 79–112 (1997) Zbl 0896.20025 MR 1461483

[BKL] Babai, L., Kantor, W. M., Lubotzky, A.: Small diameter Cayley graphs for finite simple groups. Eur. J. Combin. **10**, 507–522 (1989) Zbl 0702.05042 MR 1022771

[BS] Babai, L., Szemerédi, E.: On the complexity of matrix group problems, I. In: Proc. 25th IEEE Sympos. Foundations Comp. Sci., 229–240 (1984)

[Bau] Baumslag, G.: A finitely presented metabelian group with a free abelian derived group of infinite rank. Proc. Amer. Math. Soc. **35**, 61–62 (1972) Zbl 0269.20029 MR 0299662

[Behr] Behr, H.: Explizite Präsentation von Chevalley-gruppen über Z. Math. Z. **141**, 235–241 (1975) Zbl 0286.20056 MR 0422442

[BGHS] Bennett, C., Gramlich, R., Hoffman, C., Shpectorov, S.: Curtis–Phan–Tits theory. In: Groups, Combinatorics and Geometry (Durham, 2001), World Sci., 13–29 (2003) Zbl 1063.20012 MR 1993197

[BeS] Bennett, C. D., Shpectorov, S.: A new proof of Phan's theorem. J. Group Theory **7**, 287–310 (2004) Zbl 1055.20022 MR 2062999

[Br] Bray, J.: Personal communication

[BCLO] Bray, J., Conder, M. D. E., Leedham-Green, C. R., O'Brien, E. A.: Short presentations for alternating and symmetric groups. Trans. Amer. Math. Soc., to appear

[Bre] Breusch, R.: Zur Verallgemeinerung des Bertrandschen Postulates, dass zwischen x und $2x$ stets Primzahlen liegen. Math. Z. **34**, 505–526 (1932) Zbl 0003.24504 MR 1545270

[Bur] Burnside, W.: Theory of Groups of Finite Order. 2nd ed., Cambridge Univ. Press, Cambridge (1911) JFM 42.0151.02

[CHRR1] Campbell, C. M., Havas, G., Ramsay, C., Robertson, E. F.: Nice efficient presentations for all small simple groups and their covers. LMS J. Comput. Math. **7**, 266-283 (2004) Zbl 1053.20013 MR 2118175

[CHRR2] Campbell, C. M., Havas, G., Ramsay, C., Robertson, E. F.: On the efficiency of the simple groups of order less than a million and their covers. Experiment. Math. **16**, 347–358 (2007) Zbl 1133.20015 MR 2367323

[CHRR3] Campbell, C. M., Havas, G., Ramsay, C., Robertson, E. F.: All simple groups with order between 1 and 5 million are efficient. Preprint

[CR1] Campbell, C. M., Robertson, E. F.: Classes of groups related to $F^{a,b,c}$. Proc. Roy. Soc. Edinburgh Sect. A **78**, 209–218 (1977/78) Zbl 0378.20024 MR 0577063

[CR2] Campbell, C. M., Robertson, E. F.: A deficiency zero presentation for SL$(2, p)$. Bull. London Math. Soc. **12**, 17–20 (1980) Zbl 0393.20020 MR 0565476

[CR3] Campbell, C. M., Robertson, E. F.: The efficiency of simple groups of order $< 10^5$. Comm. Algebra **10**, 217–225 (1982) Zbl 0478.20024 MR 0674990

[CRKMW] Campbell, C. M., Robertson, E. F., Kawamata, T., Miyamoto, I., Williams, P. D.: Deficiency zero presentations for certain perfect groups. Proc. Roy. Soc. Edinburgh Sect. A **103**, 63–71 (1986) Zbl 0604.20035 MR 0858118

[CRW1] Campbell, C. M., Robertson, E. F., Williams, P. D.: Efficient presentations for finite simple groups and related groups. In: Groups—Korea 1988 (Pusan, 1988), Lecture Notes in Math. 1398, Springer, Berlin, 65–72 (1989) Zbl 0683.20026 MR 1032811

[CRW2] Campbell, C. M., Robertson, E. F., Williams, P. D.: On presentations of PSL$(2, p^n)$. J. Austral. Math. Soc. **48**, 333–346 (1990) Zbl 0705.20025 MR 1033184

[Car1] Carmichael, R. D.: Abstract definitions of the symmetric and alternating groups and certain other permutation groups. Quart. J. Pure Appl. Math. **49**, 226–270 (1923) JFM 48.1148.03

[Car2] Carmichael, R. D.: Introduction to the Theory of Groups of Finite Order. Ginn, Boston (1937) Zbl 0019.19702

[Cart] Carter, R. W., Simple Groups of Lie Type. Wiley, London (1972) Zbl 0248.20015 MR 0407163

[CoMo] Coxeter, H. S. M., Moser, W. O. J.: Generators and Relations for Discrete Groups. 3rd ed., Springer (1972) Zbl 0239.20040 MR 0349820

[Cur] Curtis, C. W.: Central extensions of groups of Lie type. J. Reine Angew. Math. **220**, 174–185 (1965) Zbl 0137.25701 MR 0188299

[DaL] Dalla Volta, F., Lucchini, A.: Generation of almost simple groups, J. Algebra **178**, 194–223 (1995) Zbl 0839.20021 MR 1358262

[Di] Dixon, J. D.: The probability of generating the symmetric group. Math. Z. **110**, 199–205 (1969) Zbl 0176.29901 MR 0251758

[Er] Erdős, P.: Über die Primzahlen gewisser arithmetischer Reihen. Math. Z. **39**, 473–491 (1935) Zbl 0010.29303 MR 1545512

[Fr] Frasch, H.: Die Erzeugenden der Hauptkongruenzgruppen für Primzahlstufen. Math. Ann. **108**, 229–252 (1933) Zbl 0006.24602 MR 1512847

[GaN] von zur Gathen, J., Nöcker, M.: Polynomial and normal bases for finite fields. J. Cryptology **18**, 337–355 (2005) Zbl 1183.11078 MR 2186408

[Ger] Gersten, S. M.: A presentation for the special automorphism group of a free group. J. Pure Appl. Algebra **33**, 269–279 (1984) Zbl 0542.20021 MR 0761633

[GLS] Gorenstein, D., Lyons, R., Solomon, R.: The Classification of the Finite Simple Groups. Number 3. Part I. Chapter A. Almost Simple K-groups. Math. Surveys Monogr. 40.3, Amer. Math. Soc., Providence (1998) Zbl 0890.20012 MR 1490581

[GrHNS] Gramlich, R., Hoffman, C., Nickel, W., Shpectorov, S.: Even-dimensional orthogonal groups as amalgams of unitary groups. J. Algebra **284**, 141–173 (2005) Zbl 1109.20015 MR 2115009

[Gri] Griess, R. L., Jr.: Schur multipliers of finite simple groups of Lie type. Trans. Amer. Math. Soc. **183**, 355–421 (1973) Zbl 0297.20023 MR 0338148

[GK] Guralnick, R. M., Kantor, W. M.: Probabilistic generation of finite simple groups. J. Algebra **234**, 743–792 (2000) Zbl 0973.20012 MR 1800754

[GKKL1] Guralnick, R. M., Kantor, W. M., Kassabov, M., Lubotzky, A.: Presentations of finite simple groups: a quantitative approach. J. Amer. Math. Soc. **21**, 711–774 (2008) Zbl pre05759135 MR 2393425

[GKKL2] Guralnick, R. M., Kantor, W. M., Kassabov, M., Lubotzky, A.: Presentations of finite simple groups: a cohomological and profinite approach. Groups Geom. Dynam. **1**, 469–523 (2007) Zbl 1135.20024 MR 2357481

[GKKL3] Guralnick, R. M., Kantor, W. M., Kassabov, M., Lubotzky, A.: Remarks on proficient groups. J. Algebra, to appear

[Hal] Hall, M., Jr.: Combinatorial Theory. Blaisdell, Waltham (1967) Zbl 0196.02401 MR 0224481

[Har] Harlander, J.: Closing the relation gap by direct product stabilization. J. Algebra **182**, 511–521 (1996) Zbl 0860.20023 MR 1391597

[Hav] Havas, G.: Personal communication

[Ho] Holt, D. F.: On the second cohomology group of a finite group. Proc. London Math. Soc. **55**, 22–36 (1987) Zbl 0624.20035 MR 0887282

[HEO] Holt, D. F., Eick, B., O'Brien, E. A.: Handbook of Computational Group Theory. Chapman & Hall, Boca Raton (2005) Zbl 1091.20001 MR 2129747

[HS] Hulpke, A., Seress, Á.: Short presentations for three-dimensional unitary groups. J. Algebra **245**, 719–729 (2001) Zbl 1062.20052 MR 1863898

[Ka] Kantor, W. M.: Subgroups of classical groups generated by long root elements. Trans. Amer. Math. Soc. **248**, 347–379 (1979) Zbl 0406.20040 MR 0522265

[KaLu] Kantor, W. M., Lubotzky, A.: The probability of generating a finite classical group. Geom. Dedicata **36**, 67–87 (1990) Zbl 0718.20011 MR 1065213

[KaS] Kantor, W. M., Seress, Á.: Computing with matrix groups. In: Groups, Combinatorics and Geometry, A. A. Ivanov et al. (eds.). World Sci., River Edge, NJ, 123–137 (2003). Zbl 1052.20001 MR 1994963

[LG] Leedham-Green, C. R.: The computational matrix group project. In: Groups and Computation III, W. M. Kantor and Á. Seress (eds.), de Gruyter, Berlin, 229–247 (2001) Zbl 1052.20034 MR 1829483

[LGO] Leedham-Green, C. R., O'Brien, E. A.: Constructive recognition of classical groups in odd characteristic. J. Algebra **322**, 833–881 (2009) Zbl 1181.20044 MR 2531225

[LiSh] Liebeck, M. W., Shalev, A.: The probability of generating a finite simple group. Geom. Dedicata **56**, 103–113 (1995) Zbl 0836.20068 MR 1338320

[Mi] Miller, G. A.: Abstract definitions of all the substitution groups whose degrees do not exceed seven. Amer. J. Math. **33**, 363–372 (1911) JFM 42.0163.03

[Mil] Milnor, J.: Introduction to Algebraic K-Theory. Ann. of Math. Stud. 72, Princeton Univ. Press, Princeton (1971) Zbl 0237.18005 MR 0349811

[Mol] Molsen, K.: Zur Verallgemeinerung des Bertrandschen Postulates. Deutsche Math. **6**, 248–256 (1941) Zbl 0026.10205 MR 0017770

[Moo] Moore, E. H.: Concerning the abstract groups of order $k!$ and $\frac{1}{2}k!$ holohedrically isomorphic with the symmetric and the alternating substitution groups on k letters. Proc. London Math. Soc. **28**, 357–366 (1897) JFM 28.0121.03

[Mor] Moree, P.: Bertrand's postulate for primes in arithmetical progressions. Comput. Math. Appl. **26**, 35–43 (1993) Zbl 0789.11001 MR 1228779

[Neu] Neumann, B. H.: On some finite groups with trivial multiplicator. Publ. Math. Debrecen **4**, 190–194 (1956) Zbl 0070.25603 MR 0078997

[Ph] Phan, K. W.: On groups generated by three-dimensional special unitary groups I. J. Austral. Math. Soc. Ser. A **23**, 67–77 (1977) Zbl 0369.20026 MR 0435247

[Ra] Ramanujan, S.: A proof of Bertrand's Postulate. J. Indian Math. Soc. **11**, 181–182 (1919) JFM 53.0030.02 MR 2280867

[Ro] Robertson, E. F.: Efficiency of finite simple groups and their covering groups. In: Finite Groups—Coming of Age, Contemp. Math. 45, Amer. Math. Soc., Providence, 287–294 (1985) Zbl 0581.20032 MR 0822243

[Sims] Sims, C. C.: Computation with Finitely Presented Groups. Cambridge Univ. Press, Cambridge (1994) Zbl 0828.20001 MR 1267733

[Soi] Soicher, L.: In preparation

[St1] Steinberg, R.: Generators for simple groups. Canad. J. Math. **14**, 277–283 (1962) Zbl 0103.26204 MR 0143801

[St2] Steinberg, R.: Lectures on Chevalley Groups (mimeographed notes). Yale Univ. (1967)

[St3] Steinberg, R.: Generators, relations and coverings of algebraic groups, II. J. Algebra **71**, 527–543 (1981) Zbl 0468.20038 MR 0630615

[St4] Steinberg, R.: Some consequences of the elementary relations in SL_n. In: Finite Groups—Coming of Age, Contemp. Math. 45, Amer. Math. Soc., Providence, 335–350 (1985) Zbl 0579.20038 MR 0822247

[Sun] Sunday, J. G.: Presentations of the groups $\mathrm{SL}(2, m)$ and $\mathrm{PSL}(2, m)$. Canad. J. Math. **24**, 1129–1131 (1972) Zbl 0253.20051 MR 0311782

[Suz] Suzuki, M.: On a class of doubly transitive groups. Ann. of Math. **75**, 105–145 (1962) Zbl 0106.24702 MR 0136646

[ThS] Threllfall, W., Seifert, H.: Aufgabe 84. Jahrhesber. Deutsch. Math.-Verein. **41**, 6–8 (1932)

[Ti1] Tits, J.: Les groupes de Lie exceptionnels et leur interprétation géométrique. Bull. Soc. Math. Belg. **8**, 48–81 (1956) Zbl 0072.38202 MR 0087889

[Ti2] Tits, J.: Buildings of Spherical Type and Finite BN-Pairs. Springer, Berlin (1974) Zbl 0295.20047 MR 0470099

[To] Todd, J. A.: A note on the linear fractional group. J. London Math. Soc. **7**, 195–200 (1932) Zbl 0005.05001

[War] Wardlaw, W. P.: Defining relations for certain integrally parameterized Chevalley groups. Pacific J. Math. **40**, 235–250 (1972) Zbl 0252.22017 MR 0333027

[We] Weil, A.: Courbes algébriques et variétés abéliennes. Hermann, Paris (1971) Zbl 0208.49202

[Wi] Wilson, J. S.: Finite axiomatization of finite soluble groups. J. London Math. Soc. **74**, 566–582 (2006) Zbl 1118.20019 MR 2286433