PRESENTATIONS OF FINITE SIMPLE GROUPS: A QUANTITATIVE APPROACH

R. M. GURALNICK, W. M. KANTOR, M. KASSABOV, AND A. LUBOTZKY

Abstract. There is a constant C_0 such that all nonabelian finite simple groups of rank n over $\mathbb{F}_q$, with the possible exception of the Ree groups ${}^2G_2(3^{2e+1})$, have presentations with at most C_0 generators and relations and total length at most $C_0(\log n + \log q)$. As a corollary, we deduce a conjecture of Holt: 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 $\mathbb{F}_pG$-module M.

Contents

2000 *Mathematics Subject Classification.* Primary 20D06, 20F05 Secondary 20J06.

The authors were partially supported by NSF grants DMS 0140578, DMS 0242983, DMS 0600244 and DMS 0354731. The authors are grateful for the support and hospitality of the Institute for Advanced Study, where this research was carried out. The research by the fourth author also was supported by the Ambrose Monell Foundation and the Ellentuck Fund.

1. Introduction

In this paper we study presentations of finite simple groups G from a quantitative point of view. Our main result provides unexpected answers to the following questions: how many relations are needed to define G, and how short can these relations be?

The classification of the finite simple groups states that every nonabelian finite simple group is alternating, Lie type, or one of 26 sporadic groups. The latter are of no relevance to our asymptotic results. Instead, we will primarily deal with groups of Lie type, which have a (relative) *rank* n over a field $\mathbb{F}_q$. In order to keep our results uniform, we view the alternating group A_n and symmetric group S_n as groups of rank $n-1$ over "the field $\mathbb{F}_1$ with 1 element" [Ti1]. With this in mind, we will prove the following

Theorem A. *All nonabelian finite simple groups of rank n over $\mathbb{F}_q$, with the possible exception of the Ree groups ${}^2G_2(q)$, have presentations with at most C_0 generators and relations and total length at most $C_0(\log n + \log q)$, for a constant C_0 independent of n and q.*[1]

We estimate (very crudely) that $C_0 < 1000$; this reflects the explicit and constructive nature of our presentations. The theorem also holds for all perfect central extensions of the stated groups.

The theorem is interesting in several ways. It already seems quite surprising that the alternating and symmetric groups have *bounded presentations*, i.e., with a bounded number of generators and relations independent of the size of the group, as in the theorem (this possibility was recently inquired about in [CHRR,

[1] Logarithms are to the base 2.

p. 281]). This was even less expected for the groups of Lie type such as $\mathrm{PSL}(n,q)$. Mann [Man] conjectured that *every finite simple group G has a presentation with $O(\log|G|)$ relations* (this is discussed at length in [LS, Sec. 2.3]). As a simple application of [BGKLP], this was proved in [Man] with the possible exception of the twisted rank 1 groups (the Suzuki groups ${}^2B_2(q) = \mathrm{Sz}(q)$, Ree groups ${}^2G_2(q) = R(q)$ and unitary groups ${}^2A_2(q) = \mathrm{PSU}(3,q)$; in fact, by [Suz, HS] Mann's result also holds for $\mathrm{Sz}(q)$ and $\mathrm{PSU}(3,q)$). Theorem A clearly goes much further than this result, providing an *absolute* bound on the number of relations.

The *length* of a presentation is the sum of the number of generators and the lengths of the relations as words in both the generators and their inverses. A presentation for a group G in the theorem will be called *short* if its length is $O(\log n + \log q)$. (A significantly different definition of "short", used in [BGKLP], involves a bound $O((\log|G|)^c)$, i.e., $O((n^2\log q)^c)$ if $q > 1$.) When $q > 1$, the standard Steinberg presentation [St1, St2] for a group in the theorem has length $O(n^4q^2)$ (with a slightly different bound for some of the twisted groups). In [BGKLP] (when combined with [Suz, HS] for the cases $\mathrm{Sz}(q)$ and $\mathrm{PSU}(3,q)$), the Curtis-Steinberg-Tits presentation [Cur], [Ti2, Theorem 13.32] was used to prove that, once again with the possible exception of the groups ${}^2G_2(q)$, all finite simple groups G have presentations of length $O((\log|G|)^2)$ (i.e., $O((n^2\log q)^2)$ if $q > 1$) – in fact of length $O(\log|G|)$ for most families of simple groups. Theorem A clearly also improves this substantially. Note that our $O(\log(nq))$ bound on length is optimal in terms of n and q, since there are at least $cnq/\log q$ groups of rank at most n over fields of order at most q (see Section 8). On the other hand, whereas our theorem shows that nonabelian finite simple groups have presentations far shorter than $O(\log|G|)$, [BGKLP] (combined with [Suz, HS]) showed that every finite group G with no ${}^2G_2(q)$ composition factor has a presentation of length $O((\log|G|)^3)$, where the constant 3 is best possible.

Theorem A seems counterintuitive, in view of the standard types of presentations of simple groups. We have already mentioned such presentations for groups of Lie type, but symmetric and alternating groups are far more familiar. The most well-known presentation for S_n is the "Coxeter presentation" (2.5) (discovered by Moore [Mo] ten years before Coxeter's birth). It uses roughly n generators and n^2 relations, and has length roughly n^2; others use $O(n)$ relations [Mo, Cox]. We know of no substantially shorter presentations in print. However, independent of this paper, [BCLO] obtained a presentation of S_n that is bounded and has *bit-length* $O(\log n)$; however, it is not short in the sense given above. (Section 1.2 contains a definition of bit-length, together with a discussion of various notions of lengths of presentations and their relationships. For example, it is straightforward to turn a presentation having bit-length N into a presentation having length $4N$, but generally at the cost of introducing an unbounded number of additional relations.)

Either bounded *or* short presentations for nonabelian simple groups go substantially beyond what one might expect. Obtaining both simultaneously was a surprise. By contrast, while abelian simple groups (i.e., cyclic groups of prime order) have bounded presentations, as well as ones that are short, they cannot have presentations satisfying both of these conditions (cf. Remark 1.2). In fact, this example led us to believe, initially, that nonabelian simple groups also would not have short bounded presentations. A hint that nonabelian finite simple groups behave differently from abelian ones came from the Congruence Subgroup Property, which can be used to obtain a short bounded presentation for $\mathrm{SL}(2,p)$ when p is

prime (see Section 3.1). This approach was later replaced by an elementary but somewhat detailed one (see Section 3.2), which became increasingly more intricate as we moved to $\mathrm{PSL}(2,q)$ and other rank 1 groups (Sections 4.1-4.4). The $\mathrm{PSL}(2,p)$ presentation also was used to obtain a bounded and short presentation for S_n (Section 3.4); when combined with presentations for rank 1 groups the latter was then used to prove the theorem for all groups of large rank (Section 6). Our arguments could have been considerably simplified if we had been willing to give up one of the two features of the presentations in Theorem A. A modified version of Theorem A is given in [GKKL2] that uses significantly fewer relations in presentations that are no longer short.

In order to emphasize our special interest in bounded presentations, we note the following consequence of the above theorem, which follows immediately from Lemma 2.1 together with the fact that, by the classification of the finite simple groups, every such group can be generated by two elements:

Corollary A′. *There is a constant C such that every finite simple group, with the possible exception of the Ree groups ${}^2G_2(q)$, has a presentation having 2 generators and at most C relations.*

In [GKKL2] we show that $C \leq 100$ by ignoring presentation length. The groups ${}^2G_2(q)$ are obstacles for all of these results: no short *or* bounded presentation is presently known for them. We can handle these offending exceptions by changing our focus either to cohomology or to profinite presentations. We say that a finite or profinite group G has a *profinite presentation* with d generators X and r relations Y if there exists a continuous epimorphism $\hat{F}_d \to G$, where $\hat{F}_d$ is the free profinite group on a set X of d generators, whose kernel is the minimal closed normal subgroup of $\hat{F}_d$ containing the r-element subset Y. Then in the category of profinite presentations there are no exceptions:

Theorem B. *There is a constant C such that every finite simple group G has a* profinite *presentation with 2 generators and at most C relations.*

The quantitative study of profinite presentations of finite simple groups was started in [Lub1], motivated by an attempt to prove the Mann-Pyber conjecture [Man, Py], which asserts that *the number of normal subgroups of index n of the free group F_d is $O(n^{cd\log n})$* for some constant c. Mann [Man] showed that his conjecture about $O(\log|G|)$-relation presentations for finite simple groups implies the Mann-Pyber conjecture; and hence he proved that conjecture except for the twisted rank 1 groups.

In [Lub1] the Mann-Pyber conjecture was proved by using a weaker version of Mann's conjecture: *There is a constant C such that every finite simple group G has a* profinite *presentation with 2 generators and at most $C\log|G|$ relations.* The crucial ingredient in the proof of the latter result was a theorem of Holt [Ho]: *There is a constant C such that, for every finite simple group G, every prime p and every irreducible $\mathbb{F}_pG$-module M,*

$$\dim H^2(G,M) \leq C(\log_p|G|_p)\dim M, \tag{1.1}$$

where $|G|_p$ denotes the p-part of the integer $|G|$. In fact, Holt proved that C can be taken to be 2, using tools including standard cohomology methods together with results in [AG]. Moreover, Holt [Ho] conjectured the following stronger result, which we will see is equivalent to Theorem B:

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

In a subsequent paper [GKKL1] we will show that the constant in Theorems B and B′ can be taken to be less than 20 (and is probably even smaller than this). The proof uses an interplay of cohomological *and* profinite presentation results.

More generally, Holt proved that (1.1) holds for any finite group G and any faithful irreducible $\mathbb{F}_pG$-module M; his proof reduced this to the simple group case. Similarly, using a different reduction, in [GKKL1] we use Theorem B′ to prove a stronger conjecture made by Holt:

Theorem B″ (Holt's Conjecture)**.** *There is a constant C such that, for every finite group G, every prime p and every faithful irreducible $\mathbb{F}_pG$-module M,* $\dim H^2(G,M) \leq C \dim M$.

As noted above, Theorem A does not require the classification of the finite simple groups: it only deals with groups having rank n over $\mathbb{F}_q$. Of course, by the classification this ignores only the 26 sporadic simple groups. On the other hand, Theorems B, B′ and B″ do use the classification, as does Holt's proof of (1.1). However, unlike many papers by some of the authors, no classification-dependent internal structural properties of these groups are required for the proof.

We emphasize that our methods are all essentially elementary: the needed structural properties of the groups are standard, and the only number theory used is "Bertrand's Postulate", a weak precursor to the Prime Number Theorem. Finally, we note that most of our presentations contain standard structural aspects of various simple groups: we do not eliminate or otherwise fine-tune relations in a manner standard in this subject (there are many examples of the latter process in [CoMo]).

1.1. **Outline of the proofs.** In Section 7 we will show that Theorems B and B′ are equivalent, and that they follow from Corollary A′ except for the Ree groups ${}^2G_2(q)$. In order to get a smaller constant in Theorem B, a detailed estimation of second cohomology groups is needed using (7.3). This is postponed to a later paper [GKKL1].

Sections 3-6 contain the proof of Theorem A. Groups of Lie type are built out of symmetric groups (and related Weyl groups) together with rank 1 groups of Lie type. Most of our efforts are directed towards these two classes of groups.

In Section 3 we obtain bounded presentations of length $O(\log n)$ for symmetric and alternating groups. Our approach is based on the groups $\mathrm{PSL}(2,p)$. Therefore we first have to handle that case in Sections 3.1 and 3.2. A short bounded presentation for $\mathrm{PSL}(2,p)$ already requires having the presentation encode exponents up to p in binary; this is accomplished using a trick from [BKL]: a group-theoretic version of "Horner's Rule" (3.3).

We combine our presentation for $\mathrm{PSL}(2,p)$ with the usual embedding $\mathrm{PSL}(2,p) < S_{p+1}$ in order to obtain a short bounded presentation for S_{p+2} whenever $p > 3$ is prime. Then we glue together such presentations for two copies of S_{p+2} in order to obtain a short bounded presentation for the general symmetric group S_n.

We deal with rank 1 groups (especially $\mathrm{SL}(2,q)$, $\mathrm{SU}(3,q)$ and $\mathrm{Sz}(q)$) in Section 4. For these groups we have to deal with several obstacles:

- *Borel subgroups do not have short bounded presentations.* However, we need to use a standard presentation of Steinberg [St2, Sec. 4], based on rank 1

BN-pairs, that is built from a presentation for a Borel subgroup B together with other data (cf. Theorem 4.34). (For example, in the case $\mathrm{SL}(2,q)$ the subgroup B is a semidirect product $\mathbb{F}_q^+ \rtimes \mathbb{F}_q^*$ and hence does not have a short bounded presentation; cf. Remark 1.2.)

As a substitute for a short bounded presentation for B, we find one for an *infinite* central extension of B, after which almost all of the center is killed when we include the remaining ingredients in a presentation for groups such as $\mathrm{SL}(2,q)$.

- Somewhat long relations, such as $x^p = 1$, need to be handled, once again using Horner's Rule.
- Elementary abelian subgroups of order q and arbitrary rank have to be *shown to be abelian by using only a bounded number of relations*. For this purpose we need to significantly generalize an idea due to Baumslag [Bau] and Campbell, Robertson and Williams [CRW1] so that we can even handle nonabelian subgroups of unitary and Suzuki groups.
- Longer relations, such as $x^{(q-1)/2} = 1$, have to be handled. This is accomplished primarily by using the minimal polynomial of a field element of the required order as a replacement for the exponent.

Thus, already in rank 1 we have to put significant effort into handling the restrictions imposed in Theorem A.

The rank 1 presentations, combined with the Curtis-Steinberg-Tits presentations [Cur], [Ti2, Theorem 13.32], are used in Section 5 to obtain short bounded presentations for bounded rank groups. Bounded (but not short) presentations for untwisted groups of bounded rank were obtained in [KoLu]. Starting with our rank 1 results, all bounded rank groups can be readily handled somewhat as in [KoLu] (and in fact this is essentially in [BGKLP]). Fortunately, the rank 1 group ${}^2G_2(q)$ does not arise as a Levi subgroup of any higher rank group.

Finally, in Section 6 we first consider $\mathrm{SL}(n,q)$, and then show how to extend the results from both $\mathrm{SL}(n,q)$ and bounded rank to the universal covers of all groups of Lie type (except ${}^2G_2(q)$), using the Curtis-Steinberg-Tits presentation. For all groups of large rank we go to some additional effort to increase our sets of generators and relations somewhat in order to obtain a presentation in which a suitable element of the center is a short word in the generators, and hence the center can be factored out within the constraints of Theorem A..

1.2. **The lengths to which we could go.** Table 1 contains a summary of the preceding outline, together with information concerning different notions of lengths of a presentation $\langle X \mid R\rangle$:

length = *word length*: $|X|$ + sum of the lengths of the words in R within the free group on X; thus, length refers to strings in the alphabet $X \cup X^{-1}$. This is the notion of length used in this paper. Mild variations, the same as ours up to a constant, are used by practitioners of the art of computing explicit presentations for groups. For example, [CHRR] just uses the above length-sum, while [Do] uses $|R|$ + the length-sum. As noted in [Sim, pp. 290-291]: "There is no universal agreement as to when one presentation is simpler than another. Usually it is considered good if the number of generators can be reduced or the total length of the relators can be made smaller.... The real challenge is to reduce both

the number of generators and the total length of the presentation." A similar notion appears in the discussion of length in [HEO, p. 184]: "As a measure of the complexity of a group presentation, we formally define its size to be the triple $(|X|, |R|, l)$, where l is the total relator length."

plength = *power length*: $|X|$ + total number of powers appearing in R.

blength = *bit-length*: total number of bits required to write the presentation, used in [BGKLP] and [BCLO]. In particular, all exponents are encoded as binary strings, the sum of whose lengths enters into the blength, as does the space required to enumerate the list of generators and relations. This definition seems especially suitable from the point of view of Computer Science. It also makes results similar to those given here significantly easier to prove (see [GKKL2]).

Some of our presentations can be modified so as to have these *other* lengths, as indicated in Table 1, where we write $q = p^e$ with p prime. However, we definitely do not provide presentations having all of the indicated types of lengths simultaneously.

Example 1: $\langle x \mid x^n \rangle$ has length $1 + n$, plength 2, and blength $2 + [\log n]$.

Example 2: If $k = \sum_0^m a_i 2^i$ with $a_i \in \{0, 1\}$, the power x^k can be replaced by $\prod_0^m (x^{2^i})^{a_i} = \prod_0^m x_i^{a_i}$ for new generators x_i and relations $x = x_0, x_i = x_{i-1}^2$ $(1 \le i \le m)$.

This multiplies bit-length by at most a factor of 4 in order to obtain length. Thus, we see that length and bit-length are essentially the same, *except if presentations are to be bounded, as in this paper.* For bounded presentations, bit-length can be far smaller than length. Clearly, plength is always at most length. We also note that the computer algebra system Magma stores relations as written: it does not introduce new relations (as in this example) unless asked to do so.

This example shows that a cyclic group of order n has a presentation of length $O(\log n)$. However, there is no such short presentation with a *bounded* number of relations:

Remark 1.2. *Each family of cyclic groups C_n whose orders are unbounded cannot have bounded presentations of length $O(\log n)$.* To see this, assume that C_n has a presentation with a bounded number of generators $x_1, \ldots, x_d$ and relations $w_1, \ldots, w_r$. Include the bounded number of additional relations $[x_i, x_j] = 1$, thereby turning our presentation into one in the category of abelian groups. Then every w_i can be viewed as a vector in the group $\mathbb{Z}^d$ with basis $x_1, \ldots, x_d$. We can renumber the w_i so that $w_1, \ldots, w_d$ generate a subgroup of finite index, at least n, in $\mathbb{Z}^d$. Then the vectors $w_1, \ldots, w_d$ are the rows of a matrix whose determinant is at least

TABLE 1. Different lengths

Sec.	groups	# relations	length	plength	bit-length
3	S_n, A_n	$O(1)$	$O(\log n)$	$O(1)$	$O(\log n)$
4	rank 1	$O(1)$	$O(\log q)$	$O(e)$	$O(\log q)$
5	rank n	$O(n^2)$	$O(n^2 \log q)$	$O(n^2 + e)$	$O(n^2 \log n \log q)$
6.1	$\mathrm{PSL}(n, q)$	$O(1)$	$O(\log n + \log q)$	$O(e)$	$O(\log n + \log q)$
6.2	large n	$O(1)$	$O(\log n + \log q)$	$O(e)$	$O(\log n + \log q)$

n, and which is also a polynomial of degree d in the entries of the w_i. Hence, at least one of the entries of some w_i is at least $n^{1/d}/d!$, which means that the corresponding exponent of some x_j appearing in w_i is at least that large. Consequently, the length of the presentation is at least $n^{1/d}/d!$, and hence is certainly not $O(\log n)$ since d is bounded.

It follows immediately that various other groups cannot have short bounded presentations. For example, *the* 1*-dimensional affine group* $\mathrm{AGL}(1,q)$ *cannot have a bounded presentation of length* $O(\log q)$. (Namely, if we mod out by the commutators of the generators in a presentation for this group then we obtain a presentation for a cyclic group of order $q-1$.)

1.3. **Historical remarks.** There seems to have been an effort, at least in the 1930's but also as far back as 1897, to obtain presentations for symmetric and alternating groups involving as few generators and relations as possible. For example, Moore [Mo] gave not only the Coxeter presentation (2.5) for S_n but also one with 2 generators and approximately n relations. Many such presentations are reproduced in [CoMo, Sec. 6.2]. In [CHRR, p. 281] we find the question: "Is there a two-generator presentation for A_n with k relators, where k is independent of n?" Our Theorem A, as well as [BCLO], answer this question. We note that [CHRR] lists various presentations of small simple groups, using as one of the criteria for "niceness" the short lengths of all relations.

There has been a great deal of research on presentations for relatively small simple groups, and for small-dimensional quasisimple groups. Extensive tables are given in [CoMo, pp. 134–141]. Special emphasis is given to $\mathrm{PSL}(2,p)$ and $\mathrm{PGL}(2,p)$ in [CoMo, Sec. 7.5]. An indication of the large amount of more recent work of this sort can be found in the references in [CCHR, CHRR, CR1, CRW1, CRW2].

On the other hand, there are few general references containing presentations for groups of Lie type; cf. [Cur, St1, St2, Ti2, Theorem 13.32] – described in [GLS, Sec. 2.9] – and their consequences [BGKLP, KoLu]. However, there are a few specific groups for which presentations have been published; [CRW2] and [CHLR] are typical. The only reference containing a hint in the direction of Corollary A′ is [Wi], which contains a more precise conjecture than the above question in [CHRR], namely, that the universal central extension of every finite simple group has a presentation with 2 generators and 2 relations.

Finally, we mention a recent, very different type of presentation for some groups of Lie type, initiated in [Ph] and continued very recently in [BS, GHNS, GHS]. While these methods can be used to obtain presentations of unitary groups with fewer relations than here (see [GKKL2]), in this paper we prefer to deal with all classical groups in a more uniform manner.

2. Elementary preliminaries

2.1. **Presentations.** We begin with some elementary observations concerning presentations. Functions will always act on the left, and we use the notation $g^h = h^{-1}gh$ and $[g,h] = g^{-1}h^{-1}gh = g^{-1}g^h$.

Let $d(G)$ denote the minimum number of elements needed to generate G.

Lemma 2.1. *If* $G = \langle D \rangle$ *is a finite group having a presentation* $\langle X \mid R \rangle$*, then* G *also has a presentation* $\langle D' \mid R' \rangle$ *such that* $|D'| = |D|$ *and the natural map* $F_{D'} \to G$

sends D' bijectively to D, and $|R'| = |D| + |R|$. In particular, G has a presentation with $d(G)$ generators and $d(G) + |R|$ relations.

Proof. 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. Let $V(D) = \{v_x(D) \mid x \in X\}$. Let $d \mapsto d'$ be a bijection from D to D'.

We claim that $\langle D' \mid d' = w_d(V(D')),\ r(V(D')) = 1, d \in D, r \in R\rangle$ is a presentation for G. For, let $\tilde{G}$ be the presented group, and let $\pi\colon F_{D'} \to \tilde{G}$ be the natural surjection.

If $H := \langle \pi(V(D'))\rangle \leq \tilde{G}$, then $H = \tilde{G}$ since $\tilde{G}$ is generated by the elements $\pi(d') = \pi(w_d(V(D')) \in H$, $d \in D$.

Also, H is a homomorphic image of G since H is generated by elements $\pi(v_x(D'))$, $x \in X$, satisfying the defining relations for G: if $r \in R$ then $r(\pi(v_x(D'))) = 1$ holds in $\tilde{G} = H$.

Finally, G is a homomorphic image of $\tilde{G}$ since the map $d' \mapsto d$ sends the generators of $\tilde{G}$ to generators of G satisfying the same relations as those satisfied in $\tilde{G}$ by $\pi(D')$. □

Lemma 2.2. *Let H be a subgroup of index m in G.*

(i) (a) *If G has a presentation with x generators and r relations, then H has a presentation with $m(x-1)+1$ generators and mr relations.*
   (b) *If the presentation in* (a) *has length l, then the corresponding presentation of H has length at most ml.*
   (c) *The length of an element of H with respect to the above generators of H is at most its length with respect to the original generators of G.*

(ii) *If H has a presentation with y generators and s relations, then G has a presentation with $d(G)$ generators and $md(G) + s$ relations (independent of y).*

Proof. (i) Part (a) follows from the standard Reidemeister-Schreier algorithm [MKS, Secs. 6.3, 6.4]: the given presentation for G produces an explicit presentation for H, with the stated numbers of generators and relations. Moreover, for (b) note that each relation of G of length l gives rise to m relations for H, each of length l in the generators of H (cf. [HEO, p. 184]). Finally, for (c) see [Ser, p. 58].

(ii) Let $\pi\colon F_{d(G)} \to G$ be a surjection with kernel N. As in (i), $L = \pi^{-1}(H)$ is a free group on a set X of $n = m(d(G)-1)+1$ generators. By Lemma 2.1, H has a presentation using the set $\pi(X)$ of these n generators together with $n+s$ relations. Consequently, N is the normal closure in L – and hence also in $F_{d(G)}$ – of a set of $n + s \leq md(G) + s$ elements. □

We note that we do not know a length version of part (ii) analogous to (b) or (c).

In the future we will want to identify subgroups of our target group with subgroups of the group given by a presentation:

Lemma 2.3. *Let $\pi\colon F_{X\cup Y} \to G = \langle X, Y \mid R, S\rangle$ and $\lambda\colon F_X \to H = \langle X \mid R\rangle$ be the natural surjections, where H is finite. Assume that $\alpha\colon G \to G_0$ is a homomorphism such that $\alpha\langle\pi(X)\rangle \cong H$. Then $\langle\pi(X)\rangle \cong H$.*

Proof. Clearly α sends $\langle\pi(X)\rangle$ onto $\alpha(\langle\pi(X)\rangle) \cong H$. If $r \in R$ then $r(\pi(X)) = 1$, so that λ induces a surjection $H = \langle X \mid R\rangle \to \langle\pi(X)\rangle$. □

From now on we will be less careful than in the preceding three lemmas: we will usually identify the generators in a presentation with their images in the presented group.

Proposition 2.4. *Suppose that G has a presentation $\langle X \mid R\rangle$ in which R has total 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 length of $\hat{R}$ is less than $4|X| + 2L + p|R|$, and $\hat{X}$ contains a generator of Z.*

Proof. Since $\hat{G}/Z = G$, we can lift the surjection $\phi\colon F = F_X \to G$, with kernel $N := \langle R^F\rangle$, to a map $\hat{\phi}\colon F \to \hat{G}$ (just send each $x \in X$ to a lift of $\phi(x)$ in $\hat{G}$); since Z is contained in the Frattini subgroup of $\hat{G}$, $\hat{\phi}$ is surjective. Then $K := \ker\hat{\phi} < N$ and $|N/K| = p$.

Let $\hat{X} = X\dot\cup\{y\}$. We may assume that $r_1 \in R$ is not in K, so that $N = K\langle r_1\rangle$. For each $r \in R$ choose e_r such that $0 \le e_r < p$ and $s_r(r_1) := rr_1^{e_r} \in K$. Then $K_1 := \langle \hat{R}^F\rangle \le K$, where $\hat{R} := \{yr_1^{-1}, [y,x], s_r(y) \mid x \in X, r \in R\}$.

We claim that $K_1 = K$. First note that $\langle r_1, K_1\rangle$ is normal in F since $[r_1, X] \subseteq K_1$. Then $N = \langle r_1, K_1\rangle$ since each $r = s_r(y)r_1^{-e_r} \in \langle r_1, K_1\rangle$, so that $N/K_1 = \langle r_1K_1\rangle$ with $r_1^p = s_{r_1}(y) \in K_1$. Since $K_1 \le K$ and $|N/K| = p$, it follows that $K_1 = K$, as claimed.

Clearly y represents a generator for Z. The total length of $\hat{R}$ is at most $(L+1) + 4|X| + L + (p-1)|R|$. □

2.2. **Some elementary presentations.** The most familiar presentation for the symmetric group S_n is the "Coxeter presentation" [Mo]:

$$S_n = \langle x_1, \dots, x_{n-1} \mid x_i^2 = (x_ix_{i+1})^3 = (x_ix_j)^2 = 1 \text{ for all possible } i \text{ and for } i+2 \le j \le n-1\rangle, \tag{2.5}$$

based on the transpositions $(i, i+1)$. We will often use the following presentation based on the transpositions $(1, i)$:

$$S_n = \langle x_2, \dots, x_n \mid x_i^2 = (x_ix_j)^3 = (x_ix_jx_ix_k)^2 = 1 \text{ for distinct } i,j,k\rangle. \tag{2.6}$$

This presentation is due to Burnside [Bur, p. 464] and Miller [Mil, p. 366] in 1911; Burnside describes it as probably "the most symmetrical form into which the abstract definition [of S_n] can be thrown", an idea that is crucial for our use of it in the next section. Carmichael [Carm, p. 169] observed that this presentation can be considerably shortened: only the relations $(x_ix_jx_ix_k)^2 = 1$ with $j = i+1$ need to be used. Before we learned of (2.6), we had been using the following presentation we had found, based on the same transpositions, which may be even more symmetrical than the preceding one:

$$S_n = \langle x_2, \dots, x_n \mid x_i^2 = (x_ix_j)^3 = (x_ix_jx_k)^4 = 1 \text{ for distinct } i,j,k\rangle.$$

For alternating groups, see [GKKL2].

There is an analogue of (2.6) for linear groups. The usual Steinberg presentation [St1] for $\mathrm{SL}(n,p)$, $n \ge 3$, is as follows (where e_{ij} represents the elementary matrix with 1 on the diagonal, i,j-entry 1 and all other entries 0):

$$\begin{aligned} &\textbf{Generators:} && e_{ij} \text{ for } 1 \le i \ne j \le n.\\ &\textbf{Relations:} && e_{ij}^p = 1, [e_{ij}, e_{ik}] = 1, [e_{ij}, e_{kj}] = 1, [e_{ij}, e_{kl}] = 1,\\ & && [e_{ij}, e_{jl}] = e_{il} \text{ for all distinct } i,j,k,l. \end{aligned} \tag{2.7}$$

The following variation, suggested by (2.6), allows us to use fewer generators:

Proposition 2.8. $\mathrm{SL}(n,p)$ *has the following presentation when* $n \geq 4$:

Generators: e_{1j} and e_{j1} for $2 \leq j \leq n$.

Relations: Whenever i,j,k are distinct and not 1, $e_{1j}^p = e_{i1}^p = [e_{i1}, e_{1j}]^p = 1$, $[e_{1j}, e_{1k}] = 1$, $[e_{i1}, e_{k1}] = 1$, $[[e_{i1}, e_{1j}], e_{k1}] = [[e_{i1}, e_{1j}], e_{1k}] = 1$, $[[e_{i1}, e_{1j}], e_{j1}] = e_{i1}$ and $[e_{1i}, [e_{i1}, e_{1j}]] = e_{1j}$.

Proof. Let $e_{ij} := [e_{i1}, e_{1j}]$ for $i \neq j$. Note that $[e_{i1}, e_{1j}]$ commutes with $[e_{k1}, e_{1l}]$ because it commutes with e_{k1} and e_{1l}. It follows easily that (2.7) holds. □

This motivates presentations in Theorems 6.1 and [GKKL2].

2.3. **Gluing.** The following lemmas contain the basic idea behind our methodology. We start with a subgroup T – with a suitable presentation – of our target group G, acting on some set with boundedly many orbits. We use this to prove that a small number of relations – each arising from one of these orbits – implies a very large number of relations.

Lemma 2.9. *Let* $T = \langle D \mid R \rangle$ *be a presentation of a transitive permutation group acting on* $\{1, 2, \ldots, n\}$ *such that* $D_1 \subset D$ *projects onto generators of the stabilizer* T_1 *of the point* 1. *Assume that every ordered triple of distinct points can be sent to* $(1,2,3)$ *or* $(1,2,4)$ *by* T. *Let* v_i *be words in* D *mapping onto permutations sending* $i \mapsto 1$ *for* $i = 2, 3, 4$. *Then the following is a presentation for a semidirect product* $T \ltimes S_{n+1}$ (*where* S_{n+1} *acts on* $\{0, 1, \ldots, n\}$):

Generators: D and z (viewed as the transposition $(0,1)$).

Relations: R, $z^2 = 1$, $[z, D_1] = 1$, $[z, v_2]^3 = 1$,
$([z, v_2][z, v_3])^2 = ([z, v_2][z, v_4])^2 = 1$.

Proof. Let G be the group defined by the above presentation. In view of the natural map $G \to T \ltimes S_{n+1}$ we can apply Lemma 2.3 to identify T with $\langle D \rangle$. That map sends z^T onto the set of all transpositions $(0,i)$, $1 \leq i \leq n$. Since $[z, D_1] = 1$ it follows that T acts on z^T as it does on $\{1, 2, \ldots, n\}$.

The relations $z^2 = [z, v_2]^3 = ([z, v_2][z, v_3])^2 = ([z, v_2][z, v_4])^2 = 1$, together with the assumed transitivity of T, imply that $x^2 = (xy)^3 = (xyxw)^2 = 1$ for any distinct $x, y, w \in z^T$. By (2.6), $N := \langle z^T \rangle \cong S_{n+1}$.

Clearly $G = TN$. Since G maps onto $T \ltimes S_{n+1}$, we have $G \cong T \ltimes S_{n+1}$. □

Note that, in fact, $G = C_G(S_{n+1}) \times S_{n+1} \cong T \times S_{n+1}$.

Lemma 2.10. *Let* $T = \langle D \mid R \rangle, T_1, D_1, v_i$, *be as in* Lemma 2.9. *Then the following is a presentation for a semidirect product* $T \ltimes \mathrm{SL}(n+1, p)$ (*where* T *permutes the last* n *rows and columns*):

Generators: D, e and f (viewed as e_{01} and e_{10}, cf. Proposition 2.8).

Relations: R, $[e, D_1] = [f, D_1] = 1$, and both $e, e^{v_2}, e^{v_3}, f, f^{v_2}, f^{v_3}$ and e, e^{v_2}, e^{v_4}, f, f^{v_2}, f^{v_4} satisfy the relations for $\mathrm{SL}(4,p)$ in Proposition 2.8 (where each of these ordered 6-tuples plays the role of $e_{12}, e_{13}, e_{14}, e_{21}, e_{31}, e_{41}$).

Proof. Let G be the group defined by the above presentation. By Lemma 2.3, we can identify T with the subgroup $\langle D \rangle$ of G.

The natural map $G \to T \ltimes \mathrm{SL}(n+1, p)$ sends $e^T \cup f^T$ onto the set of all matrices e_{1i}, e_{i1}, $2 \leq i \leq n$. Since $[\{e, f\}, D_1] = 1$, it follows that $|e^T \cup f^T| = 2n$ and that T acts on both e^T and f^T as it does on $\{1, 2, \ldots, n\}$.

The relations on $e, e^{v_2}, e^{v_3}, f, f^{v_2}, f^{v_3}$ and $e, e^{v_2}, e^{v_4}, f, f^{v_2}, f^{v_4}$, together with the assumed transitivity of T, imply that the hypotheses of Proposition 2.8 hold for $N := \langle e^G, f^G \rangle$. Then $N \cong \mathrm{SL}(n+1, p)$. As above, G is a semidirect product of N and T. $\square$

3. Symmetric and alternating groups

In this section we will show that S_n and A_n have bounded presentations of length $O(\log n)$. By Lemma 2.2(i), it suffices only to consider the symmetric groups.

We start by obtaining short bounded presentations for $\mathrm{PSL}(2, p)$ (in Sections 3.1 and 3.2). We use these in order to deal with the case $n = p + 2$ with $p > 3$ prime, and then deduce the general S_n from this case.

3.1. $\mathbf{SL(2, p)}$ and the Congruence Subgroup Property.

It is not at all obvious that even the groups $\mathrm{PSL}(2, p)$ have short bounded presentations. Our first such presentation is based on results concerning arithmetic groups. The idea of using the Congruence Subgroup Property to get bounded presentations of finite groups has already appeared in [BM, Lub3]. Here we show that the presentations can be also made to be short. In the next section we will give a more explicit and far more elementary presentation.

Proposition 3.1. *If p is an odd prime then $\mathrm{PSL}(2, p)$ has a short bounded presentation.*

Proof. We sketch a proof using the Congruence Subgroup Property. By [Ser1, Ser2] and [Mar, Theorem 2.6], $\mathrm{SL}(2, \mathbb{Z}[1/2])$ is finitely presented, has the Congruence Subgroup Property and all its non-central normal subgroups have finite index. In fact, $\mathrm{SL}(2, \mathbb{Z}[1/2])$ has a finite presentation based on its generators

$$u = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad t = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad h_2 = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix}. \tag{3.2}$$

Adding the additional relation $u^p = 1$ and using the above properties of normal subgroups produces a bounded presentation for $\mathrm{SL}(2, p)$ for $p > 2$. While this presentation appears not to be short, it can be converted to a bounded presentation of length $O(\log p)$ by using a group-theoretic version of *Horner's Rule* [BKL, p. 512]:

$$\begin{aligned} &\text{If } 2 < m \le p \text{ then } u^m = \textstyle\prod_i (u^{m_i})^{h_2^i} = u^{m_0} h_2^{-1} u^{m_1} h_2^{-1} u^{m_2} \cdots h_2^{-1} u^{m_k} h_2^k \\ &\text{has length } O(\log p), \text{ where } k = [\log_4 m] \text{ and } m = \textstyle\sum_{i=0}^{k} m_i 4^i \text{ in base 4.} \end{aligned} \tag{3.3}$$

Namely, since $u^{h_2} = u^4$ in (3.2),

$$u^m = u^{m_0} \cdots (u^{m_k})^{h_2^k} = u^{m_0} h_2^{-1} u^{m_1} h_2 \cdot h_2^{-2} u^{m_2} h_2^2 \cdot h_2^{-3} \cdots h_2{}^{-k} u^{m_k} h_2^k, \tag{3.4}$$

which collapses as in (3.3).

This produces a short bounded presentation of $\mathrm{SL}(2, p)$. In order to obtain one for $\mathrm{PSL}(2, p)$, we write -1 as a short word as in [BKL] (cf. Lemma 3.8) and add one further relation to kill it. $\square$

This method of *converting to a short bounded presentation* will be used often in the rest of this paper.

We note that the argument used in the proposition can be greatly generalized: the Congruence Subgroup Property (combined with Margulis' normal subgroup theorem) can be used in a similar manner to provide short bounded presentations

for various other families of finite simple groups of fixed rank over suitable fields (cf. [Lub3]).

Sunday [Sun] used the presentation in [BM], which was obtained using the Congruence Subgroup Property as in the above proposition, in order to produce the following presentation for $\mathrm{PSL}(2,p)$ of bounded plength and bit-length $O(\log p)$, though not of bounded length:

$$\langle u,t \mid u^p = 1, t^2 = (ut)^3, (u^4 t u^{(p+1)/2} t)^2 = 1\rangle.$$

There is no presentation for this group with smaller $|X|+|R|$; by [Sch], this follows from the fact that the Schur multiplier of $\mathrm{PSL}(2,p)$ has order 2. In view of (3.3), this presentation can be modified to a short bounded presentation by including h_2 as a new generator, and using the additional relations $u^{h_2} = u^4$ and $h_2 t = u^{\bar{2}}(u^2)^t u^{\bar{2}}$, where $\bar{2} = (p+1)/2$ is the integer between 1 and p such that $2 \cdot \bar{2} \equiv 1 \pmod p$.

3.2. **PSL(2, p).** We have separated these groups from the general case because we need them for symmetric groups and they are less involved than the general $\mathrm{PSL}(2,q)$. The latter groups require a lot more preparation (in Sections 4.1 and 4.3.1), due to headaches caused by field extensions. The ideas used in this section reappear in Sections 4.3 and 4.4 with many more bells and whistles.

Fix a prime $p > 3$ and a generator j of $\mathbb{F}_p^*$. Let $\bar{2}$ be as above, let $\bar{j}$ be the integer between 1 and p such that $j \cdot \bar{j} \equiv 1 \pmod p$, and define

$$h = \begin{pmatrix} \bar{j} & 0 \\ 0 & j \end{pmatrix}. \tag{3.5}$$

Theorem 3.6. *If $p > 3$ is prime then $\mathrm{PSL}(2,p)$ has the following presentation of length $O(\log p)$ with 4 generators and 8 relations, where integer exponents are handled using using (3.3):*

Generators: u, h_2, h, t (which we think of as the matrices in (3.2) and (3.5)).

Relations:

(1) $u^p = 1$.
(2) $u^{h_2} = u^4$.
(3) $u^h = u^{j^2}$.
(4) $t^2 = 1$.
(5) $h^t = h^{-1}$.
(6) $t = uu^t u$, $h_2 t = u^{\bar{2}}(u^2)^t u^{\bar{2}}$, $ht = u^{\bar{j}}(u^j)^t u^{\bar{j}}$.

Proof. Let G be the group defined by this presentation. Then $\mathrm{PSL}(2,p)$ is an epimorphic image of G. By (6), $G = \langle u, u^t\rangle$. Then G is perfect since $u^3 = [u, h_2] \equiv 1 \pmod{G'}$ by (2), while $u^p = 1$.

We claim that $Z := \langle z\rangle \le Z(G)$, where $z := h^{(p-1)/2}$. For, if we write $j^{p-1} = 1 + mp$ for some integer m, then (3) implies that $u^z = u^{(j^2)^{(p-1)/2}} = uu^{mp} = u$ by (1). Thus $[u,z] = 1$. Similarly, $[u^t, z] = 1$ since $(u^t)^{h^{-1}} = (u^t)^{j^2}$ by (4) and (5).

By (1) and (3), in G/Z the subgroup generated by u and h is isomorphic to a Borel subgroup of $\mathrm{PSL}(2,p)$, while the subgroup generated by t and h is dihedral of order $p-1$ by (3), (4) and (5). Also, by (6), if $k = 1$ or j then there are integers k', k'' and k''' such that

$$(u^k)^t = u^{k'} h^{k''} t u^{k'''}.$$

Conjugating these relations using $\langle h\rangle$ produces the same type of relation whenever $1 \le k \le p-1$ (by (3), $\langle h\rangle$ acts on the nontrivial powers of u with only two orbits, with orbit representatives u and u^j).

Now we have verified the standard Steinberg presentation for $\mathrm{PSL}(2,p)$ [St2, Sec. 4]. (That presentation, as well as more general ones that are also based on groups with a BN-pair of rank 1, will be very useful for us; see the beginning of Section 4.4 for more details.) Thus, $G/Z \cong \mathrm{PSL}(2,p)$, and G is a perfect central extension of $\mathrm{PSL}(2,p)$. The only perfect central extensions of $\mathrm{PSL}(2,p)$ are itself and $\mathrm{SL}(2,p)$, so that $G \cong \mathrm{PSL}(2,p)$ by (4). □

Remark 3.7. Changing (4) to $[t^2,u]=1$ produces a presentation for $\mathrm{SL}(2,p)$.

The presentation for $\mathrm{PSL}(2,p)$ at the end of Section 3.1 uses fewer relations.

For later use, the following observation is crucial [BKL, p. 512]:

Lemma 3.8. *If $p>3$ then every element of $\mathrm{PSL}(2,p)$ and $\mathrm{SL}(2,p)$ can be written as a word of length $O(\log p)$ in the generating set $\{u,h_2,h,t\}$ used above.*

Proof. Every element can be written as the product of a bounded number of powers of u and u^t (cf. [BKL, p. 512]). Now use (3.3) and relation (2). □

This lemma also holds for the presentation in Proposition 3.1.

3.3. $\boldsymbol{S_{p+2}}$**.** In this section we will obtain a short bounded presentation for S_{p+2} when $p>3$ is prime.

Start with a short bounded presentation of $\mathrm{PSL}(2,p)=\langle X \mid R\rangle$ such as the one in Theorem 3.6, so that X *consists of elements* u,h,h_2,t *corresponding to* (3.2) and (3.5).

In the action of $\mathrm{PSL}(2,p)$ on the projective line $\{0,1,\dots,p-1,\infty\}$, we identify $a\in\mathbb{F}_p$ with the 1-space $\langle(a,1)\rangle$ of $\mathbb{F}_p^2$, so that

$$u=(0,1,\dots,p-1)$$

and t interchanges 0 and ∞. Then

$$B:=\langle u^t,h\rangle$$

is a Borel subgroup fixing 0 while $\langle u,h\rangle$ fixes ∞. There are exactly two orbits of $\mathrm{PSL}(2,p)$ on ordered triples of points, with representatives $(0,-1,\infty)$ and $(0,-s,\infty)$ for any non-square s in $\mathbb{F}_p^*$.

We view S_{p+2} as acting on the set $\{\star,0,1,\dots,p-1,\infty\}$, with $\mathrm{PSL}(2,p)$ fixing $\star$.

Lemma 3.9. *S_{p+2} has the following presentation with 5 generators and 15 relations*:
Generators: X and z (viewed as the transposition $(\star,0)$).
Relations:

(1) R.
(2) $[z,u^t]=[z,h]=1$.
(3) $z^2=1$.
(4) $[z,t]^3=1$.
(5) $([z,t][z,u])^2=([z,t][z,u^s])^2=1$.
(6) $(zu)^{p+1}=1$.

Proof. Let G be the group defined by the above presentation. It is easy to see that G maps onto S_{p+2}. (The various relations reflect the facts that, when viewed as permutations, z commutes with elements having support disjoint from $\{\star,0\}$; the

supports of z and t have just one common point, as do the supports of z and u; and $[(\star,0)(\star,\infty)(\star,0)(\star,-1)]^2 = 1 = [(\star,0)(\star,\infty)(\star,0)(\star,-s)]^2$.)

By Lemma 2.3, G has a subgroup we can identify with $T = \langle X \rangle \cong \mathrm{PSL}(2,p)$. By (2), z commutes with the stabilizer $B = \langle u^t, h \rangle$ in T of 0.

By (2)–(5), we can use Lemma 2.9 (with t, u, u^s playing the roles of v_2, v_3, v_4: they send $\infty \mapsto 0, -1 \mapsto 0$ and $-s \mapsto 0$, respectively). Thus $N := \langle z^T \rangle \cong S_{p+2}$. Also, $N \trianglelefteq \langle T, z \rangle = G$.

Now $u = u^{p+1} \equiv (zu)^{p+1} = 1 \pmod{N}$ by (6), so that G/N is a homomorphic image of $\mathrm{PSL}(2,p)$ such that u is mapped to 1. Then $G/N = 1$ since $\mathrm{PSL}(2,p)$ is simple. □

Remark 3.10. This is a bounded presentation with bounded plength and bit-length $O(\log p)$. This presentation can be modified to a presentation of length $O(\log p)$ with $O(\log p)$ relations (cf. Section 1.2, Example 2). However, while it is bounded *it is not short* due to the long relation $(zu)^{p+1} = 1$ (note that u^s can be made short using (3.3)). Nevertheless, we will use the above presentation to obtain a short bounded presentation in the next theorem: we will *deduce* this long relation from shorter ones by using a second copy of $\mathrm{PSL}(2,p)$.

Remark 3.11. One of our main methods has been to use (nearly) multiply transitive permutation groups in order to "move" some relations and hence to deduce many others. For example, Lemma 2.9 is concerned with groups that are 3-transitive or close to 3-transitive. We will use this idea again in Section 6.1 and [GKKL2]. Moreover, in [GKKL1] and [GKKL2] we even use sharply 2-transitive groups for this purpose rather than more highly transitive groups.

In preliminary versions of the above lemma, when we sought bounded presentations but had not yet approached ones that are both bounded and short, we proceeded somewhat differently. For example, we used a simple and standard bounded presentation for the 3-transitive group $\mathrm{PGL}(2,p)$ instead of using its subgroup $\mathrm{PSL}(2,p)$.

An even earlier approach to symmetric groups started with a much more complicated special case, but was nevertheless interesting. First we obtained bounded presentations for $\mathrm{SL}(k,2)$ for infinitely many k, as follows. The Steinberg group $\mathrm{St}_4(\mathbb{Z}\langle x_1,\ldots,x_d\rangle)$ over the free associative (non-commutative) ring on d variables is finitely presented for every $d \geq 4$ [KM]. Consequently, if R is a boundedly presented quotient ring of $\mathbb{Z}\langle x_1,\ldots,x_d\rangle$, then $\mathrm{St}_4(R)$ is boundedly presented. Note that the matrix algebra $M_t(\mathbb{F}_q)$ is boundedly presented (e.g., $M_t(\mathbb{F}_p) = \langle A, B \mid A^t = B^t = 0,\ BA + A^{t-1}B^{t-1} = 1,\ p1 = 0\rangle$; alternatively, a bounded presentation for the general $M_t(\mathbb{F}_q)$ can be obtained using the fact that it is a cyclic algebra). Therefore, $\mathrm{St}_4(M_t(\mathbb{F}_q)) \cong \mathrm{St}_{4t}(\mathbb{F}_q) \cong \mathrm{SL}(4t,q)$ also is boundedly presented. Since $\mathrm{SL}(4t,2)$ has boundedly many orbits on 4-sets of nonzero vectors, we can proceed in a manner similar to the proof of Lemma 2.9 in order to obtain a short presentation for $S_{2^{4t}-1}$ by using the presentation (2.5) instead of (2.6).

Remark 3.12. We *need the following additional elements of* $\mathrm{PSL}(2,p)$ for future use:

- k of order $(p+1)/2$, acting with two cycles on the projective line (this is a generator of a non-split torus), and

- l such that 1 and l are representatives of the two $B, \langle k\rangle$ double cosets in $\mathrm{PSL}(2,p)$. Thus,

$$\langle X\rangle = B\langle k\rangle \cup Bl\langle k\rangle. \tag{3.13}$$

In order to relate these to our previous generators, we use Lemma 3.8 in order to *obtain two short relations expressing k and l in terms of X.*

We emphasize that *the order of k follows from the relations in* Theorem 3.6. We now use this observation.

Remark 3.14. The element zu acts as a $(p+1)$-cycle $(\star,0)(0,1,2,\dots,p-1) = (\star,0,1,2,\dots,p-1)$, and $z^t(zu)z^t$ is the $(p+1)$-cycle $(0,1,2,\dots,p-1,\infty)$. (Note that we are multiplying permutations from right to left.) Let σ be a permutation of the points of this line such that

$$k^\sigma = \left(z^t(zu)z^t\right)^2 \text{ and } \sigma = (\star)(0)(1,l(0),\dots).$$

Such a permutation exists because k fixes $\star$ and acts as the product of two disjoint $(p+1)/2$-cycles with representatives 0 and $l(0)$, while $(z^t(zu)z^t)^2$ fixes $\star$ and acts as the product of two disjoint $(p+1)/2$-cycles with representatives 0 and 1. (N.B. – It seems that σ cannot be expressed as a short word in our generating set, but this will not cause any difficulties because σ will not be used in any explicit manner.) Thus we have a relation

$$(zu)^2 = z^t k^\sigma z^t.$$

With all of this notation, we are ready to prove the following

Theorem 3.15. *S_{p+2} has a presentation of length $O(\log p)$ with 9 generators and 26 relations.*

Proof. We will use two copies $\langle X \mid R\rangle$ and $\langle X' \mid R'\rangle$ of the presentation for $\mathrm{PSL}(2,p)$ obtained in Theorem 3.6 and Remark 3.12 (where $X \cap X' = \emptyset$); the natural map $x \mapsto x'$, $x \in X$, extends to an isomorphism "prime":$\langle X \mid R\rangle \to \langle X' \mid R'\rangle$. We will show that S_{p+2} has the following presentation:

Generators: $X \cup X'$ and z. (We view z as $(\star,0)$ and X' as X^σ in the preceding remark.)

Relations:

(1) $R \cup R'$.
(2) $[z,u^t] = [z,h] = [z,u'^{t'}] = [z,h'] = 1$.
(3) $z^2 = 1$.
(4) $[z,t]^3 = 1$.
(5) $([z,t][z,u])^2 = ([z,t][z,u^s])^2 = 1$.
(6) $[z,u^{-1}] = [z,l'^{-1}]$ and $(zu)^2 = z^t k' z^t$.

Relations (6) are especially important since they link the two copies for $\mathrm{PSL}(2,p)$. Note that the first relation can be rewritten $z^{u^{-1}} = z^{l'^{-1}}$.

Let G be the group defined by this short, bounded presentation. By the preceding remark, there is a surjection $\pi\colon G \to S_{p+2}$. (Note that this uses the permutation σ that is not actually in the presentation! The first relation in (6) is based on the fact that both u^{-1} and $(l^\sigma)^{-1}$ conjugate $(\star,0)$ to $(\star,1)$ since $l^\sigma(0) = \sigma^{-1}(l(0)) = 1$.)

We claim that X *and z satisfy the relations in* Lemma 3.9. This is clear for Lemma 3.9(1)-(5). Moreover, the present relation (6) implies Lemma 3.9(6):

$$(zu)^{p+1} = (z^t k' z^t)^{(p+1)/2} = z^t k'^{(p+1)/2} z^t = 1$$

since $k'^{(p+1)/2} = 1$ follows from the relations R'. Thus, by Lemma 2.3 we can identify $G^* := \langle z, X\rangle$ with S_{p+2} in such a way that z is a transposition and $|z^{\langle z,X\rangle}| = p+1$. Then $G^* = \langle z^{\langle z,X\rangle}\rangle$.

We can also identify $\langle X'\rangle$ with $\mathrm{PSL}(2,p)$. In place of (3.13) we will use $\langle X'\rangle = B'\langle k'\rangle \cup B'l'^{-1}\langle k'\rangle$, where $B' = \langle u'^{t'}, h'\rangle$. Since z commutes with B' by (2), it follows that

$$z^{\langle X'\rangle} = z^{\langle k'\rangle} \cup (z^{l'^{-1}})^{\langle k'\rangle} = z^{\langle k'\rangle} \cup (z^{u^{-1}})^{\langle k'\rangle} \subseteq z^{\langle z,X\rangle}$$

by both parts of (6). Then

$$p+1 = |\pi(z^{\langle z,X'\rangle})| \le |z^{\langle z,X'\rangle}| \le |z^{\langle z,X\rangle}| = p+1,$$

so that $z^{\langle z,X'\rangle} = z^{\langle z,X\rangle}$ and hence $G^* = \langle z^{\langle z,X\rangle}\rangle = \langle z^{\langle z,X,X'\rangle}\rangle = \langle z^G\rangle \trianglelefteq G$.

Clearly, G/G^* is a homomorphic image of $\langle X'\rangle \cong \mathrm{PSL}(2,p)$ in which k' is mapped to 1, since $k' \in \langle X, z\rangle = G^*$ by (6). Therefore $G/G^* = 1$ since $\mathrm{PSL}(2,p)$ is simple.

Replace k' and l' by the words representing them in order to obtain a presentation with 9 generators and 26 relations. □

We need further properties of the preceding generators in order to handle general symmetric groups:

Lemma 3.16. *The following elements of S_{p+2} can be written as words of length $O(\log p)$ in $X \cup X'$ whenever $1 \le i \le p-1$:*

$$\begin{aligned} &\pi_i := (\star, 0, 1, \dots, i) && \lambda_i := (i+1, \dots, p-1, \infty)\\ &\theta_i := (i, i+1) && \tau := (\star, \infty)\\ &\theta := (p-1, \infty). \end{aligned}$$

Proof. In each case we will express the permutation in terms of elements that can be moved into one of our copies of $\mathrm{PSL}(2,p)$ and hence can be written as short words by Lemma 3.8. We start with the elements

$$u = (0,1,\dots,p-1), k' = (z^t(zu)z^t)^2 = (0,1,\dots,p-1,\infty)^2,\ z = (\star,0),\ t = (\infty,0)\cdots$$

in $X \cup X'$. Then

$$\begin{aligned} &\tau = z^t && \theta\ = (0,\infty)^u = (z^u)^\tau && (\star,1) = uzu^{-1}\\ &(0,1) = (uzu^{-1})^z && \theta_i = u^i(uzu^{-1})^z u^{-i}. \end{aligned}$$

If $1 \le 2j-1 \le p-1$, then

$$(\infty, 0, 1, \dots, 2j-1) = (0,1,\dots,p-1)^{2j}(0,1,\dots,p-1,\infty)^{-2j} = u^{2j}k'^{-j},$$

so that

$$\begin{aligned} &\pi_{2j-1} = z^t(\infty,0,1,\dots,2j-1)z^t = z^t u^{2j} k'^{-j} z^t\\ &\pi_{2j} = \pi_{2j-1}\theta_{2j-1}\\ &\lambda_{p-1-2j} = (0,1,\dots,p-1,\infty)^{-2j}(0,1,\dots,p-1)^{2j} = k'^{-j}u^{2j}\\ &\lambda_{(p-1-2j)-1} = \theta_{p-1-2j}\lambda_{p-1-2j}. \end{aligned}$$

By Lemma 3.8, the elements $u^i, u^{2j} \in \langle X\rangle$ and $k'^j \in \langle X'\rangle$ can be expressed as short words in our generators. □

Note that the same argument shows that all cycles $(i, i+1, \dots, j)$, and all permutations with bounded support, can be written as short words.

3.4. $\boldsymbol{S_n}$. Finally, we consider all symmetric groups:

Theorem 3.17. *S_n has a presentation of length $O(\log n)$ with 18 generators and 58 relations.*

Proof. By Bertrand's Postulate, if $n \geq 6$ then there is a prime p such that $(n-2)/2 < p < n-2$. Then $i := 2(p+2) - n - 2 \geq 1$. We will use two copies $\langle X \mid R\rangle$ and $\langle \tilde{X} \mid \tilde{R}\rangle$ of the presentation of S_{p+2} in Theorem 3.15 (where $X \cap \tilde{X} = \emptyset$), and identify them along a subgroup S_{i+2}.

The map $x \mapsto \tilde{x}$, $x \in X$, extends to an isomorphism "tilde":$\langle X \mid R\rangle \to \langle \tilde{X} \mid \tilde{R}\rangle$. Then the elements $\lambda_i, \theta_i, \tilde{\lambda}_i, \tilde{\tau}$ are as in the preceding lemma. We will show that S_n has the following presentation:

Generators: $X, \tilde{X}$.

Relations:

(1) $R \cup \tilde{R}$.
(2) $z = \tilde{z}$, $\pi_i = \tilde{\pi}_i$.
(3) $[\lambda_i, \tilde{\lambda}_i] = [\lambda_i, \tilde{\tau}] = [\theta_i, \tilde{\lambda}_i] = [\theta_i, \tilde{\tau}] = 1$.

Let G be the group defined by this bounded presentation of length $O(\log n)$. By Lemma 2.3, G is generated by two copies $\langle X\rangle$ and $\langle \tilde{X}\rangle$ of S_{p+2}. These are linked as follows:

$$\begin{array}{lllllll} & & i+1, & i+2, & \dots & p-1, & \infty \\ \star, 0, 1, 2, \dots, i-1, i, & & & & & & \\ & & \widetilde{i+1}, & \widetilde{i+2}, & \dots & \widetilde{p-1}, & \widetilde{\infty} \end{array} \tag{3.18}$$

Namely, the first copy of S_{p+2} acts on the first and second rows, and the second copy on the second and third rows; the relations (2) identify the intersection, which is S_{i+2} acting on the second row. Thus, we will abuse notation and identify $s = \tilde{s}$ for $s \in \{\star, 0, \dots, i\}$. Note that $\mathrm{Sym}\{i, i+1, i+2, \dots, p-1, \infty\} = \langle \lambda_i, \theta_i\rangle \cong S_{p-i+1}$ commutes with $\mathrm{Sym}\,\{\widetilde{i+1}, \dots, \widetilde{p-1}, \widetilde{\infty}, \star\} = \langle \tilde{\lambda}_i, \tilde{\tau}\rangle \cong S_{p-i+1}$, by (3), since $i \geq 1$.

In order to prove that this defines S_n we will use the Coxeter presentation (2.5) with the following ordering of our n points:

$$\infty, p-1, \dots, i+1, i = \tilde{i}, i-1, \dots, 2, 1, 0, \star, \widetilde{\infty}, \widetilde{p-1}, \dots, \widetilde{i+1}. \tag{3.19}$$

The first copy of S_{p+2} acts from ∞ to $\star$, and the second copy acts from $i = \tilde{i}$ to $\widetilde{i+1}$.

For any adjacent points a, b in (3.19), there is a well-defined element $g_{(a,b)}$ of G that acts as a transposition in at least one of our two symmetric groups S_{p+2}. The elements $g_{(a,b)}$ generate G since they generate both symmetric groups.

The natural surjection $G \to S_n$ maps $g_{(a,b)}$ to the transposition (a, b). In order to prove that $G \cong S_n$, it suffices to prove that the elements $g_{(a,b)}$ satisfy the Coxeter relations (2.5).

TABLE 2. Pairs of pairs of points

a, b	c, d	reason
$\infty, \dots, i$	$\infty, \dots, \star$	S_{p+2}
$\infty, \dots, i$	$\star, \dots, \widetilde{i+1}$	$[\langle \lambda_i, \theta_i\rangle, \langle \tilde{\lambda}_i, \tilde{\tau}\rangle] = 1$ by (3)
$i, i-1, \dots, \widetilde{i+1}$	$i, i-1, \dots, \widetilde{i+1}$	S_{p+2}

Consider four points a, b, c, d occurring in this order in (3.19), with a, b and c, d two distinct pairs of adjacent points. (We allow the possibility $b = c$.) We tabulate the various cases in Table 2, including the reason that the required relation in (2.5) holds.

Thus, we obtain a short presentation for S_n with $2*9$ generators and $2*26+2+4$ relations. □

Corollary 3.20. *S_n has a bounded presentation with bounded plength.*

Proof. Use the preceding presentation *without* converting the various powers of u and k' into short words in the proof of Lemma 3.16. (Of course, we should also use Lemma 3.9 in place of Theorem 3.15, since the latter is not relevant here.) □

Remark 3.21. Neither n nor $\log n$ appears explicitly in this presentation. However, n is encoded in the presentation (in binary): p certainly is, and i is encoded through the word for λ_i in Lemma 3.16. In other words, n can be reconstructed from p and i, which are visible in the presentation.

Remark 3.22. For use in Section 6.1, for $n \geq 8$ we calculate a number of additional elements of S_n that can be written as short words in $X \cup \tilde{X}$.

(i) In Section 6.1 we will use the following *renumbering of our points*:

$$\star, 0, 1, \ldots, i-1, \widetilde{i+1}, \ldots, \widetilde{p-1}, \widetilde{\infty}, i, \ldots, p-1, \infty \;=\; \mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{5}, \ldots, \mathbf{n},$$

where $\mathbf{n} := n+1$: we will be considering the stabilizer of $\mathbf{1}$ in the symmetric group $S_{\mathbf{n}} = \mathrm{Sym}\{\mathbf{1}, \ldots, \mathbf{n}\}$. Then the permutations

$$z = (\mathbf{2}, \mathbf{3}) \qquad \sigma := \pi_i \tilde{\lambda}_{i-1} \lambda_{i-1} = (\mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{5}, \ldots, \mathbf{n})$$
$$z^{\sigma^{-4}} = (\mathbf{6}, \mathbf{7}) \qquad z z^{\sigma^{-1}} z^{\sigma^{-2}} = (\mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{5})$$
$$(z z^{\sigma^{-1}} z^{\sigma^{-2}} z^{\sigma^{-3}})^{-1} \sigma = (\mathbf{6}, \mathbf{7}, \ldots, \mathbf{n}) \qquad (z z^{\sigma^{-1}} z^{\sigma^{-2}} z^{\sigma^{-3}} z^{\sigma^{-4}})^{-1} \sigma = (\mathbf{7}, \ldots, \mathbf{n})$$

can be written as words of length $O(\log \mathbf{n})$ in our generators.

(ii) We will also need information concerning parities of some of our generators: the stabilizer of the point $\star = \mathbf{2}$ is generated by a set $X \cup \tilde{X}$ of even permutations (the generators of our two copies of $\mathrm{PSL}(2, p)$) together with the odd permutation $y := z^{\sigma^{-1}} = (\mathbf{3}, \mathbf{4})$ that we *will now include among our $18+1$ generators.*

(iii) Since $y y^{\sigma^{-1}} = (\mathbf{3}, \mathbf{4}, \mathbf{5})$ we see that, in the alternating group $A_{\mathbf{n-1}}$ on $\{\mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{5}, \ldots, \mathbf{n}\}$, $\langle X, \tilde{X}, y y^{\sigma^{-1}} \rangle$ is the stabilizer of $\mathbf{2}$, and $\langle u^t, h, \tilde{X}, (y y^{\sigma^{-1}})^{\sigma^{-1}} \rangle$ $= \langle B, \tilde{X}, (y y^{\sigma^{-1}})^{\sigma^{-1}} \rangle$ is the stabilizer of both $\mathbf{2}$ and $\mathbf{3}$.

Corollary 3.23. *Every perfect central extension of A_n has a presentation with a bounded number of relations and length $O(\log n)$.*

Proof. For A_n this follows from Theorem 3.17 and Lemma 2.2(i). When $n > 6$ the center of the universal perfect central extension of A_n has order 2 [GLS, Theorem 5.2.3], so that Proposition 2.4 can be applied. □

4. Rank 1 groups

Simple groups of Lie type are built out of two pieces: rank 1 subgroups such as $\mathrm{SL}(2, q)$, and Weyl groups such as S_n. Therefore, as stated earlier, we have placed special emphasis on these types of groups.

4.1. **The BCRW-trick.** The following lemma is based on an idea in [CRW1], which [KoLu] called *the CRW-trick*. A year after first using this we learned that a similar idea had been used earlier (for a different goal) in [Bau]. We therefore rename this *the BCRW-trick*. This makes remarkably effective use of a simple observation: if u, v, w are group elements such that $w = uv$, then any element commuting with two of them also commutes with the third. This observation is used to show that, under suitable conditions, a subgroup can be proven to be *abelian* using a very small number of relations (see Section 4.2 for a different viewpoint and a generalization):

Lemma 4.1. *Let G be a group generated by elements u, a, b satisfying the following relations*:

(1) $[a, b] = 1$;
(2) $[u, u^a] = 1$ *and* $[u, u^b] = 1$; *and*
(3) *each of* u, u^a, u^b *can be expressed as a word in the other two.*

Then $\langle u^{\langle a,b\rangle}\rangle$ *is abelian.*

It is clear that (3) and the first part of (2) imply the second part. We have stated the result in the present manner in order to emphasize the symmetry between a and b.

Proof. Define $u_{s,t} := u^{a^s b^t}$ for all integers s, t. It suffices to prove that

(4.2) $u = u_{0,0}$ *and* $u_{s,t}$ *commute for all integers* s, t.

First note that, by (1) and (3),

(4.3) *If u commutes with two members of a triple $u_{i,j}$, $u_{i+1,j}$, $u_{i,j+1}$ then it commutes with the third.*

Also, *if u commutes with $u_{i,j}$ then $u^{b^{-j}a^{-i}} = u_{-i,-j}$ commutes with u* (using (1)). It follows that

(4.4) *If u commutes with two members of a triple $u_{i,j}$, $u_{i-1,j}$, $u_{i,j-1}$ then it commutes with the third.*

For, u commutes with two members of the triple $u_{-i,-j}$, $u_{-i+1,-j}$, $u_{-i,-j+1}$, so by (4.3) it commutes with the third and hence with the third member of the triple in (4.4).

We will prove (4.2) by induction on $n := |s| + |t|$.

When $n = 1$ our hypothesis states that $u = u_{0,0}$ commutes with $u_{1,0}$ and $u_{0,1}$; these and $u_{-1,0}$ and $u_{0,-1}$ are the only possible elements having $n = 1$, and we have already seen that u commutes with the preceding two elements.

Assume that (4.2) holds for some $n \geq 1$ and consider nonnegative integers s, t such that $s + t = n + 1$.

Case: $0 = s \leq t = n + 1$. By induction, u commutes with the last two members of the triple $u_{1,n}$, $u_{0,n}$, $u_{1,n-1}$, hence by (4.4) with the first. Then, by induction and (4.3), u commutes with the first two members of the triple $u_{0,n}$, $u_{1,n}$, $u_{0,n+1}$, hence with the third.

Case: $1 \leq s \leq t$. By induction and (4.4), u commutes with the last two members of the triple $u_{s,t}$, $u_{s-1,t}$, $u_{s,t-1}$, hence with the first.

By the symmetry of our hypotheses, this completes the proof when $s, t \geq 0$, and hence also when $s, t \leq 0$. By symmetry, it remains to consider the case $s < 0 < t$, where we use (4.3) and the triple $u_{s,t-1}, u_{s+1,t-1}, u_{s,t}$. $\square$

We will need various nonabelian versions of the preceding lemma. The most straightforward one is as follows:

Lemma 4.5. *Let U_0 and W_0 be subgroups of a group G, and let u, w, a, b be elements of G satisfying the following conditions*:

(1) $[a, b] = 1$;
(2) *each of u, u^a, u^b can be expressed as a word in the other two together with elements of U_0;*
(3) *each of w, w^a, w^b can be expressed as a word in the other two together with elements of W_0;*
(4) $[u, w] = [u^a, w] = 1$;
(5) $U_0^a = U_0^b = U_0$ *and* $W_0^a = W_0^b = W_0$; *and*
(6) $[U_0, w] = [u, W_0] = 1$.

Then $[\langle u^{\langle a,b\rangle}\rangle, \langle w^{\langle a,b\rangle}\rangle] = 1$.

This time we have not preserved the symmetry between a and b; but note that $[u^b, w] = 1$ by (2), (4) and (6).

Proof. Define $u_{s,t} := u^{a^s b^t}$ and $w_{s,t} := w^{a^s b^t}$ for all integers s, t. It suffices to prove that

(4.6) $\quad u_{s,t}$ *and* w *commute for all integers* s, t.

The proof parallels that of Lemma 4.1. This time,

(4.7) *If u commutes with two members of a triple $w_{i,j}, w_{i+1,j}, w_{i,j+1}$ then it commutes with the third.*
If w commutes with two members of a triple $u_{i,j}, u_{i+1,j}, u_{i,j+1}$ then it commutes with the third.

For the first of these, by (1), (3) and (5) each of the three indicated elements is a word in the others together with elements of $W_0^a = W_0^b = W_0$; now use (6). Similarly, the second assertion follows from (1), (2), (5) and (6).

Even more trivially, in view of (1), *if u commutes with $w_{i,j}$ then $u^{b^{-j}a^{-i}} = u_{-i,-j}$ commutes with w*; and *if w commutes with $u_{i,j}$ then $w^{b^{-j}a^{-i}} = w_{-i,-j}$ commutes with u*. It follows that

(4.8) *If w commutes with two members of a triple $u_{i,j}$, $u_{i-1,j}$, $u_{i,j-1}$ then it commutes with the third.*

Namely, u commutes with two members of the triple $w_{-i,-j}$, $w_{-i+1,-j}$, $w_{-i,-j+1}$, hence, by (4.7), u commutes with the third, and hence w commutes with the third member of the triple in (4.8).

As before we prove (4.6) by induction on $n := |s| + |t|$.

When $n = 0$, $u = u_{0,0}$ and $w = w_{0,0}$ commute by (4).

Consider the case $n = 1$. By (4), w commutes with the first two members of the triple $u_{0,0}$, $u_{1,0}$, $u_{0,1}$, and hence by (4.7) with $u_{0,1}$. Then w commutes with the first and last members of the triple $u_{0,1}$, $u_{-1,1}$, $u_{0,0}$, and hence by (4.8) with $u_{-1,1}$. Then w also commutes with the last two members of the triple $u_{-1,0}$, $u_{0,0}$, $u_{-1,1}$,

and hence by (4.7) with $u_{-1,0}$. Finally, w commutes with the first two members of the triple $u_{0,0}$, $u_{-1,0}$, $u_{0,-1}$, and hence by (4.7) with $u_{0,-1}$.

Assume that (4.6) holds for some $n \geq 1$, and consider nonnegative integers s, t such that $s + t = n + 1$.

Case: $0 = s \leq t = n + 1$. By induction, w commutes with the last two members of the triple $u_{1,n}$, $u_{0,n}$, $u_{1,n-1}$, and hence by (4.8) with the first. Then, by induction and (4.7), w commutes with the first two members of the triple $u_{0,n}$, $u_{1,n}$, $u_{0,n+1}$, hence with the third.

Case: $1 \leq s \leq t$. By induction and (4.8), w commutes with the last two members of the triple $u_{s,t}$, $u_{s-1,t}$, $u_{s,t-1}$, hence with the first.

By the symmetry of our hypotheses, this completes the proof when $s, t \geq 0$. When $s, t \leq 0$, proceed as in the above two cases but negate all subscripts. By symmetry, it remains to consider the case $s < 0 < t$, where we use (4.7) and the triple $u_{s,t-1}, u_{s+1,t-1}, u_{s,t}$ to prove that w commutes with $u_{s,t}$. □

In the proof of Proposition 4.25 we will use this lemma for more than one choice of U_0 and W_0. The simplest use has $W_0 = 1$ and w in the abelian group U_0.

4.2. **Some combinatorics behind the BCRW-trick.** The proof of Lemma 4.1 was concerned primarily with triples of subscripts. In order to handle Suzuki and Ree groups, we will need to generalize this subscriptology. For these purposes, and in order both to see the connection with the original BCRW Trick and to *use* the original one, we state the desired situation in some generality.

Definition 4.9. Let $\mathcal{M}$ be a finite collection of finite sets of vectors in $\mathbb{Z}^k$.

(1) A set $S \subseteq \mathbb{Z}^k$ is called *closed* provided that, for any $v \in \mathbb{Z}^k$ and any $M \in \mathcal{M}$, if S contains all but perhaps one of the vectors in the translate $v+M = \{v+m \mid m \in M\}$ then S contains the entire translate.

(2) Fix a length function $|v| \geq 0$ for vectors $v \in \mathbb{Z}^k$ such that each ball about 0 contains only finitely many vectors.

A vector $v \in \mathbb{Z}^k$ is called *reducible* if there exist $u \in \mathbb{Z}^k$ and $M \in \mathcal{M}$ such that v is the *unique* longest vector in $u + M$. In this case we will say that v *is reducible via* $u + M$.

A vector is called *irreducible* if it is not reducible.

(3) The collection $\mathcal{M}$ is called of *finite type* (with respect to the length function $|\cdot|$) if there are only finitely many irreducible vectors in $\mathbb{Z}^k$.

(4) If a collection is denoted $\mathcal{M}_{\mathcal{X}}$, then $\mathcal{M}_{-\mathcal{X}}$ denotes $\{-M \mid M \in \mathcal{M}_{\mathcal{X}}\}$.

Example 4.10. In order to relate this to Lemmas 4.1 and 4.5, we use $\mathcal{M}_{\mathrm{BCRW}} = \big\{\{(0,0),(1,0),(0,1)\}\big\}$, consisting of a single triple in $\mathbb{Z}^2$, together with

$$\mathcal{M}_{\pm\mathrm{BCRW}} = \mathcal{M}_{\mathrm{BCRW}} \cup \mathcal{M}_{-\mathrm{BCRW}} = \big\{\{(0,0),(1,0),(0,1)\}, \{(0,0),(-1,0),(0,-1)\}\big\}.$$

The two collections $\mathcal{M}_{\pm\mathrm{BCRW}}$ arise naturally in Lemma 4.1, as follows.

Let $u = v$ and the group $H := \mathbb{Z}^2$ acts on $U = \langle u^{\langle a,b\rangle}\rangle$ via $u^{(i,j)} = u^{a^i b^j}$. The set $\mathcal{M}_{\mathrm{BCRW}} = \{M\}$ corresponds to the relation $u^b = uu^a$, and hence to the fact that each element in $\{u, u^a, u^b\} = u^M = \{u^m \mid m \in M\}$ lies in the subgroup generated by the remaining elements (as in Lemma 4.1(3)). If $v \in H$ then $(u^M)^v = u^{v+M}$. Hence, the translates of $\mathcal{M}_{\mathrm{BCRW}}$ are exactly the triples of subscripts appearing in (4.3), while the translates of $\mathcal{M}_{\pm\mathrm{BCRW}}$ are exactly the triples of subscripts appearing in (4.3) and (4.4).

Let $S := \{h \in H \mid [u^h, u] = 1\}$. Since $u^{v+(0,1)} = u^v u^{v+(1,0)}$ for any $v \in H$, if u commutes with two of the elements $u^{v+(0,1)}$, u^v, $u^{v+(1,0)}$ then it commutes with the third, so that S is closed under $\mathcal{M}_{\mathrm{BCRW}}$. Moreover, if $[u^h, u] = 1$ then $[u^{-h}, u] = 1$, so that $S = -S$, and hence S is also closed under $\mathcal{M}_{-\mathrm{BCRW}}$.

We used the length function $|(x, y)| = |x| + |y|$ in the proof of Lemma 4.1. Moreover, the way we used (4.3) and (4.4) emphasized the fact that $\mathbb{Z}^2$ is the smallest $\mathcal{M}_{\pm\mathrm{BCRW}}$-closed subset of $\mathbb{Z}^2$ containing the vectors $(0, 0), (\pm 1, 0), (0, \pm 1)$ (a notion generalized below in Lemma 4.13).

The notion of "reduced" was involved in our induction: commuting with two elements of a triple of group elements implies commuting with the third; in the induction, the third vector was "longer" than the other two. Our hypotheses gave us the above initial set of 5 irreducible elements of $\mathbb{Z}^2$, from which we "generated" all of $\mathbb{Z}^2$ using the notion of closed.

The situation in Lemma 4.5 was very similar. There we implicitly used $S := \{h \in H \mid [u^h, w] = 1\}$, and our hypotheses were designed to imply that S is closed under $\mathcal{M}_{\pm\mathrm{BCRW}}$. We will formalize this somewhat more in Lemma 4.12.

Example 4.11. We now reconsider the set $\mathcal{M}_{\mathrm{BCRW}}$ in the preceding example, but this time using the Euclidean length $|v|^2 = x^2 + y^2$, $v = (x, y)$. We wish to determine when v is irreducible. Note that

if $y < 0 < x$ then (x, y) reduces via $(x - 1, y), (x, y), (x - 1, y + 1)$,
if $x < 0 < y$ then (x, y) reduces via $(x, y - 1), (x + 1, y - 1), (x, y)$,
if $x, y < 0$ then (x, y) reduces via $(x, y), (x + 1, y), (x, y + 1)$,
if $0 \le y < x - 1$ then (x, y) reduces via $(x - 1, y), (x, y), (x - 1, y + 1)$, and
if $0 \le x < y - 1$ then (x, y) reduces via $(x, y - 1), (x + 1, y - 1), (x, y)$.

(Compare the proof of Lemma 4.1.) The only vectors (x, y) not covered by these observations are $(0, y)$ for $y \le 0$, $(x, 0)$ for $x \le 0$, and (x, y) for $x, y \ge 0$ and $|x - y| \le 1$, all of which are irreducible.

For Lemma 4.1 we needed to have a vector (x, y) such that both (x, y) and $(-x, -y)$ have the above form. It is clear that for this to happen one of x, y has to be 0 and the other one $0, 1$ or -1. This proves that *the only irreducible vectors for* $\mathcal{M}_{\mathrm{BCRW}}$ *are the* 5 *points inside a ball of radius* 1 *around* $(0, 0)$.

The next results deal with the ways these notions will be used later. First we formalize the use in a group-theoretic setting already hinted at in Example 4.10:

Lemma 4.12. *Let $\mathcal{M}_{\mathcal{U}}$ and $\mathcal{M}_{\mathcal{W}}$ be finite collections of finite sets of elements of a subgroup $H \cong \mathbb{Z}^k$ of a group G. Let $U_0, W_0 \le G$ and $u, w \in G$ satisfy the following conditions*:

(1) *If $M \in \mathcal{M}_{\mathcal{U}}$ and $h \in M$ then u^h can be expressed as a word in U_0 and $\{u^l \mid l \in M \setminus \{h\}\}$;*
(2) *If $M \in \mathcal{M}_{\mathcal{W}}$ and $h \in M$ then w^h can be expressed as a word in W_0 and $\{w^l \mid l \in M \setminus \{h\}\}$; and*
(3) $[U_0^H, w] = [u, W_0^H] = 1$.

Then $S := \{h \in H \mid [u^h, w] = 1\}$ is closed under $\mathcal{M}_{\mathcal{U}}$ and $\mathcal{M}_{-\mathcal{W}}$.

Condition (1) essentially says that, for each $M \in \mathcal{M}_{\mathcal{U}}$, there is a relation among the elements $\{u^h \mid h \in M\}$ that can be solved for each of these elements in terms of the others. Of course, we will want S to be nonempty, which will always be forced by having $[u, w] = 1$.

Proof. We will write H additively. First note that S is closed under $\mathcal{M}_{\mathcal{U}}$. For, if $h \in M \in \mathcal{M}_{\mathcal{U}}$ and $v \in H$, by (2) we can express u^{v+h} as a word in U_0^v and $\{u^{v+l} \mid l \in M \setminus \{h\}\}$. Therefore, by (3), if w commutes with u^{v+l} for each $l \in M \setminus \{h\}$ then it commutes with u^{v+h}.

The preceding argument shows that $\bar{S} := \{h \in H \mid [u, w^h] = 1\}$ is closed under $\mathcal{M}_{\mathcal{W}}$, so that $S = -\bar{S}$ is also closed under $\mathcal{M}_{-\mathcal{W}}$ (see Definition 4.9(4)). □

Next we observe how the notions "closed" and "finite type" (Definition 4.9(1),(3)) will be used (compare the proofs of Lemmas 4.1 and 4.5):

Lemma 4.13. *Fix a length function as in* Definition 4.9(2)*. If S is a closed subset of $\mathbb{Z}^k$ containing the set S_0 of all irreducible vectors for $\mathcal{M}$, then $S = \mathbb{Z}^k$.*

Proof. List all vectors in $\mathbb{Z}^k$ by non-decreasing length: $v_1, v_2, \dots$ (recall the finiteness assumption in Definition 4.2(2)). Using induction on k we will show that $v_k \in S$. If v_k is irreducible it is in S_0 and therefore in S. If v_k is reducible then there is a translate $v + M$, $v \in \mathbb{Z}^k$, $M \in \mathcal{M}$, in which v_k is the longest vector. By induction the other vectors in $v + M$ are in S. Since S is closed, also $v_k \in S$. □

Corollary 4.14. *If the set $\mathcal{M}$ is of finite type then there is a finite set $S_0 \subseteq \mathbb{Z}^k$ such that, if a subset $S \subseteq \mathbb{Z}^k$ is $\mathcal{M}$-closed and contains S_0, then $S = \mathbb{Z}^k$.*

Although the length function does not appear in the statement of the corollary it was needed in the proof. Moreover:

Remark. If we only consider lengths arising from positive definite bilinear forms on $\mathbb{Z}^k$, then the property that $\mathcal{M}$ is of finite type does not depend on the specific choice of length. Of course, the precise number of irreducible vectors does depend on the choice of length function.

4.3. **Central extensions of Borel subgroups.** In this section we apply the results in the preceding sections to go one step closer to simple groups of rank 1: in Propositions 4.19, 4.25 and 4.30 we will provide short bounded presentations for *infinite* central extensions of their Borel subgroups. (Recall from Remark 1.2 that Borel subgroups do not have such presentations.)

4.3.0. *Polynomial notation.* All of our presentations for groups of Lie type involve dealing with polynomials over finite fields.

1. If γ lies in an extension field of $\mathbb{F}_p$, then $m_\gamma(x)$ will 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 $\deg f_{\delta;\gamma} < \deg m_\gamma$.

2. For any polynomial $g(x) = \sum_0^e g_i x^i \in \mathbb{Z}[x]$ and any two elements u, h in a group G, define "powers" as follows, multiplying in the stated order:

$$[[u^{g(x)}]]_h = (u^{g_0})(u^{g_1})^{h^1} \cdots (u^{g_e})^{h^e}, \tag{4.15}$$

so that

$$[[u^{g(x)}]]_h = u^{g_0} h^{-1} u^{g_1} h^{-1} u^{g_2} \cdots h^{-1} u^{g_e} h^e. \tag{4.16}$$

If $0 \le g_i < p$ then this is a word of length $O(pe)$ in u, h; (3.4) explains the collapse of the exponents here (cf. [BKL, p. 512]).

Moreover, if $0 \le g_i < p$ and some $h_2 \in G$ satisfies $u^{h_2} = u^4$ or $u^{h_2} = u^2$, then we can apply (a version of) (3.3) for each u^{g_i}, and hence we can write $[[u^{g(x)}]]_h$ as

a word of length $O(e \log p)$ in u, h_2 and h. Note that $[[u^{f}(x)]]_h$ then depends only on the image of $f(x)$ in $\mathbb{F}_p[x]$.

Care is needed with (4.15): we cannot use it to calculate $[[u^{g(x)+k(x)}]]_h$ or $[[u^{g(x)k(x)}]]_h$ in the obvious manner unless we know that all of the conjugates u^{h^i} commute with one another. If this occurs and $u^p = 1$, then $[[u^{g(x)+k(x)}]]_h = [[u^{g(x)}]]_h[[u^{k(x)}]]_h$, and $[[u^{f(x)g(x)}]]_h = 1$ if $[[u^{f(x)}]]_h = 1$.

We will usually use these formulas when h is such that the conjugates u^{h^i} generate an isomorphic image of the additive group of $\mathbb{F}_q$ and the action of h corresponds to multiplication by some γ on $\mathbb{F}_q = \mathbb{F}_p[\gamma]$. In this case, we have $[[u^{m_\gamma(x)}]]_h = 1$, and an isomorphism is given by

$$\delta \mapsto [[u^{f_{\delta;\gamma}(x)}]]_h \quad \text{for} \quad \delta \in \mathbb{F}_q.$$

(Relations used in Propositions 4.25, 4.30 and 4.33 contain variations on this situation in nonabelian settings.)

4.3.1. *Special linear groups.* Let ζ be a generator of $\mathbb{F}_q^*$, where $q = p^e$ for a prime p. Our first result (Proposition 4.19) requires additional data that exists only when $q \neq 2, 3, 5, 9$: we need $a, b \in \mathbb{F}_q^*$ such that

$$b^2 = a^2 + 1 \quad \text{and} \quad \mathbb{F}_q = \mathbb{F}_p[a^2]. \tag{4.17}$$

Such elements exist for the stated q: if $\zeta^2 + 1$ and $\zeta^2 - 1$ are nonsquares then $\zeta^4 - 1$ is a nonzero square a^2; and this generates $\mathbb{F}_q$ for the stated q since ζ^4 is in no proper subfield. We emphasize that we will be working with a^2 rather than ζ; the latter will be needed only to obtain an element of the desired order $q-1$. See Lemma 4.20 for a variation on the present theme that does not require the equation $b^2 = a^2 + 1$.

We base our presentation on the following matrices:

$$u = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \; h_\star = \begin{pmatrix} \star^{-1} & 0 \\ 0 & \star \end{pmatrix} \quad \text{for } \star \in \{2, a, b, \zeta\}. \tag{4.18}$$

(If $p = 2$ discard h_2.) Then $B := \langle u, h_\zeta \rangle$ *is a Borel subgroup of* $\mathrm{SL}(2, q)$, of order $q(q-1)$.

With this preparation, using the polynomial notation in Section 4.3.0 we can consider the following presentation.

Generators: $u, h_2, h_a, h_b, h_\zeta$. (If $p = 2$ discard h_2.)
Relations:

(1) $[h_\star, h_\bullet] = 1$ for $\star, \bullet \in \{2, a, b, \zeta\}$.
(2) $u^{h_b} = uu^{h_a} = u^{h_a}u$.
(3) $u^{h_2} = u^4$.
(4) $u^p = 1$.
(5) $[[u^{m_{a^2}(x)}]]_{h_a} = 1$.
(6) $u^{h_\zeta} = [[u^{f_{\zeta^2;a^2}(x)}]]_{h_a}$.

Here and in the remainder of this section *we view each of our relations as having length* $O(\log q)$ by using (3.3) and (4.16).

Proposition 4.19. *If* $q \neq 2, 3, 5, 9$, *then this produces a bounded presentation of length* $O(\log q)$ *for an infinite central extension of a Borel subgroup* B *of* $\mathrm{SL}(2, q)$.

Proof. Let $\hat{B}$ be the group defined by this presentation. Using the matrices written above it is easy to check that the natural map $\pi\colon \hat{B} \to B$ sending $u, h_\zeta, h_a, h_b, h_2$

to "themselves" is a homomorphism. In particular, (2) arises from the fact that h_b acts on $\left(\begin{smallmatrix}1 & \star\\ 0 & 1\end{smallmatrix}\right)$ as multiplication of $\star$ by $b^2 = a^2 + 1$.

By (1), (2) and Lemma 4.1 (with h_a and h_b in place of a and b), $U := \langle u^{\langle h_a\rangle}\rangle = \langle u^{h_a^i h_b^j} \mid i, j \in \mathbb{Z}\rangle$ is abelian, and hence is an elementary abelian p-group by (4). In particular, the definition of $[[u^{g(x)}]]_{h_a}$ is independent of the order of the terms for $g(x) \in \mathbb{F}_p[x]$.

Clearly h_a acts as a linear transformation on U. By (5), its minimal polynomial divides the irreducible polynomial $m_{a^2}(x)$. Since the images of u under powers of this linear transformation span U, we have $|U| \le q$, and then we obtain equality using π.

Now $\delta \mapsto [[u^{f_{\delta;a^2}(x)}]]_{h_a}$ is an isomorphism $\mathbb{F}_q^+ \to U$. Since $f_{\delta a^2;a^2}(x) \equiv x f_{\delta;a^2}(x) \pmod{m_{a^2}(x)}$ by the definition of $f_{\delta;a^2}(x)$ in Section 4.3.0, this map extends to an $\mathbb{F}_p[x]$-module isomorphism with x acting on $\mathbb{F}_q$ as multiplication by a^2 and on U as conjugation by h_a. (Note that there are choices for this isomorphism: in place of a^2 we could use any conjugate of it by a field automorphism.)

By (6), the linear transformation induced by h_ζ on U has order $(q-1)/(2, q-1)$. Thus, h_ζ acts on U exactly as occurs in B. All $h_\star$ do as well. It follows from (1) that $\hat{B}$ is a central extension of B, as required.

In view of (4.15), this presentation involves integer exponents that can be much larger than $\log p$. However, using (3) these can be shortened as in (3.3). Similarly, exponents that are polynomials are dealt with in (4.16) using short words. Thus, this presentation produces one that is short and bounded.

Finally, there is a semidirect product $U \rtimes \mathbb{Z}$ that satisfies the above relations: we are dealing with an infinite group. □

Note that

$$Z(\hat{B}) = \langle h_\zeta^d, h_2 h_\zeta^{-n(2)}, h_a h_\zeta^{-n(a)}, h_b h_\zeta^{-n(b)}\rangle,$$

where $d = |\zeta^2| = (q-1)/(2, q-1)$ and $\zeta^{n(\star)} = \star$ for $\star \in \{2, a, b\}$. Almost all of $Z(\hat{B})$ will disappear when we deal with simple groups.

As we saw above, we can choose $a = \zeta$ or $b = \zeta$ or $b = \zeta^2$; correspondingly $h_a = h_\zeta$ or $h_b = h_\zeta$ or $h_b = h_\zeta^2$. This allows us to obtain a shorter presentation by removing one of the generators h_a or h_b, and hence also three relations in (1).

The following is a simple variation on this proposition.

Lemma 4.20. *If $q \ne 2$, then the following produces a bounded presentation of length $O(\log q)$ for an infinite central extension of the 1-dimensional affine group* $\mathrm{AGL}(1, q) = \{x \mapsto \gamma x + \delta \mid \gamma \in \mathbb{F}_q^*, \delta \in \mathbb{F}_q\}$:

Generators: $u, h_2', h_\zeta', h_{\zeta+1}'$. (If $p = 2$ discard h_2'.)

Relations:

(1) $[h_\star', h_\bullet'] = 1$ for $\star, \bullet \in \{2, \zeta, \zeta+1\}$.
(2) $u^{h_{\zeta+1}'} = u u^{h_\zeta'} = u^{h_\zeta'} u$.
(3) $u^{h_2'} = u^2$.
(4) $u^p = 1$.
(5) $[[u^{m_\zeta(x)}]]_{h_\zeta'} = 1$ (cf. (4.15)).

Proof. This is similar to the preceding proof, using the same u but based on the new diagonal matrices $h_\star' = \left(\begin{smallmatrix}1 & 0\\ 0 & \star\end{smallmatrix}\right)$ for $\star \in \{2, \zeta, \zeta+1\}$. □

4.3.2. *Unitary groups.* We now turn to nonabelian versions of the same theme: the unitary groups $\mathrm{SU}(3,q)$. This time let $F = \mathbb{F}_{q^2}$, where $q = p^e$ and $F^* = \langle\zeta\rangle$ as usual. We will use the following matrices that preserve the form $x_1\bar{x}_3+x_2\bar{x}_2+x_3\bar{x}_1$:

$$u(\alpha,\beta) = \begin{pmatrix} 1 & \alpha & \beta \\ 0 & 1 & -\bar{\alpha} \\ 0 & 0 & 1 \end{pmatrix},\ h_2 = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix},\ h_\delta = \begin{pmatrix} \bar{\delta}^{-1} & 0 & 0 \\ 0 & \bar{\delta}/\delta & 0 \\ 0 & 0 & \delta \end{pmatrix}, \tag{4.21}$$

for any field elements such that $\beta + \bar{\beta} = -\alpha\bar{\alpha}$ and $\delta \neq 0$. Let U be the set of all matrices $u(\alpha,\beta)$; its center has $\alpha = 0$. Moreover,

$$u(\alpha,\beta)^{h_\zeta} = u(\zeta^{2q-1}\alpha, \zeta^{q+1}\beta), \text{ and } u(\alpha,\beta)^{h_2} = u(2\alpha, 4\beta) \text{ if } p \neq 2. \tag{4.22}$$

Then $B = U \rtimes \langle h_\zeta\rangle$ is a Borel subgroup of $\mathrm{SU}(3,q)$. Its structure is as follows: $|U| = q^3$, $Z(U) = U' = \Phi(U)$ is elementary abelian of order q, and $U/Z(U)$ is elementary abelian of order q^2; the action of h_ζ on U is given above, and it has order $(q^2-1)/d$ in that action, where $d = (2q-1, q^2-1) = (3, q+1)$.

We start with a property of F somewhat analogous to (4.17):

Lemma 4.23. *If $q \neq 3, 5$ and q is odd, then there are elements $a, b \in F$ satisfying*

$$a^{2q-1} + b^{2q-1} = 1 \quad \textit{and} \quad a^{q+1} + b^{q+1} = 1, \tag{4.24}$$

such that $F = \mathbb{F}_p[a^{2q-1}]$ and $\mathbb{F}_q = \mathbb{F}_p[a^{q+1}]$.

If $q = 2^e$ with $e \geq 2$, then there are elements $a, b \in F$ satisfying (4.24) *such that $\mathbb{F}_q = \mathbb{F}_p[a^{q+1}] = \mathbb{F}_p[a^{2q-1}]$.*

We give a proof in Appendix A.

Proposition 4.25. *There is an infinite central extension of the above Borel subgroup B of $\mathrm{SU}(3,q)$ having a bounded presentation of length $O(\log q)$.*

Proof. Until the end of this proof we assume that $q \neq 2, 3, 5$. We use the abbreviations $\gamma' = \gamma^{q+1}$, $\gamma'' = \gamma^{2q-1}$ for $\gamma \in F$. Section 4.3.0 contains the polynomial notation used below.

Fix elements $a, b \in F$ as in the preceding lemma. We base our presentation on arbitrary elements $u = u(\alpha_0, \beta_0) \notin Z(U)$ and $1 \neq w \in Z(U)$.

We will use the following presentation with (at most) 6 generators and 20 relations.

Generators: $u, w, h_2, h_a, h_b, h_\zeta$. (If $p = 2$ discard h_2.)

Relations:

(1) $[h_\star, h_\bullet] = 1$ for $\star, \bullet \in \{2, a, b, \zeta\}$.
(2) $w = w^{h_a}w^{h_b} = w^{h_b}w^{h_a}$.
(3) $w^{h_2} = w^4$.
(4) $w^p = 1$.
(5) $[[w^{m_{a'}(x)}]]_{h_a} = 1$.
(6) $[[w^{f_{\zeta';a'}(x)}]]_{h_a} = w^{h_\zeta}$.
(7) $u = u^{h_a}u^{h_b}w_1 = u^{h_b}u^{h_a}w_2$.
(8) $[u, w] = [u^{h_a}, w] = 1$.
(9) $u^{h_2} = u^2 w_3$.
(10) $u^p = w_4$.
(11) $[[u^{m_{a''}(x)}]]_{h_a} = w_5$.
(12′) $[[u^{f_{\zeta'';a''}(x)}]]_{h_a} = u^{h_\zeta}w_6$ if q is odd.

(12″) $([[u^{f_{\alpha;a''}(x)}]]_{h_a})^{h_\zeta}[[u^{f_{\beta;a''}(x)}]]_{h_a} = u^{h_\zeta^2}w_6$ if q is even and ζ'' satisfies $\zeta''^2 = \alpha\zeta'' + \beta$ for $\alpha, \beta \in \mathbb{F}_q$.

(13) $[u, u^{h_\zeta}] = w_7$ and $[u^{h_a}, u^{h_\zeta}] = w_8$ if q is even.

Here $w_1, \dots, w_8$ are elements of $W := \langle w^{\langle h_a\rangle}\rangle$; they are determined by ζ and our chosen matrices u, w and $h_\star$.

There is a surjection $\pi\colon \hat{B} \to B$ from the group $\hat{B}$ defined by this presentation onto B; this is the significance of the specific elements w_i. Note that (12″) uses the fact that ζ'' is not in $\mathbb{F}_q$, and hence has a quadratic minimal polynomial over $\mathbb{F}_q$.

By (1), (2) and Lemma 4.1, W is abelian (using h_a, h_b in place of the elements a, b in that lemma) and hence elementary abelian by (4).

Using π we see that $|W| \geq q$. Now (2), (5) and (6) let us identify W with the additive group of $\mathbb{F}_q$ in such a way that each $h_\star$ acts as multiplication by a conjugate of $\star' = \star^{q+1}$.

By (1), (7) and (8), we can apply Lemma 4.5 to deduce that $[u, W] = 1$ (using $U_0 = W$, $W_0 = 1$ and h_a, h_b in place of the elements a, b in that lemma).

If $U := \langle u^{\langle h_a\rangle}, W\rangle$, then $[U, W] = 1$ since $W^{\langle h_a, h_b\rangle} = W$ by (2).

In view of (7), another application of Lemma 4.1 shows that U/W is abelian, and by (10) it is elementary abelian.

Using π we see that $|U/W| \geq |\mathbb{F}_p[a^{2q-1}]|$. Now (11) lets us identify U/W with $\mathbb{F}_p[a^{2q-1}]$ in such a way that each $h_\star$ acts as multiplication by $\star'' = \star^{2q-1}$.

Case q odd. Since $F = \mathbb{F}_q[\zeta'']$, (12′) implies that h_ζ acts irreducibly on U/W, and hence $|U/W| \leq q^2$. Thus, $|U| = q^3$. Since $W = \Phi(U)$, the action of h_ζ on U/W forces its action on W. Since $U \trianglelefteq \hat{B}$, by (1) it follows that $\hat{B}$ is a central extension of B, as required.

Case q even. We use Lemma 4.5 in order to bound $|\langle U, U^{h_\zeta}\rangle/W|$: in that lemma we use $G = \langle h_a, h_b, U, U^{h_\zeta}\rangle/W$, also uW and $u^{h_\zeta}W$ in place of of u and w, and finally $U_0 = W_0 = 1$. The only condition that needs to be checked is Lemma 4.5(4), and this is just (13). Thus, $[u^{\langle h_a, h_b\rangle}, u^{h_\zeta\langle h_a, h_b\rangle}] \leq W$. It follows from (1) that $[U, u^{h_\zeta}] \leq W$ since $U = \langle u^{\langle h_a\rangle}, W\rangle$.

Then $|\langle U, U^{h_\zeta}\rangle| \leq q^3$ by (1), and hence equality holds using π. By (12″), $\langle U, U^{h_\zeta}\rangle^{h_\zeta} = \langle U, U^{h_\zeta}\rangle$. Now we can proceed as in the preceding case, using $\langle U, U^{h_\zeta}\rangle$ in place of U.

Once again $\hat{B}$ is infinite. There are at most 20 relations: 20 if q is odd, and only 17 if q is even since we delete 3 from (1) together with (3) and (9).

We still need to show that the length is $O(\log q)$. By (3), (3.3) and (4.16), each element of W has length $O(\log q)$ in our generators. Similarly, by (9), (10), (3.3) and (4.16), every element in $\langle U, U^{h_\zeta}\rangle/W$ has length $O(\log q)$, and hence the relations (11), (12′) and (12″) can be viewed as short.

This completes the proof of the proposition, but for future reference we include presentations for the omitted cases $q = 2, 3, 5$. When $q = 2$, $B \cong Q_8 \times C_3$. For the remaining q we use the irreducibility of h_ζ on $Q/Z(Q)$.

Generators: u, w, h_ζ.

Relations:

(1) $w^p = 1$.
(2) $[u, u^{h_\zeta}] = w$.
(3) $u^p = 1$.
(4) $[u, w^{h_\zeta^i}] = 1$, $i = 0, 1$.

(5) $[[w^{m_{\zeta'}(x)}]]_{h_\zeta} = 1$.
(6) $[[u^{m_{\zeta''}(x)}]]_{h_\zeta} = w_1$.
Here, $w_1 \in W := \langle w^{\langle h_\zeta \rangle} \rangle$.

In the group with this presentation, W is cyclic of order $q = p$, and by (5) it is normalized by h_ζ. By (4), W commutes with $U := \langle u^{\langle h_\zeta \rangle} \rangle$. By (2), (3) and (6), $\langle U, W \rangle / W$ is elementary abelian, and we can proceed as before. □

4.3.3. *Suzuki groups.* We start with analogues of (4.21) and (4.22), based on [Suz, p. 133]. This time, $F = \mathbb{F}_q$, $q = 2^{2k+1}$ and $F^* = \langle \zeta \rangle$. Let $\theta \colon x \mapsto x^{2^{k+1}}$ be the field automorphism such that $\theta^2 = 2$. Then $B = U \rtimes \langle h_\zeta \rangle$, where U consists of all $(\alpha, \beta) \in F^2$ with multiplication rule

$$(\alpha, \beta)(\gamma, \delta) = (\alpha + \gamma, \beta + \delta + \alpha\gamma^\theta), \tag{4.26}$$

and $W := Z(U) = U'$ is all $(0, \beta)$. Moreover, if $\epsilon \in \mathbb{F}^*$ let h_ϵ denote the automorphism of U defined by

$$(\alpha, \beta)^{h_\epsilon} = (\epsilon\alpha, \epsilon^{\theta+1}\beta). \tag{4.27}$$

Then $\langle h_\zeta \rangle$ is transitive on the nontrivial elements of both $Z(U)$ and $U/Z(U)$.

If $x \in U$ and $v = (i, j, k, l) \in \mathbb{Z}^4$ we write

$$x^v = x^{(i,j,k,l)} = x^h \text{ where } h = h_\zeta^i h_{\zeta^\theta}^j h_{\zeta+1}^k h_{\zeta^\theta+1}^l. \tag{4.28}$$

By (4.27), the following trivial identities in the field F

$$(\zeta) + 1 = (\zeta + 1) \qquad (\zeta^\theta) + 1 = (\zeta^\theta + 1)$$

translate into relations modulo W:

$$u^{(0,0,0,0)} u^{(1,0,0,0)} \equiv u^{(0,0,1,0)} \qquad u^{(0,0,0,0)} u^{(0,1,0,0)} \equiv u^{(0,0,0,1)}$$

for any $u \in U$. These relations correspond to the following collection $\mathcal{M}_\mathcal{U}$ of subsets of $\mathbb{Z}^4$:

$$\big\{\{(0,0,0,0), (1,0,0,0), (0,0,1,0)\}, \{(0,0,0,0), (0,1,0,0), (0,0,0,1)\}\big\}.$$

Since $\theta^2 = 2$ we have $\lambda^{(\theta-1)(\theta+1)} = \lambda$ and $\lambda^{(2-\theta)(\theta+1)} = \lambda^\theta$ for any $\lambda \in F$. Then

$$(\zeta^{-1})^{\theta+1} (\zeta^\theta)^{\theta+1} + 1 = ((\zeta + 1)^{-1})^{\theta+1} (\zeta^\theta + 1)^{\theta+1}$$

$$(\zeta^2)^{\theta+1} ((\zeta^\theta)^{-1})^{\theta+1} + 1 = ((\zeta + 1)^2)^{\theta+1} ((\zeta^\theta + 1)^{-1})^{\theta+1},$$

which, by (4.27), imply that

$$w^{(0,0,0,0)} w^{(-1,1,0,0)} = w^{(0,0,-1,1)} \text{ and } w^{(0,0,0,0)} w^{(2,-1,0,0)} = w^{(0,0,2,-1)}$$

for any $w \in W$. These correspond to another collection $\mathcal{M}_\mathcal{W}$ of subsets of $\mathbb{Z}^4$:

$$\big\{\{(0,0,0,0), (-1,1,0,0), (0,0,-1,1)\}, \{(0,0,0,0), (2,-1,0,0), (0,0,2,-1)\}\big\}.$$

Our presentation of a Borel subgroup of $\mathrm{Sz}(q)$ uses Lemma 4.12 to show that U is nilpotent of class 2. This relies heavily on the observation that the collections $\mathcal{M}_\mathcal{U} \cup \mathcal{M}_{-\mathcal{U}}$, $\mathcal{M}_\mathcal{W} \cup \mathcal{M}_{-\mathcal{W}}$ and $\mathcal{M}_\mathcal{U} \cup \mathcal{M}_{-\mathcal{W}}$ are of finite type. This is almost trivial for the first two collections because they contain copies of $\mathcal{M}_{\pm\mathrm{BCRW}}$ up to a linear transformation. The next lemma, whose proof is in Appendix B, gives the same for the last collection.

Lemma 4.29. *Let $\mathcal{M}_{\mathrm{Sz}} := \mathcal{M}_{\mathcal{U}} \cup \mathcal{M}_{-\mathcal{W}}$ consist of the following four triples of vectors of $\mathbb{Z}^4$:*

(a) $= \{(0,0,0,0),(1,0,0,0),(0,0,1,0)\}$ (b) $= \{(0,0,0,0),(0,1,0,0),(0,0,0,1)\}$
(c) $= \{(0,0,0,0),(1,-1,0,0),(0,0,1,-1)\}$ (d) $= \{(0,0,0,0),(-2,1,0,0),(0,0,-2,1)\}$.

Then $\mathcal{M}_{\mathrm{Sz}}$ is of finite type and there are exactly 139 irreducible vectors for $\mathcal{M}_{\mathrm{Sz}}$, with respect to the length function $|(x_1,x_2,x_3,x_4)|^2 = x_1^2 + 2x_2^2 + x_3^2 + 2x_4^4$ for $(x_1,x_2,x_3,x_4) \in \mathbb{Z}^4$.

Moreover, there is a set $S_0 \subset \mathbb{Z}^4$ consisting of 16 vectors such that, if $S \subseteq \mathbb{Z}^4$ contains S_0 and is $\mathcal{M}_{\mathrm{Sz}}$-closed, then $S = \mathbb{Z}^4$.

The proof of the first part of the lemma involves tedious computations requiring only high school algebra and bookkeeping. More effort is needed for the refinement in the second part (see Appendix B).

Proposition 4.30. *There is an infinite central extension of the above Borel subgroup B of $\mathrm{Sz}(q)$ having a bounded presentation of length $O(\log q)$.*

Proof. We base our presentation on an arbitrary choice of $u \in U \setminus Z(U)$ and on $w = u^2 \in Z(U)$. Let S_0 denote the set of vectors appearing in Lemma 4.29. We will use the following 5 generators and $17 + |S_0| = 33$ relations:

Generators: $u, h_\star$ for $\star \in \{\zeta, \zeta^\theta, \zeta+1, \zeta^\theta+1\}$.

Relations:

(1) $[h_\star, h_\bullet] = 1$ for $\star, \bullet \in \{\zeta, \zeta^\theta, \zeta+1, \zeta^\theta+1\}$.
(2) $w^2 = 1$, where $w := u^2$.
(3) $w^{(0,0,-1,1)} = ww^{(-1,1,0,0)}$ in the notation of (4.28).
(4) $w^{(0,0,2,-1)} = ww^{(2,-1,0,0)}$.
(5) $[w, w^{(-1,1,0,0)}] = 1$.
(6) $[[w^{m_\zeta(x)}]]_a = 1$ where $a := h_\zeta^{-1} h_{\zeta^\theta}$.
(7) $w^{(1,0,0,0)} = [[w^{x^{2^{k+1}+1}}]]_a$, where x^{2^k} is reduced mod $m_\zeta(x)$ in order to obtain a short relation.
(8) $u^{(1,0,0,0)} = uu^{(0,0,1,0)} w_1$.
(9) $u^{(0,1,0,0)} = uu^{(0,0,0,1)} w_2$.
(10) $[u^{(i,j,k,l)}, w] = 1$ for every $(i,j,k,l) \in S_0$.
(11) $[u, u^{(1,0,0,0)}] = w_3$.
(12) $[[u^{m_\zeta(x)}]]_{h_\zeta} = w_4$.
(13) $u^{(0,1,0,0)} = [[u^{x^{2^{k+1}}}]]_{h_\zeta} w_5$.

Here, $w_1, \ldots$ are suitable elements of $W := \langle w^{\langle a \rangle} \rangle$ determined by our choice of u and ζ.

First of all, these relations are satisfied in the Borel subgroup B: if $w = (0,\beta)$ then $(0,\beta)^{h_{\zeta^\theta} h_\zeta^{-1}} = (0, \zeta\beta)$ by (4.27), which yields (6) and (7); (4.27) also yields (12) and (13). We have already seen (3), (4), (8) and (9), while (1), (2), (5), (10) and (11) are clear.

By (1), (3), (5) and Lemma 4.1 (with a as above and $b := h_{\zeta+1}^{-1} h_{\zeta^\theta+1}$), W is abelian. By (2), (3) and (6), W can be identified with F in such a way that a and b act as multiplication by ζ and $\zeta+1$, respectively. By (7), h_ζ acts as multiplication by $\zeta^{\theta+1}$, which together with (3) and (4) implies that $h_\star$ acts on W as $\star^{\theta+1}$ for $\star \in \{\zeta, \zeta^\theta, \zeta+1, \zeta^\theta+1\}$.

Once again we use the notation (4.28). We proceed essentially as in Section 4.1. Let

$$S := \{v \in \mathbb{Z}^4 \mid [u^v, w] = 1\}.$$

Let $\mathcal{M}_{\mathrm{Sz}}$ be as in Lemma 4.29.

By Lemma 4.12 (with $U_0 = W$ and $W_0 = 1$), combined with (3), (4), (7), (8) and the fact that $[w, W] = 1$, the set S is $\mathcal{M}_{\mathrm{Sz}}$-closed (Definition 4.9(1)). Since S contains S_0 by (10), we have $S = \mathbb{Z}^4$ by Lemma 4.29, so that $[u^{(i,j,k,l)}, w] = 1$ for all $(i, j, k, l) \in \mathbb{Z}^4$. Thus, each $[u, w^{(i,j,k,l)}] = 1$ and hence $[u, W] = 1$.

It follows that $W = W^{h_\zeta}$ commutes with $U := \langle u^{\langle h_\zeta \rangle} \rangle$. By (1), (8), (9) and (13), U is invariant under each $h_\star$, and hence contains $\langle (u^2)^{\langle a \rangle} \rangle = \langle w^{\langle a \rangle} \rangle = W$. Thus, $W \leq Z(U)$.

By (1), (8) and (11), we can apply Lemma 4.1 to U/W. Then this group is elementary abelian since $w = u^2$. By (12), $|U| \leq q^2$. As in Proposition 4.25, it follows that the presented group $\hat{B}$ is a central extension of B. Once again $\hat{B}$ is infinite. □

4.3.4. *Ree groups.* We cannot prove Theorem A for the Ree groups ${}^2G_2(3^{2k+1})$, but we can at least handle their Borel subgroups. Let $F = \mathbb{F}_{3^{2k+1}}$, $\theta\colon x \mapsto x^{3^{k+1}}$ and $F^* = \langle \zeta \rangle$. By [KLM], $B = U \rtimes \langle h_\zeta \rangle$ where $U = \{x(a, b, c) \mid a, b, c \in F\}$ and

$$x(a_1, b_1, c_1)x(a_2, b_2, c_2) = x(a_1 + a_2, b_1 + b_2 + a_1 a_2^\theta, \\ c_1 + c_2 - a_1^2 a_2^\theta + a_2 b_1 + a_1 a_2^{\theta+1}) \tag{4.31}$$

$$x(a, b, c)^{h_\zeta} = x(\zeta^{2-\theta} a, \zeta^{\theta-1} b, \zeta c). \tag{4.32}$$

Proposition 4.33. *There is an infinite central extension of a Borel subgroup B of ${}^2G_2(q)$ having a bounded presentation of length $O(\log q)$.*

A sketch of a proof is given in Appendix C. These groups are harder to work with than previous ones. In particular, whereas we used $\mathbb{Z}^2$ for $\mathrm{SL}(2, q)$, and $\mathbb{Z}^4$ for Suzuki groups, for these Ree groups we use $\mathbb{Z}^6$.

4.4. **Presentations for rank 1 groups.** We are now ready to consider the rank 1 groups $\mathrm{SL}(2, q)$, $\mathrm{PSL}(2, q)$, $\mathrm{PGL}(2, q)$, $\mathrm{PSU}(3, q)$, $\mathrm{SU}(3, q)$ and $\mathrm{Sz}(q)$. Except for headaches due to field extensions or nonabelian p-groups, the idea in this section is the same as the one in Section 3.2

In each case there is 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^{-q} in the unitary case). These lead us to the *Steinberg presentation* for these groups [St2, Sec. 4]:

Theorem 4.34. (Steinberg presentation) *Each of the above groups has a presentation using*

1. *any presentation for B,*
2. *any presentation for $\langle h, t \rangle$, and*
3. *$|U| - 1$ relations of the form*

$$u_0^t = u_1 h_0 t u_2, \tag{4.35}$$

with u_0, u_1, u_2 nontrivial elements of U and $h_0 \in \langle h \rangle$ (one such relation for each choice of u_0).

We have already used a special case of this presentation in Section 3.2. We now use this presentation in order to prove the following

Theorem 4.36. (a) *All of the groups* $\mathrm{SL}(2,q), \mathrm{PSL}(2,q), \mathrm{PGL}(2,q), \mathrm{PSU}(3,q)$, $\mathrm{SU}(3,q)$ *and* $\mathrm{Sz}(q)$ *have bounded presentations of length* $O(\log q)$.

(b) *Each element of each of the preceding groups can be written as a word of length* $O(\log q)$ *in the generating set used in the above presentation.*

Proof. We will treat the groups separately. In each case we will show that we only need some of the relations (4.35) in order to deduce all of those relations.

4.4.1. *Special linear groups.* Here we will be more explicit than in the unitary or Suzuki cases. We may assume that $q > 9$. We will use the notation leading up to Proposition 4.19. We will show that $\mathrm{SL}(2,q)$ has the following presentation using (at most) 6 generators and 22 relations.

Generators: $u, t, h_2, h_a, h_b, h_\zeta$. (If $p = 2$ discard h_2.)

Relations:

(1) $[h_\star, h_\bullet] = 1$ for $\star, \bullet \in \{2, a, b, \zeta\}$.
(2) $u^{h_b} = uu^{h_a} = u^{h_a}u$.
(3) $u^{h_2} = u^4$.
(4) $u^p = 1$.
(5) $[[u^{m_{a^2}(x)}]]_{h_a} = 1$.
(6) $u^{h_\zeta} = [[u^{f_{\zeta^2;a^2}(x)}]]_{h_a}$.
(7) $h_\star^t = h_\star^{-1}$ for $\star \in \{2, a, b, \zeta\}$.
(8) $[t^2, u] = 1$.
(9) $t = uu^t u$.
(10) $h_\star t = [[u^{f_{\bullet;a^2}(x)}]]_{h_a} \cdot [[u^{f_{\star;a^2}(x)}]]_{h_a}^t \cdot [[u^{f_{\bullet;a^2}(x)}]]_{h_a}$, where $\bullet = \star^{-1}$ for $\star \in \{2, a, b, \zeta\}$.

This certainly involves more relations than one might expect ([CRW1] uses at most 13 relations). In [GKKL2] we give a similar presentation (of bit-length $O(\log q)$ but *not* of length $O(\log q)$) using only 9 relations.

In order to show that the group G presented in this manner is isomorphic to $\mathrm{SL}(2,q)$, note that $\mathrm{SL}(2,q)$ is a homomorphic image of G such that t is in (3.2) and $\hat{B} := \langle u, h_2, h_a, h_b, h_\zeta \rangle$ (which we identify with a subgroup of G) maps onto a Borel subgroup B of $\mathrm{SL}(2,q)$. By (1)-(6), we can apply Proposition 4.19: $\hat{B}$ is isomorphic to a central extension of B.

In particular, $U := \langle u^{\langle h_a \rangle} \rangle$ is a normal subgroup of $\hat{B}$, and h_a acts irreducibly on U. Then $U = [U, h_a] \le \hat{B}' \le G'$, so that $\langle U, U^t \rangle \le G'$. By (9), $t \in \langle U, U^t \rangle$, and then all $h_\star \in \langle U, U^t \rangle$ by (10), so that $G = \langle U, U^t \rangle$ and G is a perfect group.

In the identification of U with $\mathbb{F}_q$ in Proposition 4.19, h_ζ acts as multiplication by ζ^2. By (7), since $h_\zeta^{(q-1)/(2,q-1)}$ centralizes U it also centralizes U^t and hence G. Similarly, by (8) we have $Z := \langle h_\zeta^{(q-1)/(2,q-1)}, t^2 \rangle \le Z(G)$.

We claim that G/Z satisfies the presentation for $\mathrm{PSL}(2,q)$ in Theorem 4.34. For, $\langle h_\zeta, t \rangle / Z$ is dihedral of order $2(q-1)/(2,q-1)$ by (7). Also, $U\langle h_\zeta, t^2 \rangle / Z$ is isomorphic to a Borel subgroup of $\mathrm{PSL}(2,q)$. Moreover, $\langle h_\zeta \rangle$ acts on the non-trivial elements of U with at most two orbits; orbit-representatives are u^1 and $[[u^{f_{\zeta;a^2}(x)}]]_{h_a}$. By (9) and (10), relation (4.35) holds for $u_0 = u^1$ or $[[u^{f_{\zeta;a^2}(x)}]]_{h_a}$;

conjugating these relations by $\langle h_\zeta \rangle$ produces all of the required relations (4.35). This proves our claim.

The only perfect central extensions of $\mathrm{PSL}(2,q)$ are $\mathrm{PSL}(2,q)$ and $\mathrm{SL}(2,q)$ (since $q > 9$). Hence, $G \cong \mathrm{SL}(2,q)$.

This presentation is clearly bounded. It can be made short: relations (5), (6) and (10) have length $O(\log q)$ by (4.16) and (3.3). (Note that the generators h_a and h_b are not needed here.)

We still need to verify Theorem 4.36(b). Every element of $\mathrm{SL}(2,q)$ is in UU^tUU^tU. Thus, we only need to verify (b) for U, and again this follows from (4.16) and (3.3).

Case $\mathrm{PSL}(2,q)$. Replace the previous relation (8) by $t^2 = 1$.

Case $\mathrm{PGL}(2,q)$. This is proved as in the preceding case, using Lemma 4.20 in place of Proposition 4.19.

4.4.2. *Unitary groups.* We may assume that $q > 5$. We use the matrices in (4.21), together with the additional matrix

$$t = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{4.37}$$

In the notation of (4.21), if $\beta \neq 0$ and $\beta + \bar{\beta} + \alpha\bar{\alpha} = 0$ then (4.35) becomes

$$u(\alpha,\beta)^t = u(-\alpha/\bar{\beta}, 1/\beta)\, h_\beta\, t\, u(-\alpha/\beta, 1/\beta). \tag{4.38}$$

We will make fundamental use of a result of Hulpke and Seress [HS], who showed that the presentation in Theorem 4.34 can be shortened by using (1), (2) and a carefully chosen set of at most 7 of the relations (3).

The following presentation uses (at most) 14 generators and 52 relations. We will show that it defines a group G that is isomorphic to $\mathrm{SU}(3,q)$.

Generators: $u, w, t, h_2, h_a, h_b, h_\zeta, \breve{h}_i$ for $1 \le i \le 7$. (If $p = 2$ discard h_2.)

Relations:

(0) All relations (1)-(13) used in Proposition 4.25.
(14) $t^2 = 1$.
(15) $h_\star = v_{\star 1} v_{\star 2}^t v_{\star 3} t$ for $\star \in \{1, 2, a, b, \zeta\}$, where $h_1 := 1$.
(16) $u^{h_\star^t} = [[u^{f_\star(x)}]]_{h_\zeta}$ for $\star \in \{2, a, b, \zeta\}$.
(17) $u_i^t = u_{i1} \breve{h}_i t u_{i2}$ for $1 \le i \le 7$, relations due to Hulpke and Seress [HS].
(18) $[h_\zeta, h_\zeta^t] = 1$.
(19) $[h_\zeta, \breve{h}_i] = 1$ for $1 \le i \le 7$.
(20) $u^{\breve{h}_i} = [[u^{g_i(x)}]]_{h_\zeta}$ and $u^{\breve{h}_i^t} = [[u^{g_i^\bullet(x)}]]_{h_\zeta}$ for $1 \le i \le 7$.

Here $u_i, u_{i1}, u_{i2}, v_{\star 1}, v_{\star 2}, v_{\star 3}$ are specific elements of $U := \langle u^{\langle h_\zeta \rangle} \rangle$, and $f_\star(x)$, $g_i(x)$, $g_i^\bullet(x)$ are specific polynomials in $F[x]$.

There is a surjection $\pi\colon G \to \mathrm{SU}(3,q)$ taking each of the generators to "itself". This is clear for (0) and (14), while (15) comes from (4.38). Relations (16) and (18)-(20) encode the fact that, in $\mathrm{SU}(3,q)$, $h_\star$ and $\breve{h}_i$ are powers of h_ζ, while $h_\star^t = h_\star^{-q}$ and $\breve{h}_i^t = \breve{h}_i^{-q}$. There are at most 7 relations of the form (17) in the presentation for $\mathrm{SU}(3,q)$ given in [HS]. (All of these relations are given far more explicitly in [BGKLP, (11) and (12) on p. 103] and [HS], using (4.38).)

Moreover, π maps $\hat{B} := \langle u, w, h_2, h_a, h_b, h_\zeta \rangle$ (which we identify with a subgroup of G) onto a Borel subgroup B of $\mathrm{SU}(3,q)$. Proposition 4.25 implies that $\hat{B}$ is a central extension of B. In particular, since $q > 5$, U is a normal subgroup of $\hat{B}$ of order q^3 and h_ζ acts irreducibly on its Frattini quotient $U/Z(U)$. It follows that $U = [U, h_\zeta] \le G'$. By (15) for h_1, we have $t \in \langle U, U^t \rangle$, and then all $h_\star, \breve{h}_i \in \langle U, U^t \rangle$ by (15) and (17), so that $G = \langle U, U^t \rangle$ and G is a perfect group.

In particular, in the identification of $U/Z(U)$ with $\mathbb{F}_{q^2}$ in Proposition 4.25, h_ζ acts as multiplication by ζ^{2q-1} (cf. (4.22)), and hence has order $d := (q^2-1)/(3, q+1)$ in its action on U.

By (16) and (18), h_ζ^t also acts on U, and is irreducible on $U/Z(U)$. Thus, $U = \langle u^{\langle h_\zeta^t \rangle} \rangle$.

Let $Z = Z(G)$. We claim that $h \in \langle h_\zeta \rangle Z$ if h is any of the elements $h_\star, \breve{h}_i$. By Proposition 4.25, (19) and (20), h normalizes U and there is an integer k such that $1 \le k < d$ and hh_ζ^k centralizes U. By (16), (19) and (20), since h^t commutes with h_ζ^t it also normalizes $U = \langle u^{\langle h_\zeta^t \rangle} \rangle$, and there is also an integer k' such that $1 \le k' < d$ and $(hh_\zeta^{k'})^t$ centralizes U. Since $\pi(G) = \mathrm{SU}(3,q)$, we have $k = k'$, so that hh_ζ^k centralizes $\langle U, U^t \rangle = G$, which proves our claim.

We can now use Theorem 4.34 to show that $G/Z \cong \mathrm{PSU}(3,q)$. For, (16) implies that $\langle h_\zeta, t, Z \rangle / Z$ has order $2(q^2-1)/(3, q+1)$, and hence is isomorphic to the required subgroup of $\mathrm{PSU}(3,q)$. Also, $\hat{B}Z/Z = U\langle \text{all } h_\star \rangle Z/Z = U\langle h_\zeta \rangle Z/Z$ is isomorphic to a Borel subgroup of $\mathrm{PSU}(3,q)$. The only relations remaining to be checked for G/Z are those in (4.35), and these are (17) (read modulo Z).

Since $\mathrm{SU}(3,q)$ is an epimorphic image of G, it follows that $G \cong \mathrm{SU}(3,q)$ (see [GLS, Corollary 6.1.7]). This presentation is clearly bounded. It can be made short using (4.16) and (3.3).

Finally, Theorem 4.36(b) holds when $g \in U$ by (4.16) and (3.3), then also when $g \in U^t$, and hence whenever $g \in UU^tUU^tU = \mathrm{SU}(3,q)$. $\square$

By Theorem 4.36(b), we also obtain a short bounded presentation for $\mathrm{PSU}(3,q)$ since a generator of $Z(\mathrm{SU}(3,q))$ has length $O(\log q)$.

4.4.3. *Suzuki groups.* We may assume that $q > 8$. This time we will use a remarkable result of Suzuki [Suz, p. 128], who showed that the presentation in Theorem 4.34 can be shorted by using (1), (2) and a single carefully chosen relation (3) (in fact, this is just the relation needed to define $\mathrm{Sz}(2)$).

We use the following presentation having 7 generators and 43 relations.

Generators: $u, w, t, h_\zeta, h_{\zeta+1}, h_{\zeta^\theta}, h_{\zeta^\theta+1}$.

Relations:

(0) All 33 relations (1)-(13) in Proposition 4.30.
(14) $t^2 = 1$.
(15) $h_\star^t = h_\star^{-1}$ for $\star \in \{\zeta, \zeta+1, \zeta^\theta, \zeta^\theta+1\}$.
(16) $t = u_1 u_2^t u_3$, a relation due to Suzuki [Suz, (13)].
(17) $h_\star t = u_{\star 1} u_{\star 2}^t u_{\star 3}$ for $\star \in \{\zeta, \zeta+1, \zeta^\theta, \zeta^\theta+1\}$.
Here $u_1, u_2, u_3, u_{\star 1}, u_{\star 2}, u_{\star 3}$ are specific elements of $U := \langle u^{\langle h_\zeta \rangle} \rangle$ that depend on the choices of ζ and u in Section 4.3.3.

Note that, in the Suzuki group case, we do not need additional relations as in Section 4.4.2(19), since the power of h_ζ occurring in relation (16) is 1.

It order to show that the group G presented in this manner is isomorphic to $\mathrm{Sz}(q)$, first note that $\mathrm{Sz}(q)$ is a homomorphic image of G such that $\hat{B} := \langle u, h_\zeta, h_{\zeta+1}, h_{\zeta^\theta}, h_{\zeta^\theta+1}\rangle$ projects onto a Borel subgroup B of $\mathrm{Sz}(q)$. (For more explicit versions of (16) and (17), see [Suz, (13) and (38)].)

The rest of the proof is essentially the same as in Sections 4.4.1 and 4.4.2. Both $U\langle h_\zeta\rangle/\langle h_\zeta^{q-1}\rangle$ and $\langle h_\zeta, t\rangle/\langle h_\zeta^{q-1}\rangle$ are as required in Theorem 4.34. Suzuki [Suz, p. 128] proved that a single relation (16) is the *only* one of the relations (4.35) needed in order to deduce all q^2-1 of them. Hence, $G/\langle h_\zeta^{q-1}\rangle \cong \mathrm{Sz}(q)$. Finally, $G \cong \mathrm{Sz}(q)$ since G has no proper perfect central extensions (recall that $q > 8$; cf. [GLS, pp. 312-313]).

This presentation can be made short by using (4.16) and (3.3). It remains to consider Theorem 4.36(b). Each element of U or U^t is a word of length $O(\log q)$ in our generators. Now (b) follows from the fact that $\mathrm{Sz}(q) = UU^tUU^tU$.

This completes the proof of Theorem 4.36. □

We have not dealt with the Ree groups in Theorem A. The obstacle to both short and bounded presentations of these groups is the fact that all presentations presently known use more than q relations of the form (4.35).

5. Fixed rank

In this section we will prove the following result, which implies Theorem A when the rank is bounded:

Theorem 5.1. *All perfect central extensions of finite simple groups of Lie type and given rank n, except perhaps the Ree groups ${}^2G_2(q)$, have presentations of length $O(\log|G|)$ with $O(n^2)$ relations.*

In view of Theorem 4.36, we may assume that the $n \geq 2$. The idea in the proof is simple: replace the relations in the Curtis-Steinberg-Tits presentation (Sections 5.1 and 5.2) involving rank 1 subgroups with the short bounded presentation in Theorem 4.36.

5.1. **Curtis-Steinberg-Tits presentation.** Steinberg [St1] gave a presentation for the groups of Lie type involving all roots groups corresponding to the root system. Curtis [Cur] and Tits [Ti2, Theorem 13.32] independently obtained similar presentations involving only rank 1 and 2 subsystems of the root system. It is the latter type of presentation that we will use, since less data is required. There are two slightly different versions of the presentation: the "high-level" one, due to Tits and based on buildings, that does not explicitly deal with root groups, and the more involved one that uses root groups and their commutator relations. For now we will state a version of the higher level one, explaining the more familiar one with formulas in the next section.

Each group G of Lie type has a suitably defined root system Φ and fundamental system $\Pi = \{\alpha_1, \ldots, \alpha_n\}$ of roots. We view the twisted groups of type ${}^2A_m(q)$ (with $m \geq 3$ and $m = 2n$ or $m = 2n+1$), ${}^2D_m(q)$ (with $m \geq 4$), ${}^2E_6(q)$ or ${}^3D_4(q)$ as having root systems of type C_n, B_{m-1}, F_4 or G_2, respectively. There is also a suitable sense in which ${}^2F_4(q)$ has 16 roots [GLS, Corollary 2.4.6].

Theorem 5.2. (Curtis-Steinberg-Tits presentation [Cur, St1, Ti2] and [GLS, Theorem 2.9.3]) *Let G be a simple group of Lie type and rank at least 3. Let Π be a*

fundamental system of roots in the root system. Whenever $\alpha \in \Pi$ *let* L_α *be the corresponding rank* 1 *subgroup of* G*; whenever* $\alpha, \beta \in \Pi$, $\beta \neq \pm\alpha$, *let* $L_{\alpha,\beta} := \langle L_\alpha, L_\beta \rangle$ *be the corresponding rank* 2 *subgroup of* G*. Then the amalgamated product of the rank* 2 *subgroups* $L_{\alpha,\beta}$*, amalgamated along the rank* 1 *subgroups* L_α*, is a central extension of* G*.*

Of course, the above version of the Curtis-Steinberg-Tits presentation also gives no information concerning presentations of rank 2 groups. For these we need to use commutator relations. The commutator version of the Curtis-Steinberg-Tits presentation is in [Cur], as well as [GLS, Theorem 2.9.3], and will be described in the next section.

5.2. **Commutator relations.** Bounded rank groups were considered in [BGKLP]. We will use the method of that paper, inserting our presentations for rank 1 groups in place of the standard ones used there. In that paper, all commutator relations are correctly presented, as are the definitions of the toral elements h_β. However, there were significant errors in that paper: for most of the twisted groups, the actions of h_β on the root group elements $x_\alpha(a)$ (or $x_\alpha(a,b)$ for odd-dimensional unitary groups) were not correct. The methodology used was, however, correct, and will be used here without any explicit formulas. The same methodology was used in [KoLu], though not for all possible types of group and without considerations of lengths of presentations.

We assume that G is a universal group of Lie type [GLS, p. 38]. We will factor out part or all of $Z(G)$ later in this section; this point of view occurred in the Section 4.4, and will also be used in Section 6. We will ignore the tiny number of cases in which G has a perfect central extension by a group whose order is the characteristic [GLS, p. 313].

With each root $\alpha \in \Phi$ there is associated a rank 1 group $L_\alpha \cong \mathrm{SL}(2,q)$, $\mathrm{PSL}(2,q)$, $\mathrm{SL}(2,q^2)$, $\mathrm{PSL}(2,q^2)$, $\mathrm{SL}(2,q^3)$, $\mathrm{SU}(3,q)$ or $\mathrm{Sz}(q)$. As indicated above, we use the presentation $\mathbf{P}_\alpha$ for L_α obtained in Section 4.4. Among our generators of L_α we have the following: a torus $\langle h_\alpha \rangle$; a p-element u_α lying in a root group U_α of L_α and of G (note that $U_\alpha = \langle u_\alpha^{\langle h_\alpha \rangle} \rangle$ except when $L_\alpha = \mathrm{SU}(3,2)$); and a "reflection" r_α (called "t" in Section 4.4) that is an involution modulo $Z(L_\alpha)$ such that r_α normalizes $\langle h_\alpha \rangle$ and $U_\alpha^{r_\alpha} = U_{-\alpha}$ is the opposite root group (relative to Φ). When q is tiny the generic presentations in Section 4.4 are not applicable, but there certainly are short, bounded presentations using at least the above generators.

We have $L_{-\alpha} = L_\alpha$. Since G is universal, so is L_α. Checking all cases, we find that

$$|h_\alpha| = |\bar{U}_\alpha^\#|, \text{ where } \bar{U}_\alpha^\# := \bar{U}_\alpha \backslash \{1\} \text{ for } \bar{U}_\alpha := U_\alpha / \Phi(U_\alpha) \tag{5.3}$$

(here $\Phi(U_\alpha)$ denotes the Frattini subgroup of U_α).

There is a Borel subgroup B associated with the positive roots in Φ. It contains U_α for all positive α and has a maximal torus $H := \langle h_\alpha \mid \alpha \in \Pi \rangle$. Moreover, H normalizes each root group U_α, and acts on $\bar{U}_\alpha$ as multiplications by elements in an extension field $\mathbb{F}_{q^{e(\alpha)}}$ of $\mathbb{F}_q$, where $|\bar{U}_\alpha| = q^{e(\alpha)}$, $e(\alpha) \leq 3$, and 3 occurs only for ${}^3D_4(q)$.

By Theorem 5.2 we need to consider all of the subgroups L_{α_i,α_j}, $i \neq j$, arising from our fundamental system $\Pi = \{\alpha_1, \ldots, \alpha_n\}$. Here L_{α_i,α_j} is a rank 2 group whose root system Φ_{α_i,α_j} has the fundamental system $\{\alpha_i, \alpha_j\}$. Crucial for our

argument (and those in [BGKLP, KoLu]) is the fact that H has boundedly many orbits on the direct product of suitable pairs $\bar{U}_\alpha$, $\bar{U}_\beta$. For this we need the following

Lemma 5.4. *Assume that $L := L_{\alpha_i,\alpha_j}$ is not* $\mathrm{SU}(5,2)$. *Let α and $\beta \neq \pm\alpha$ be in Φ_{α_i,α_j}. If $u_\alpha \in U_\alpha \backslash \Phi(U_\alpha)$ and $u_\beta \in U_\beta \backslash \Phi(U_\beta)$, then $C_{\alpha,\beta} := C_{\langle h_\alpha, h_\beta\rangle}(\langle u_\alpha, u_\beta\rangle)$ is in the center of the rank 2 subgroup $X_{\alpha,\beta} := \langle U_\alpha, U_{-\alpha}, U_\beta, U_{-\beta}\rangle$ of L. Moreover, either*

(a) *$C_{\alpha,\beta}$ centralizes L,*
(b) *q is odd and $X_{\alpha,\beta} = \mathrm{SL}(2,q) \circ \mathrm{SL}(2,q)$ or $\mathrm{SL}(2,q) \times \mathrm{SL}(2,q)$ with center $C_{\alpha,\beta}$, inside $\mathrm{Sp}(4,q)$, $\mathrm{SU}(4,q)$ or $G_2(q)$,*
(c) *q is odd and $X_{\alpha,\beta} = \mathrm{SL}(2,q) \circ \mathrm{SL}(2,q^3)$ with center $C_{\alpha,\beta}$, inside ${}^3D_4(q)$,*
(d) *q is odd and $X_{\alpha,\beta} = \mathrm{SU}(4,q)$ with center $C_{\alpha,\beta}$, or*
(e) *$X_{\alpha,\beta} = \mathrm{SL}(3,q)$ with $3|q-1$, inside $L = G_2(q)$ or ${}^3D_4(q)$.*

Proof. As already noted, $U_\alpha = \langle u_\alpha^{\langle h_\alpha\rangle}\rangle$ since L is not $\mathrm{SU}(5,2)$. Since H acts on L_α as a group of automorphisms fixing U_α and $U_{-\alpha}$, we then have $C_H(u_\alpha) = C_H(U_\alpha) = C_H(L_\alpha)$. The latter equality can be seen by inspecting each of the rank 1 groups L_α; but it is more easily understood by noting that $C_H(U_\alpha)$ normalizes each member of $\{U_\alpha\} \cup (U_{-\alpha})^{U_\alpha} = (U_\alpha)^{L_\alpha}$.

Consequently, $C_{\alpha,\beta}$ centralizes the rank 2 subgroup $X_{\alpha,\beta}$; we may assume that $C_{\alpha,\beta}$ does not centralize L. There are very few orbits of unordered distinct pairs $\{\pm\alpha\}, \{\pm\beta\}$ under the action of the Weyl group of L on Φ_{α_i,α_j}. When G does not have type ${}^2F_4(q)$ it straightforward to calculate $X_{\alpha,\beta}$ in each case and then to check that it appears in (b)–(e). The case $G = L = {}^2F_4(q)$ cannot occur in view of the orders of the centralizers of semisimple elements obtained in [Shi]. □

We now describe the version of the Curtis-Steinberg-Tits presentation using commutator relations [Cur] (cf. [GLS, Sec. 2.9]). For all roots $\alpha, \beta \neq \pm\alpha$ lying in the same rank 2 subsystem spanned by a pair of members of Π we use the following two types of relations:

(P_α) A presentation $\mathbf{P}_\alpha$ of L_α, including the elements u_α, h_α and r_α described above. Let U_α be as above.
($B_{\alpha,\beta}$) $[x_\alpha, x_\beta]$ = product of elements of root groups $U_{i\alpha+j\beta}$ with $i\alpha+j\beta \in \Phi$, $i,j > 0$, for each nontrivial $x_\alpha \in U_\alpha, x_\beta \in U_\beta$.

Here and below, the terms of the product are assumed to be exactly the ones that occur in the given group G. Precise formulas for these products are given in [St1, Sec. 10], [Gr] and [GLS, Theorem 2.4.5].

This gives the standard presentation for G. We will soon shorten this presentation with the help of one further type of relations for the same pairs α, β:

($H_{\alpha,\beta}$) $x_\alpha^{h_\beta} = \star_\alpha$ for each nontrivial $x_\alpha \in U_\alpha$, where $\star_\alpha \in U_\alpha$ depends on both x_α and h_β.

The crucial if simple observation is that, since $\langle h_\alpha, h_\beta\rangle$ normalizes U_α and U_β by (P_α), (P_β) and ($H_{\alpha,\beta}$), it acts by conjugation on the set of relations ($B_{\alpha,\beta}$). Hence we will use (some of) the relations ($H_{\alpha,\beta}$) to deduce all of the relations ($B_{\alpha,\beta}$) from a small number of them. Specifying the latter subset is mostly a matter of bookkeeping.

We break the proof of Theorem 5.1 into several cases.

Rank 2 case: Define $z = z(G)$ as follows:

$$\begin{array}{rccccccc} G = & \mathrm{SL}(3,q) & \mathrm{Sp}(4,q) & G_2(q) & \mathrm{SU}(4,q) & \mathrm{SU}(5,q) & {}^3D_4(q) & {}^2F_4(q) \\ z = & (3,q-1) & (2,q-1) & (3,q-1) & (4,q+1) & (20,q+1) & (3,q-1) & 1 \end{array}$$

so that $z \le 20$.

Recall that $H = \langle h_{\alpha_1}, h_{\alpha_2}\rangle$ acts on each $\bar{U}_\alpha$ as a group of field multiplications in an extension field $\mathbb{F}^*_{q^{e(\alpha)}}$. Let $x_{\alpha\mu} \in U_\alpha$, $1 \le \mu \le z$, denote elements corresponding to coset representatives in $\mathbb{F}^*_{q^{e(\alpha)}}$ of the zth powers of all elements of $\mathbb{F}^*_{q^{e(\alpha)}}$ (or projecting onto such coset representatives in $\bar{U}_\alpha$ for nonabelian U_α in the $\mathrm{SU}(5,q)$ and ${}^2F_4(q)$ cases).

We now replace the preceding presentation by the following one for a group J (using the same pairs α, β as above).

- (P_α) As before. This also provides us with the various $x_{\alpha\mu}$, which we may assume are among the generators of L_α.
- $(B'_{\alpha,\beta})$ $[x_{\alpha\mu}, x_{\beta\nu}]$ = product of elements of root groups $U_{i\alpha+j\beta}$ with $i\alpha + j\beta \in \Phi$, $i, j > 0$, for all μ, ν.
- $(H'_{\alpha,\beta})$ $x_{\alpha\mu}^{h_\beta} = \star_\alpha$ for all μ, where $\star_\alpha \in U_\alpha$ depends on both $x_{\alpha\mu}$ and h_β.

We will show that $J \cong G$. Since G is a homomorphic image of J, for each root α we have a subgroup L_α behaving as in (5.3).

Once again, $U_\alpha = \langle x_\alpha^{\langle h_\alpha\rangle}\rangle$ if $x_\alpha \in U_\alpha \backslash \Phi(U_\alpha)$. The action of H on each $\bar{U}_\alpha$ as a group of field multiplications is encoded in (P_α), (P_β) and $(H'_{\alpha,\beta})$.

For some perpendicular pairs α, β the relations $(B'_{\alpha,\beta})$ state that $[x_{\alpha\mu}, x_{\beta\nu}] = 1$. For such a pair α, β we only need one such relation with $x_{\alpha\mu} \in U_\alpha \backslash \Phi(U_\alpha)$ and $x_{\beta\nu} \in U_\beta \backslash \Phi(U_\beta)$, since conjugating by all elements of $\langle h_\alpha, h_\beta\rangle$ then implies that $[U_\alpha, U_\beta] = 1$, and hence that all relations $(B_{\alpha,\beta})$ hold.

We deal with $(B'_{\alpha,\beta})$ for all remaining pairs α, β in a somewhat similar manner. Lemma 5.4 implies that $|C_{\alpha,\beta}|$ *divides* z. By (5.3), each orbit of $\langle h_\alpha, h_\beta\rangle$ on $\bar{U}_\alpha^\# \times \bar{U}_\beta^\#$ has size $|\bar{U}_\alpha^\# \times \bar{U}_\beta^\#|/|C_{\alpha,\beta}|$. Hence, our pairs $(x_{\alpha\mu}, x_{\beta\nu})$ include representatives for all $\langle h_\alpha, h_\beta\rangle$–orbits on $\bar{U}_\alpha^\# \times \bar{U}_\beta^\#$.

Conjugating the relations $(B'_{\alpha,\beta})$ by all elements of H, and using $(H'_{\alpha,\beta})$, we obtain all relations $(B_{\alpha,\beta})$ if U_α and U_β are abelian. When U_α or U_β is nonabelian, for each pair $(x_\alpha\Phi(U_\alpha), x_\beta\Phi(U_\beta))$ of cosets we still obtain a relation of the form $(B_{\alpha,\beta})$. By using the elementary identity $[x, uv] = [x, v][x, u]^v$ in both G and the presented group J, we see that our conjugates of the relations $(B'_{\alpha,\beta})$ imply all relations $(B_{\alpha,\beta})$.

Thus, all relations (P_α) and $(B_{\alpha,\beta})$ required in the Curtis-Steinberg-Tits presentation hold for J, so that $J \cong G$.

There are at most 20^2 relations $(B'_{\alpha,\beta})$ and 20 relations $(H'_{\alpha,\beta})$ for each choice of the roots α, β. The length of this presentation is $O(\log q)$: each factor appearing in $(B'_{\alpha,\beta})$, and each $\star_\alpha$, has that length by Theorem 4.36(b).

General case (excluding $\mathrm{SU}(2n+1,2)$)**:** Let $z(G)$ be the least common multiple of the integers $z(L_{\alpha_i,\alpha_j})$ for the various rank 2 groups L_{α_i,α_j} just considered. Introduce elements $x_{\alpha\mu} \in U_\alpha$, $1 \le \mu \le z(G)$, that behave as before.

Theorem 5.2 uses all L_{α_i,α_j}, including ones of type $A_1 \times A_1$ occurring when the roots α_i and α_j are perpendicular. Therefore we need to consider the corresponding relations $(B'_{\alpha,\beta})$ and $(H'_{\alpha,\beta})$ for α and β in the root system spanned by α_i and α_j. As above, in the $A_1 \times A_1$ case one relation $(B'_{\alpha,\beta})$ with $x_{\alpha\mu} \in U_\alpha \backslash \Phi(U_\alpha)$ and $x_{\beta\nu} \in U_\beta \backslash \Phi(U_\beta)$ implies all relations $(B_{\alpha,\beta})$.

As before we see that all (P_α), $(B'_{\alpha,\beta})$ and $(H'_{\alpha,\beta})$ imply all relations $(B_{\alpha,\beta})$. Once again, $J \cong G$.

The case $\mathrm{SU}(2n+1,2)$**:** For each long root α lying in the span of two fundamental roots, we introduce two elements $x_{\alpha\mu}$ that generate the quaternion group U_α. We can now proceed exactly as above.

Summary: Each relation $(B'_{\alpha,\beta})$ involves a bounded number of elements of various rank 1 groups L_γ; each of these elements has length $O(\log q)$ in the generators of L_γ by Theorem 4.36(b). A similar statement holds for $(H'_{\alpha,\beta})$. We needed to consider $O(n^2)$ pairs α, β.

Centers: While this is a presentation for the universal group G, the center can be killed exactly as in [BGKLP, Sec. 5.2]. For example, for groups of type A_n there is a standard product of the form $\prod_1^{n-1} h_{\alpha_i}(\xi^i)$, $|\xi| = (q-1, n+1)$, that generates the center. By Theorem 4.36(b), this expression has length $O(n \log q)$ in our generators. Thus, within our length requirements we can factor out all or part of the resulting cyclic group (a very different approach to this is used later in Section 6). For the remaining types it is also easy to write the required central elements as short words in our generators.

We have now handled all cases of Theorem 5.1. □

Note that this presentation often involves more elements $x_{\alpha\mu}$ and relations $(B'_{\alpha,\beta})$ than are actually needed. For example, for odd-dimensional unitary groups we never need 20 coset representatives: we only need at most 4 or 5 for any pair α, β of roots (cf. Lemma 5.4). Different examples occur in [GKKL2].

Remark 5.5. For future reference we mention two elements of $\mathrm{SL}(n-1,q)$ that can be readily found using the above generators. By Theorem 4.36(b),

$$c_{23} := r_{\alpha_2}^2 r_{\alpha_1} = \begin{pmatrix} 1 & 0 & 0 & \\ 0 & 0 & 1 & O \\ 0 & -1 & 0 & \\ & O & & I \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & \\ -1 & 0 & 0 & O \\ 0 & 0 & 1 & \\ & O & & I \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 & 0 & \\ 0 & 0 & 1 & O \\ 0 & 1 & 0 & \\ & O & & I \end{pmatrix}$$

is an element of $L_{\alpha_2} L_{\alpha_1}$, and hence has length 3 in our generators. In the monomial action on the standard basis $e_1, \dots, e_n$, we have $c_{23} = (e_1, -e_1)(e_2, e_3)$. Hence,

$$c_{2345} := (e_1, -e_1)(e_2, e_3) \cdot (e_1, -e_1)(e_3, e_4) \cdot (e_1, -e_1)(e_4, e_5) = (e_1, -e_1)(e_2, e_3, e_4, e_5)$$

has length 9. Compare Proposition 5.6.

5.3. **Word lengths.** The preceding remark is a very special case of the following observation:

Proposition 5.6. *If n is bounded in* Theorem 5.1, *then every element of every central extension of G is a word of length $O(\log q)$ in the generators used in the theorem.*

Proof. The elements r_{α_i} generate the Weyl group W modulo H. By hypothesis, $|\Phi|$ and $|W|$ are bounded. By the Bruhat decomposition, $\tilde{G} = \bigcup_{w \in W} BwB$ with $B = UH$ a Borel subgroup. Here U and H are products of fewer than $|\Phi|$ subgroups of $L_\alpha, \alpha \in \Phi$. Since each L_α is a W-conjugate of some L_{α_i}, Theorem 4.36(b) implies the result. □

6. Theorem A

We now turn to Theorem A. By Theorem 5.1, we only need to consider the classical groups of rank greater than 8.

6.1. **A presentation for $\mathrm{SL}(n,q)$.** We begin with the case $\mathrm{SL}(n,q)$. We would like to use the Weyl group S_n in our presentation, but $\mathrm{SL}(n,q)$ does not have a natural subgroup S_n when q is odd. There are various ways around this difficulty, such as using a subgroup $(2^n \rtimes S_n) \cap \mathrm{SL}(n,q)$ or the alternating group (compare [GKKL2]). We have chosen to use S_{n-1} as an adequate substitute for S_n.

Theorem 6.1. *All groups* $\mathrm{SL}(n,q)/Z$, *where* $Z \le Z(\mathrm{SL}(n,q))$, *have bounded presentations of length* $O(\log n + \log q)$.

Clearly the most interesting cases are $Z = 1$ or $Z(\mathrm{SL}(n,q))$. However, later we will also need the case $|Z| = 2$.

Proof. We start with two bounded presentations of length $O(\log n + \log q)$ given in Theorems 3.17 and 5.1:

- $T = \langle X \mid R(X)\rangle$ of S_{n-1} (acting on $\{1,2,\dots,n\}$, fixing 1), and
- $F = \langle \tilde{X} \mid \tilde{R}\rangle$ of $\mathrm{SL}(5,q)$.

  The notation $R(X)$ is used here since later we will need to use a second copy $\langle \bar{X} \mid R(\bar{X})\rangle$ of S_n.

  We assume that X and $\tilde{X}$ are disjoint, and that these presentations satisfy the following additional conditions for some $X_1, X_2 \subseteq X$ and $\tilde{Y} \subset \tilde{X}$:
  - (i) $\langle X_1 \cup X_2\rangle$ projects onto the stabilizer T_2 of 2.
  - (ii) X_1 projects into A_{n-1}, and each element of X_2 projects outside A_{n-1}.
  - (iii) $x_{(2,3)}$, $x_{(2,3,4,5)}$, $x_{(6,7)}$, $x_{(5+\delta,\dots,n)}$ are words of length $O(\log n)$ in X that project onto $(2,3)$, $(2,3,4,5)$, $(6,7)$, $(5+\delta,\dots,n)$, respectively, where $\delta = (2, n-1)$.
  - (iv) x_σ is a word of length $O(\log n)$ in X and projects onto $\sigma = (2,\dots,n)$ (this is needed only for handling the center of $\mathrm{SL}(n,q)$).
  - (v) The elements u, t, h_ζ in (3.2) and (4.18), that generate a subgroup $L = \mathrm{SL}(2,q)$, are in $\tilde{X}$. When q is odd, we also assume that $h_2 \in \tilde{X}$. (Here we are using the subgroup $\begin{pmatrix} \mathrm{SL}(2,q) & O \\ O & I \end{pmatrix}$ of $\mathrm{SL}(5,q)$.)
  - (vi) c_{23} and c_{2345} are words of length $O(\log q)$ in $\tilde{X}$ acting as in Remark 5.5.
  - (vii) $\langle \tilde{Y}\rangle = \mathrm{SL}(4,q)$, $|\tilde{Y}| = 2$ and the members of $\tilde{Y}$ have length $O(\log q)$ in $\tilde{X}$. (Here we are using the subgroup $\begin{pmatrix} 1 & O \\ O & \mathrm{SL}(4,q) \end{pmatrix}$ of $\mathrm{SL}(5,q)$.)
  - (viii) X contains a set of generators of the stabilizer $T_{23} = T_2 \cap T_3$ of both 2 and 3. (This will only be needed later when we deal with some orthogonal groups in Section 6.2, Case 2.)

Existence: In Remark 3.22(i) we constructed the permutations in (iii) and (iv); in Remark 3.22(ii, iii) we noted that conditions (i) and (ii) hold; and Remark 3.22(iii) also takes care of (viii). Remark 5.5 finds the elements in (vi). Note that the group $L = L_{\alpha_1}$, and a presentation for it using the generators in (v) among others, were essential ingredients in Section 5.2 when $q > 9$. For (vii) see Proposition 5.6.

Now our presentation is as follows:

Generators: $X \cup \tilde{X}$ (we are thinking of T and F as embedded in $\mathrm{SL}(n,q)$ as $\begin{pmatrix} \pm 1 & O \\ O & \text{permutations} \end{pmatrix}$ using permutation matrices, and $\begin{pmatrix} \mathrm{SL}(5,q) & O \\ O & I \end{pmatrix}$, respectively).

Relations:

(1) $R(X) \cup \tilde{R}$.
(2) $u^{x_1} = u, \bar{u}^{x_1} = \bar{u}$, for all $x_1 \in X_1$ (where we have abbreviated $\bar{u} := u^t$).
(3) $u^{x_2} = u^{-1}, \bar{u}^{x_2} = \bar{u}^{-1}$, for all $x_2 \in X_2$.
(4) $[h_\zeta, X_1 \cup X_2] = 1$.
(5) $x_{(2,3)} = c_{23}, x_{(2,3,4,5)} = c_{2345}$.

We will show that *the group G defined by this presentation is isomorphic to* $\mathrm{SL}(n,q)$. There is a natural surjection $\pi\colon G \to \mathrm{SL}(n,q)$. (For, in view of (ii), relation (3) express the fact that each odd permutation maps the first basis vector to its negative, while the even permutations in (2) fix that vector, as do the elements of $\tilde{Y}$.)

By (1) and Lemma 2.3, G has subgroups we can identify with $T = \langle X \rangle$ and $F = \langle \tilde{X} \rangle$.

We have $\langle u, \bar{u}, h_\zeta, x_{(2,3,4,5)} \rangle = F$ since $\langle u, \bar{u}, h_\zeta \rangle$ is $L = \mathrm{SL}(2,q)$. By (i), (2), (3) and (4), $X_1 \cup X_2$ acts on $\{u^{\pm 1}, \bar{u}^{\pm 1}, h_\zeta\}$, so that the elements $x_{(6,7)}$ and $x_{(5+\delta,\ldots,n)}$ of $T_2 = \langle X_1 \cup X_2 \rangle$ act on F. Moreover, $\langle x_{(2,3)}, x_{(2,3,4,5)} \rangle = \langle c_{23}, c_{2345} \rangle < F$ by (5). Thus,

$$|N_T(F)| \geq |\langle x_{(2,3)}, x_{(2,3,4,5)}, x_{(6,7)}, x_{(5+\delta,\ldots,n)} \rangle| = |S_4 \times S_{n-5}|,$$

while $|F^T| \geq |\pi(F^T)| = \binom{n-1}{4}$. It follows that T acts on F^T as it does on the set of all 4-sets in $\{2, \ldots, n\}$.

The subgroups $U_{12} = \langle u^{\langle h_\zeta \rangle} \rangle$ and $U_{21} = \langle \bar{u}^{\langle h_\zeta \rangle} \rangle$ of F are root groups order q. Once again, by (i), (2), (3) and (4), $T_2 = \langle X_1 \cup X_2 \rangle$ normalizes each of them. As above, it follows that T acts on both $(U_{12})^T$ and $(U_{21})^T$ as it does on $\{2, \ldots, n\}$.

Both $U_{31} = (U_{21})^{c_{23}}$ and $U_{23} = [U_{12}, U_{31}]$ are also root groups of F. Since $(T_2)^{c_{23}} = T_3$ normalizes $(U_{21})^{c_{23}} = U_{31}$, it follows that $T_2 \cap T_3$ normalizes U_{23}. As above, we find that $|(U_{23})^T| = (n-1)(n-2)$.

Thus, T acts on both $(U_{12})^T$ and $(U_{21})^T$ as it does on singletons, and on $(U_{23})^T$ as it does on ordered pairs from our $(n-1)$-set. By the 4-transitivity of T, any given pair from $(U_{12})^T \cup (U_{21})^T \cup (U_{23})^T$ can be conjugated into F by a single element of T.

Consequently, $N := \langle L^T \rangle \cong \mathrm{SL}(n,q)$ by the Curtis-Steinberg-Tits presentation (as in Section 5.2 or Lemma 2.10). Moreover, $N \trianglelefteq G$, and G/N is a homomorphic image of $\langle T \rangle \cong S_{n-1}$ in which, by (5), a transposition is sent to 1. Thus, $G/N = 1$.

This proves the theorem when $Z = 1$.

Now that we have a presentation for $\mathrm{SL}(n,q)$, we need to factor out an arbitrary subgroup Z of its center. However, a generator of Z probably cannot be written as a short word in our present generators. In order to deal with this obstacle we will use the following lemma to increase our generating set:

Lemma 6.2. *Suppose that G is a finite group containing the wreath product $H \wr S_m = S_m \ltimes H^m$ for some group H. Let $\langle Y \mid S \rangle$ be a presentation for G of length $< l$ and $\langle \bar{Y} \mid \bar{S} \rangle$ a presentation for S_m of length $< l$. Fix $\epsilon \in H$ and let $h_{i,j}$ be the element in H^m with coordinates 1 everywhere aside from ith coordinate ϵ and jth coordinate ϵ^{-1}. Assume that*

(a) *The following elements of the standard subgroup $S_m < S_m \ltimes H^m$ can be written as words of length $< l$ in Y:* $\sigma = (1, 2, \dots, m)$, $\mu = (2, \dots, m)$ *and* $\tau_2 = (2, 3)$;
(b) *The elements $\bar\sigma$ and $\bar\tau_2$ in $\langle \bar Y \mid \bar S\rangle$ corresponding to the above ones can be written as words of length $< l$ in $\bar Y$; and*
(c) *$h_{1,2}$ can be written as a word of length $< l$ in Y.*

Then the following is a presentation for G, of length $< 13l$:

Generators: $Y \cup \bar Y \cup \{d\}$.

Relations:

(1) $S \cup \bar S$.
(2) $\bar\sigma = \sigma d$.
(3) $\bar\tau_2 = \tau_2 h_{2,3}$.
(4) $[d, \tau_2] = [d, \mu] = [d, h_{2,3}] = 1$.

Moreover, d maps to the element $(\epsilon^{1-m}, \epsilon, \dots, \epsilon) \in H^m < G$.

Note that Y and $\bar Y$ are unrelated. It is the subgroups $\langle \sigma, \tau\rangle$ and $\langle \bar\sigma, \bar\tau\rangle$, both isomorphic to S_m, that are related via the isomorphism "bar"; and we will see that they are conjugate in $S_m \ltimes H^m$. Also note that ϵ is used to define the element $h_{2,3}$ and hence is essential for (3) and (4).

Proof. If $\pi \in S_m$ and $(a_1, \dots, a_m) \in H$, then in $S_m \ltimes H^m$ we have

$$(a_1, \dots, a_m)\pi = \pi(a_{\pi(1)}, \dots, a_{\pi(m)}).$$

In particular, $(a_1, \dots, a_m)^{\sigma^{-1}} = (a_m, a_1, \dots, a_{m-1})$ is "pushing to the right".

Let J be the group presented above. It is straightforward to check that J surjects onto G, by taking d to be $(\epsilon^{1-m}, \epsilon, \dots, \epsilon)$ and letting the new copy of S_m be conjugate to the standard copy via $(\epsilon^{-1}, \epsilon^{m-2}, \dots, \epsilon^2, \epsilon, 1) \in H^m$. It is also easy to check that this presentation has length $< 13l$.

By Lemma 2.3, J has subgroups we can identify with $G = \langle Y\rangle$ and $S_m = \langle \bar Y\rangle$. It suffices to prove that $J \le G$.

Relations (4) imply that d commutes with $\tau_k = (k, k+1)$ and $h_{k,k+1}$ whenever $2 \le k < m$. Hence, by (2) and (3),

$$\bar\tau_3 = (\bar\tau_2)^{\bar\sigma^{-1}} = (\tau_2 h_{2,3})^{d^{-1}\sigma^{-1}} = \left(\tau_2^{d^{-1}} h_{2,3}^{d^{-1}}\right)^{\sigma^{-1}} = (\tau_2 h_{2,3})^{\sigma^{-1}} = \tau_3 h_{3,4}.$$

Similarly, induction gives $\bar\tau_k = \tau_k h_{k,k+1}$ whenever $2 \le k < m$. A similar equation also holds for $\tau_m := (m, 1)$:

$$\bar\tau_m = (\bar\tau_{m-1})^{\bar\sigma^{-1}} = (\tau_{m-1} h_{m-1,m})^{d^{-1}\sigma^{-1}} = (\tau_{m-1} h_{m-1,m})^{\sigma^{-1}} = \tau_m h_{m,1}.$$

Consequently, $\bar\tau_2, \dots, \bar\tau_m \in S_m \ltimes H^m \le G$. It follows that our second copy $\langle \bar Y\rangle$ of S_m lies in G, and hence so do $\bar\sigma$, $\bar Y$ and d (by (2)). Thus, $J = \langle Y \cup \bar Y \cup \{d\}\rangle \le G$, as required.

In order to explain the motivation used here, we include the following direct calculation that d behaves as desired:

$$\begin{aligned}
\bar\sigma &= \bar\tau_2 \bar\tau_3 \cdots \bar\tau_{m-1} \bar\tau_m \\
&= \tau_2 h_{2,3} \tau_3 h_{3,4} \cdots \tau_{m-1} h_{m-1,m} \tau_m h_{m,1} \\
&= \tau_2 \tau_3 \cdots \tau_{m-1} \tau_m h_{2,1} h_{3,1} \cdots h_{m-1,1} h_{m,1} \\
d &= \sigma^{-1}\bar\sigma = h_{2,1} h_{3,1} \cdots h_{m-1,1} h_{m,1}. \quad \square
\end{aligned}$$

When we use this lemma, the bound l will be taken to be $O(\log n + \log q)$.

Completion of the proof of Theorem 6.1: Let $Z = \langle \epsilon I \rangle$. In the preceding lemma let $m = n - 1$ and let $H = \langle \epsilon \rangle$ be the subgroup of $\mathbb{F}_q^*$ of order $|\epsilon|$. Embed H^{n-1} into $G = \mathrm{SL}(n, q)$ as all $\mathrm{diag}(\epsilon_1, \ldots, \epsilon_n)$ with $\epsilon_i \in \langle \epsilon \rangle$ and $\prod \epsilon_i = 1$. We already used a presentation $\langle X \mid R(X) \rangle$ for S_{n-1} above in (i), where S_{n-1} permutes the last $n-1$ coordinates. We use the corresponding presentation $\langle \bar{X} \mid R(\bar{X}) \rangle$ for a second copy of S_{n-1} (with X and $\bar{X}$ disjoint).

First we need to verify conditions (a)-(c) of the lemma. We temporarily use the notation in Remark 3.22(i): $\sigma = (\mathbf{2}, \mathbf{3}, \ldots, \mathbf{n})$ and $z = (\mathbf{2}, \mathbf{3})$ have the required length, hence so do $\mu = (\mathbf{2}, \mathbf{3})\sigma$ and $(\mathbf{3}, \mathbf{4}) = (\mathbf{2}, \mathbf{3})^{\sigma^{-1}}$. It follows that Lemma 6.2(a) holds, and hence so does (b). It remains to consider (c). The group $L = \mathrm{SL}(2, q)$ was defined above in (v). Use Theorem 4.36(b) to write the element $d' = \mathrm{diag}(\epsilon, \epsilon^{-1}, 1, \ldots, 1) \in L$ as a word of length $O(\log q)$ in our generators (cf. (v)). Then $h_{2,3} = d'^{\sigma^{-1}}$ has the length required in (c).

The lemma provides a new bounded presentation for G of length $O(\log n + \log q)$, including a new generator d representing the diagonal matrix $\mathrm{diag}(1, \epsilon^{2-n}, \epsilon, \ldots, \epsilon)$. Then $Z = \langle dd' \rangle$, so that the additional relation $dd' = 1$ produces the desired factor group. □

6.2. **Generic case.** In this section we obtain short bounded presentations for the universal central extensions of the simple groups of Lie type. This is significantly simpler than dealing with the simple groups, which involves factoring out the centers of the universal extensions. However, in some cases the latter is easy: it is a matter of choosing slight variations on the subgroups we amalgamate.

We begin with some general properties of these universal central extensions:

Lemma 6.3. *Let G be a simple classical group defined on a vector space of dimension > 8, and let $\hat{G}$ denote its universal central extension.*

(1) *If $G = \mathrm{PSL}(n, q)$ then $\hat{G} = \mathrm{SL}(n, q)$.*
(2) *If $G = \mathrm{PSU}(n, q)$ then $\hat{G} = \mathrm{SU}(n, q)$.*
(3) *If $G = \mathrm{PSp}(2n, q)$ then $\hat{G} = \mathrm{Sp}(2n, q)$.*
(4) *If $G = \Omega(2n + 1, q)$ then $|Z(\hat{G})| = (2, q - 1)$.*
(5) *If $G = \mathrm{P}\Omega^+(2n, q)$ then $|Z(\hat{G})| = (4, q^n - 1)$, and $Z(\hat{G})$ is cyclic unless n is even.*
(6) *If $G = \mathrm{P}\Omega^-(2n, q)$ then $Z(\hat{G})$ is cyclic of order $(4, q^n + 1)$.*
(7) *Suppose that $G = \Omega(V)$ is an orthogonal group over a field of odd characteristic. Then $Z(\hat{G})$ has an involution, the* spin involution, *lying in $Z(\hat{H})$ whenever $H = \Omega(U)$ for a nondegenerate subspace U of V of dimension at least 3.*

Proof. See [GLS, pp. 312-313] for (1)-(6) and [GLS, Proposition 6.2.1(b)] for (7). □

Note that there are various exceptions to (1)-(6) when the dimension is at most 8 [GLS, p. 313].

We will use the following crucial theorem as well as variations on its proof.

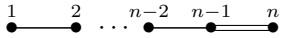
Theorem 6.4. *All universal central extensions of groups of Lie type of rank n over $\mathbb{F}_q$ have bounded presentations of length $O(\log n + \log q)$.*

Proof. By Theorem 5.1, we may assume that G is a classical simple group of rank $n > 8$. We will use the root system of G and the Curtis-Steinberg-Tits presentation

(Theorem 5.2). For each root α there is a corresponding subgroup L_α that is a central extension of $\mathrm{PSL}(2,q)$, $\mathrm{PSL}(2,q^2)$ or $\mathrm{PSU}(3,q)$. We may assume that $\Pi = \{\alpha_1, \dots, \alpha_n\}$ is a fundamental system of roots (in the standard order), where $\alpha_1, \dots, \alpha_{n-1}$ have the same length.

Case 1: G does not have type D_n. Let

$$\Pi_1 = \{\alpha_1, \dots, \alpha_{n-1}\},\ \Pi_2 = \{\alpha_1, \dots, \alpha_{n-2}, \alpha_n\} \text{ and } \Pi_3 = \{\alpha_{n-1}, \alpha_n\}.$$



Remark 6.5. *These sets of roots have the following properties*:

(a) For $i = 1, 2, 3$, either $|\Pi_i| = 2$ or $G_i := \langle L_\beta \mid \beta \in \Pi_i \rangle$ is of type SL_d, $\mathrm{SL}_d \times \mathrm{SL}_2$ or $\mathrm{SL}_d \times \mathrm{SU}_3$ for some d; and
(b) Each pair from Π lies in some Π_i.

Namely, (b) is clear, and G_1 is a homomorphic image of $\mathrm{SL}(n,q)$ or $\mathrm{SL}(n,q^2)$ (as is seen from the commutator relations in Section 5 – or more precisely in the case of odd-dimensional unitary groups, from the explicit relations in [Gr, BGKLP]). Moreover, G_2 has type $\mathrm{SL}_{n-1} \times \mathrm{SL}_2$ or $\mathrm{SL}_{n-1} \times \mathrm{SU}_3$, while $|\Pi_3| = 2$, which proves the remark.

The root α_i $(1 \le i < n)$ can be identified with the ordered pair $(i, i+1)$ of elements of the set $\{1, \dots, n\}$ on which the Weyl group S_n of G_1 acts. The corresponding root groups $X_{\pm\alpha_i}$ are just groups of elementary matrices of G_1, and generate a subgroup $L_{\alpha_i} = L_{i,i+1} \cong \mathrm{SL}(2,q)$ or $\mathrm{SL}(2,q^2)$ acting on the span of two of the standard basis vectors of the usual n-dimensional module for G_1.

We use the following groups and presentations.

- G_1 is the universal central extension obtained in Theorem 6.1 (but *reversing* the order of the set $\{1, \dots, n\}$), given with a bounded presentation $\langle X_1 \mid R_1 \rangle$ of length $O(\log n + \log q)$ including
  (i) generators inside X_1 for $T = S_{n-1}$ acting on $\{1, \dots, n\}$ and fixing n, and for T'_{n-1}, the stabilizer of $n-1$ in $T' = A_{n-1}$;
  (ii) a bounded presentation $\langle X(n-1) \mid R(n-1) \rangle$ of length $O(\log q)$ for the subgroup $L_{\alpha_{n-1}}$; and
  (iii) a set $X(n-2)$ of generators for the subgroup $L_{\alpha_{n-2}}$ of G_1.

  Existence: Note that Remark 3.22(iii) provides short generators for T'_{n-1}, as required in (i). The groups $L_{\alpha_{n-1}}$ and $L_{\alpha_{n-2}}$ were used in the presentation for F in the proofs of Theorems 5.1 and 6.1, so that we also already have $X(n-2) \cup X(n-1) \subseteq X_1$ and $R(n-1) \subseteq R_1$, as required in (ii) and (iii).

  Then $\langle L_{\alpha_{n-2}}, T'_{n-1} \rangle \cong \mathrm{SL}(n-1,q)$ or $\mathrm{SL}(n-1,q^2)$, since T'_{n-1} fixes the corresponding standard basis vector (whereas each odd permutation in T_{n-1} sends that vector to its negative; cf. Section 6.1).
- $G_3 = \langle X_3 \mid R_3 \rangle$ was obtained in Section 5 by using the root system Π_3 and bounded presentations of length $O(\log q)$ for two rank 1 groups, namely
  (iv) the *same* presentation $\langle X(n-1) \mid R(n-1) \rangle$ used above for $L_{\alpha_{n-1}}$; and
  (v) a bounded presentation $\langle X(n) \mid R(n) \rangle$ of length $O(\log q)$ for L_{α_n}.
  Hence, we already have $X(n-1) \cup X(n) \subseteq X_3$ and $R(n-1) \cup R(n) \subseteq R_3$.

Our presentation is as follows:

Generators: $X_1 \cup X_3$.
Relations:

(1) $R_1 \cup R_3$.
(2) Identify $L_{\alpha_{n-1}}$ inside G_1 and inside G_3 using the identity map.
(3) $[L_{\alpha_{n-2}}, L_{\alpha_n}] = [T'_{n-1}, L_{\alpha_n}] = 1$.

Then $G_2 := \langle L_{\alpha_{n-2}}, T'_{n-1}\rangle \cdot L_{\alpha_n} = \langle L_\beta \mid \beta \in \Pi_1 \cap \Pi_2 \text{ or } \Pi_2 \cap \Pi_3\rangle$ is a central product of an SL_{n-1} and an SL_2 or SU_3. In view of Remark 6.5(b), we obtain the desired universal central extension $\hat{G}$ by Theorem 5.2. This presentation is clearly short and bounded.

Case 2: G has type D_n. We assume that α_{n-1} and α_n are connected to α_{n-2} in the Dynkin diagram. This time we use

$$\Pi_1 = \{\alpha_1, \dots, \alpha_{n-1}\},\ \Pi_2 = \{\alpha_1, \dots, \alpha_{n-3}, \alpha_n\} \text{ and } \Pi_3 = \{\alpha_{n-2}, \alpha_{n-1}, \alpha_n\}.$$

Once again Remark 6.5 holds, this time with G_1, G_2 and G_3 the universal central extensions $\mathrm{SL}(n,q)$, $\mathrm{SL}(n-2,q) \times \mathrm{SL}(2,q)$ and $\mathrm{SL}(4,q)$, respectively. This time we use the following presentations.

- G_1 is the universal central extension obtained in Theorem 6.1 (but once again *reversing* the order of the set $\{1, \dots, n\}$), given with a bounded presentation $\langle X_1 \mid R_1\rangle$ of length $O(\log n + \log q)$ including
  (i) a bounded presentation $\langle X(n-1) \mid R(n-1)\rangle$ of length $O(\log q)$ for $L_{\alpha_{n-1}}$;
  (ii) a bounded presentation $\langle X(n-2) \mid R(n-2)\rangle$ of length $O(\log q)$ for $L_{\alpha_{n-2}}$;
  (iii) a bounded presentation $\langle X(n-3) \mid R(n-3)\rangle$ of length $O(\log q)$ for $L_{\alpha_{n-3}}$; and
  (iv) generators inside X_1 for $T = S_{n-1}$ acting on $\{1, \dots, n\}$ and fixing n, and for the stabilizer $T'_{n-1,n-2}$ of both $n-1$ and $n-2$ in $T' = A_n$.
  The groups $L = L_{\alpha_{n-1}}, L_{\alpha_{n-2}}, L_{\alpha_{n-3}}$ were used in the presentation for F in the proof of Theorem 6.1, so that $\cup_{n-3}^{n-1} X(i) \subseteq X_1$ and $\cup_{n-3}^{n-1} R(i) \subseteq R_1$. Generators for $T'_{n-1,n-2}$ were obtained in Remark 3.22(iii). (Recall that the numbering in that remark has been reversed in this proof.) This time $\langle L_{\alpha_{n-3}}, T'_{n-1,n-2}\rangle \cong \mathrm{SL}(n-2,q)$.
- $\langle X_3 \mid R_3\rangle$ is a bounded presentation of length $O(\log q)$ for $G_3 = \mathrm{SL}(4,q)$, obtained as in Section 5 using the presentations (i) and (ii), together with the root system Π_3 and
  (v) a bounded presentation $\langle X(n) \mid R(n)\rangle$ of length $O(\log q)$ for L_{α_n}.
  Note that $\cup_{n-2}^{n} X(i) \subseteq X_3$ and $\cup_{n-2}^{n} R(i) \subseteq R_3$.

Our presentation is as follows:

Generators: $X_1 \cup X_3$.
Relations:

(1) $R_1 \cup R_3$.
(2) Identify $\langle L_{\alpha_{n-2}}, L_{\alpha_{n-1}}\rangle \cong \mathrm{SL}(3,q)$ inside G_1 and inside G_3 using the identity map.
(3) $[L_{\alpha_{n-3}}, L_{\alpha_n}] = [T'_{n-1,n-2}, L_{\alpha_n}] = 1$.

Once again $G_2 := \langle L_{\alpha_{n-3}}, T'_{n-1,n-2}\rangle \cdot L_{\alpha_n} = \langle L_\beta \mid \beta \in \Pi_1 \cap \Pi_2 \text{ or } \Pi_2 \cap \Pi_3\rangle$ behaves as required. In view of Remark 6.5(b), we obtain the desired universal central extension $\hat{G}$ by Theorem 5.2. □

6.3. **Symplectic groups.** We will need to factor out the centers of the various groups in Theorem 6.4. The following simple observation will make this easy in many cases:

Lemma 6.6. *Suppose that q is odd.*

(i) *If n is even then $\mathrm{SL}(n,q)$ has a bounded presentation of length $O(\log n + \log q)$ in which the involution z in the center of $\mathrm{SL}(n,q)$ is a word of length $O(\log n + \log q)$ in the generators.*

(ii) *If n is odd then $\mathrm{SL}(n,q)$ has a bounded presentation of length $O(\log n + \log q)$ in which the central involution z in any given Levi subgroup $\mathrm{SL}(n-1,q)$ is a word of length $O(\log n + \log q)$ in the generators.*

Proof. Let $G = \mathrm{SL}(n,q)$.

(i) In the completion of the proof of Theorem 6.1, we found a presentation for G in which $z = dd'$ behaves as stated when ϵ is chosen to be -1.

(ii) Here we proceed exactly as in the proof of Theorem 6.4 for the present group G instead of the other classical groups G considered in that theorem, using the presentation for $G_1 = \mathrm{SL}(n-1,q)$ obtained in (i). □

Note that (ii) also can be proved using Lemma 6.2.

Corollary 6.7. $\mathrm{PSp}(2n,q)$ *has a bounded presentation of length $O(\log n + \log q)$.*

Proof. The proof of Theorem 6.4 for $\hat{G} = \mathrm{Sp}(2n,q)$ involved a presentation for $G_1 = \mathrm{SL}(n,q)$. This time we will use the presentation for $\mathrm{SL}(n,q)$ in the preceding lemma.

If n is even then the additional relation $z = 1$ in Lemma 6.6(i) produces the desired presentation.

Suppose that n is odd. Let $e_1, \dots, e_n, f_1, \dots, f_n$ be a hyperbolic basis of the $\mathbb{F}_q$-space underlying the group $\hat{G}$. Choose this basis so that L_{α_i} has support $\langle e_i, e_{i+1}, f_i, f_{i+1}\rangle$ for $i < n$ and $\langle e_n, f_n\rangle$ for $i = n$. Let $H = \langle X_{\pm\alpha_i} \mid i \le n-2\rangle = \mathrm{SL}(n-1,q)$, with support the orthogonal complement $\langle e_1, \dots, e_{n-1}, f_1, \dots, f_{n-1}\rangle$ of the support $\langle e_n, f_n\rangle$ of $L_{\alpha_n} \cong \mathrm{Sp}(2,q)$. By Lemma 6.6(ii), the involution $z \in Z(H)$ is a word of length $O(\log n + \log q)$ in our generators. On the other hand, the central involution z' of L_{α_n} has length $O(\log q)$ by Theorem 4.36(b). Hence, the additional relation $z = z'$ produces the desired presentation. □

6.4. **Orthogonal groups.**

Theorem 6.8. *All perfect central extensions of simple orthogonal groups of dimension N over $\mathbb{F}_q$ have bounded presentations of length $O(\log N + \log q)$.*

Proof. We may assume that $N > 8$ and the center of $\hat{G}$ is nontrivial. By Lemma 6.3, q is odd and there are various quotient groups to consider.

6.4.1. *Factoring out the spin involution.* For each simple orthogonal group G we need to factor out $\langle s\rangle$ from $\hat{G}$, where s is the spin involution in Lemma 6.3(7).

Assume that G does not have type D_n. By Lemma 6.3(7), s is the central involution in $L_{\alpha_n} \cong \mathrm{SL}(2,q)$ or $\mathrm{SL}(2,q^2)$ (where the central quotient groups are

the orthogonal groups $\mathrm{P}\Omega(3,q)$ or $\mathrm{P}\Omega^-(4,q)$). By Theorem 4.36(b), s has length $O(\log q)$ in our generators. Thus, the additional relation $s=1$ produces the desired presentation.

If G has type D_n we use $\langle L_{\alpha_{n-1}}, L_{\alpha_n}\rangle \cong \mathrm{SL}(2,q)\times\mathrm{SL}(2,q)$. This time s is the central involution lying in neither $\mathrm{SL}(2,q)$ factor. Once again s has length $O(\log q)$ in our generators, and the additional relation $s=1$ produces the desired presentation.

By Lemma 6.3, we have now provided a short bounded presentation for $\hat{G}/\langle s\rangle$ in all cases; this proves Theorem 6.8 when G is $\Omega(2n+1,q)$, $\mathrm{P}\Omega^-(2m,q)$ with m even, or $\mathrm{P}\Omega^\epsilon(2m,q)$ with m odd and $q\equiv -\epsilon 1 \pmod 4$, where $\epsilon=\pm$.

6.4.2. *$D_n(q)$ with n even.* Here we merely repeat the argument in Case 2 of the proof of Theorem 6.4, this time using the presentation for $G_1=\mathrm{SL}(n,q)$ provided by Lemma 6.6(i). The additional relation $z=1$ produces the desired presentation.

At this point we have a presentation for $G=\mathrm{P}\Omega^+(2n,q)$, as well as for its universal cover $\hat{G}$. The center of $\hat{G}$ is elementary abelian of order 4. Starting with G, use Proposition 2.4 twice (with $p=2$) in order to obtain a presentation for $\hat{G}$ in which we have words of length $O(\log n+\log q)$ for the generators a,b of $Z(\hat{G})$. Then each group $\hat{G}/\langle a\rangle$, $\hat{G}/\langle b\rangle$, $\hat{G}/\langle a,b\rangle$ can be obtained by adding one or two relations of length $O(\log n+\log q)$. (Of course, one of these groups is $\hat{G}/\langle s\rangle$, which was dealt with in Section 6.4.1.)

Note that earlier we passed from the universal central extension of a simple group to quotients of it. The above argument went in the opposite direction: from a simple group to all perfect central extensions, in the case where the center of the universal central extension was bounded. This same idea already occurred in Corollary 3.23.

6.4.3. *Factoring out -1.* It remains to consider the orthogonal group $\mathrm{P}\Omega^\epsilon(2m,q)$ where m is odd and $q\equiv\epsilon 1 \pmod 4$.

It is convenient to rename G as $\hat{G}/\langle s\rangle = \Omega^\epsilon(2m,q)$. We will obtain a new presentation for G in order to factor out $\langle -1\rangle$.

Case 1: $G=\Omega^-(2m,q)$ with m and q odd. Here the rank is $n=m-1$, which is even. Use Lemma 6.6(i) to obtain a bounded presentation of length $O(\log n+\log q)$ for $G_1=\mathrm{SL}(n,q)$ such that its central involution z has length $O(\log n+\log q)$ in the generators.

Use this presentation for G_1 in Case 1 of the proof of Theorem 6.4 in order to obtain a presentation for $\hat{G}$ of length $O(\log n+\log q)$. Pass modulo the spin involution as in Section 6.4.1 in order to obtain a presentation for G. We still need a short bounded presentation for $G/\langle -1\rangle$.

Let $e_1,\dots,e_m,f_1,\dots,f_m$ be a basis of the $\mathbb{F}_q$-space underlying G, where $e_1,\dots,e_n,f_1,\dots,f_n$ is a hyperbolic basis of its span and is perpendicular to the anisotropic 2-space $\langle e_m,f_m\rangle$. We can choose these bases so that L_{α_i} has support $\langle e_i,e_{i+1},f_i,f_{i+1}\rangle$ for each $i\le n$. Then $G_1=\langle L_{\alpha_1},\dots,L_{\alpha_{n-1}}\rangle\cong\mathrm{SL}(n,q)$ has support $\langle e_1,\dots,e_n,f_1,\dots,f_n\rangle$, and z is -1 on this subspace. Note that $\langle e_1,\dots,e_n,f_1,\dots,f_n\rangle^\perp=\langle e_m,f_m\rangle$ has type $\Omega^-(2,q)$, and $-1\in\mathrm{GL}(\langle e_m,f_m\rangle)$ lies in $\Omega^-(2,q)$ since $q\equiv 3 \pmod 4$.

The group $\langle L_{\alpha_{n-1}},L_{\alpha_n}\rangle\cong\Omega^-(6,q)$ has an element z' with support $\langle e_m,f_m\rangle$ and acting as -1 on this 2-space. Use Proposition 5.6 to write z' as a word of length

$O(\log q)$ in our generators. Now add the relation $z = z'$ in order to obtain the desired presentation for $G/\langle -1\rangle = \mathrm{P}\Omega^-(2m, q)$.

Case 2: $G = \Omega^+(2n, q)$ with n and q odd. Here the rank is n. Let $e_1, \dots, e_n, f_1, \dots, f_n$ be a hyperbolic basis of V such L_{α_i} has support $\langle e_i, e_{i+1}, f_i, f_{i+1}\rangle$ for each $i \le n$. Let $G_1 = \mathrm{SL}(n, q) < G$ preserve each of the subspaces $\langle e_1, \dots, e_n\rangle$ and $\langle f_1, \dots, f_n\rangle$. Let $H = \mathrm{SL}(n-1, q)$ preserve $\langle e_1, \dots, e_{n-1}\rangle$ and $\langle f_1, \dots, f_{n-1}\rangle$, inducing 1 on $\langle e_1, \dots, e_{n-1}, f_1, \dots, f_{n-1}\rangle^\perp = \langle e_n, f_n\rangle$.

Use this subgroup H in Lemma 6.6(ii) in order to obtain a presentation for G_1 such that the central involution z of H has length $O(\log n + \log q)$. Use this presentation for G_1 in Case 2 of the proof of Theorem 6.4 in order to obtain a presentation for $\hat{G}$ of length $O(\log n + \log q)$. Pass modulo the spin involution as in Section 6.4.1 in order to obtain a presentation for G.

The element $-1 \in \mathrm{GL}(\langle e_n, f_n\rangle)$ lies in $\Omega^+(2, q)$ since $q \equiv 1 \pmod 4$. This time $\langle L_{\alpha_{n-1}}, L_{\alpha_n}\rangle \cong \Omega^+(4, q)$ has an element z' with support $\langle e_n, f_n\rangle$ and acting as -1 on this 2-space. Use Theorem 4.36(b) to write z' as a word of length $O(\log q)$ in our generators. Now add the relation $z = z'$ in order to obtain the desired presentation for $G/\langle -1\rangle = \mathrm{P}\Omega^+(2n, q)$. $\square$

6.5. **Unitary groups.** We handle unitary groups as in Section 6.2, once again choosing a presentation more carefully.

Theorem 6.9. *All groups $\mathrm{SU}(N, q)/Z$, where $Z \le Z(\mathrm{SU}(N, q))$, have bounded presentations of length $O(\log N + \log q)$.*

Proof. As above we rename G as $\mathrm{SU}(N, q)$, and let V be the underlying $\mathbb{F}_{q^2}$-space. Let $n = [N/2]$, and let the vectors $e_1, \dots, e_n$, $f_1, \dots, f_n$ be a hyperbolic basis for the subspace they span; moreover, let these be perpendicular to a final basis vector v if $N = \dim V$ is odd. We will focus on the case where $N = 2n + 1$ is odd, the even case possibly being slightly easier. We write matrices with our basis ordered $e_1, \dots, e_n, f_1, \dots, f_n, v$.

We can choose the basis so that L_{α_i} has support $\langle e_i, e_{i+1}, f_i, f_{i+1}\rangle$ for $1 \le i \le n-1$, and $\langle e_n, f_n, v\rangle$ when $i = n$. Then the support of $G_1 = \mathrm{SL}(n, q^2)$ is $\langle e_1, \dots, e_n, f_1, \dots, f_n\rangle$, while the support of $G_3 \cong \mathrm{SU}(5, q)$ is $\langle e_{n-1}, e_n, f_{n-1}, f_n, v\rangle$.

In Case 1 of the proof of Theorem 6.4 we used the presentation for G_1 appearing in Theorem 6.1. This time we use the presentation for G_1 deduced from it via Lemma 6.2, using $m = n - 1$, $Z = \langle \epsilon I\rangle$ and $H = \langle \epsilon\rangle \le \mathbb{F}_{q^2}^*$. Here $H^m = H^{n-1}$ is embedded in $\mathrm{SL}(n, q)$ as the diagonal matrices $\mathrm{diag}(\epsilon_1, \dots, \epsilon_n, \epsilon_1, \dots, \epsilon_n, 1)$ with $\epsilon_i \in H$ and $\prod_i \epsilon_i = 1$ (note that $\bar{\epsilon}^{-1} = \epsilon$: this matrix is, indeed, an isometry); while $S_m = S_{n-1}$ is embedded as all matrices

$$\begin{pmatrix} \pi & 0 & & O & O \\ 0 & \pm 1 & & & \\ & & \pi^{-1} & 0 & \\ & O & 0 & \pm 1 & O \\ & & & & \\ & O & & O & 1 \end{pmatrix}$$

using permutation matrices π of sign ± 1. Lemma 6.2 produces a new generator d in a slightly modified short bounded presentation for G_1; the matrix for d is $\mathrm{diag}(\epsilon, \dots, \epsilon, \epsilon^{2-n}, 1, \epsilon, \dots, \epsilon, \epsilon^{2-n}, 1, 1)$ (where we have reordered the coordinates used in that lemma).

Since $\epsilon \in Z(G)$ we have $\epsilon^{2n+1} = 1$, and there is an element $d' \in G_3$ with matrix

$$\mathrm{diag}(1, \dots, 1, \epsilon^{n-1}, \epsilon, 1, \dots, 1, \epsilon^{n-1}, \epsilon, \epsilon)$$

such that $dd' = \epsilon I$. Express d' as a word of length $O(\log q)$ in the generators for G_3 by using Proposition 5.6. The additional relation $dd' = 1$ produces the desired presentation for the quotient group $G/\langle \epsilon I \rangle = G/Z$. □

6.6. **Perfect central extensions.** The Curtis-Steinberg-Tits presentation produces the universal central extension of a group of Lie type. Once we obtained a version of this presentation, we had to factor out all or part of the center. In the process we proved the following additional

Theorem 6.10. *Each perfect central extension of a finite simple group G of rank n over $\mathbb{F}_q$ has a bounded presentation of length $O(\log n + \log q)$.*

Proof. In the preceding parts of Section 6 we proved this theorem for the various classical groups. We dealt with the alternating groups in Corollary 3.23. □

7. Theorems B and B′

In this section, we prove Theorems B and B′ (Holt's Conjecture for simple groups). We recall the following well-known observation (cf. [Ho, Lemma 1.1]):

Lemma 7.1. *If a finite group G has a presentation F/R with F free and R normally generated by r elements, then $\dim_{\mathbb{F}_p} H^2(G, M) \leq r \dim_{\mathbb{F}_p} M$ for any $\mathbb{F}_p G$-module M.*

Proof. A short exact sequence

$$1 \to M \to E \to F/R \to 1$$

is determined (up to equivalence) by a G-homomorphism $R \to M$. Since R can be generated by r elements (as a normal subgroup), any homomorphism is determined by the images of the r generators. Thus, there are most $|M|^r$ such homomorphisms and therefore at most that many inequivalent extensions. Consequently, $\dim_{\mathbb{F}_p} H^2(G, M) \leq r \dim_{\mathbb{F}_p} M$. □

Proof of Theorem B′: By the lemma, Holt's Conjecture follows from Corollary A′ except for the case ${}^2G_2(3^{2e+1})$, which we now handle.

Note that, if D is a subgroup of G that contains a Sylow p-subgroup of G, then the restriction map $H^2(G, M) \to H^2(D, M)$ is injective [Gru, p. 91].

In order to use this, we consider the cases $p = 2$, $p = 3$ and $p > 3$ separately.

If $p > 3$, a Sylow p-subgroup D of G is cyclic, and hence can be presented with one generator and one relation. Applying the preceding lemma to D yields the result for G with constant 1.

A Sylow 2-subgroup of G is elementary abelian of order 8, and hence has a presentation with 3 generators and 6 relations.

Finally, for $p = 3$, by Proposition 4.33 a Borel subgroup of G has a presentation with a bounded number of relations, whence the above lemma produces a bound in characteristic 3 as well. □

We will provide a different argument for this Ree case in [GKKL1].

Working in the profinite category, Theorems B and B′ are equivalent. Namely, in [Lub2] there is a formula for the minimal number $\hat{r}(G)$ of relations needed for a profinite presentation of a finite group G:

$$\hat{r}(G) = \sup_p \sup_M \Big(\Big\lceil \frac{\dim H^2(G,M) - \dim H^1(G,M)}{\dim M} \Big\rceil + d(G) - \xi_M \Big), \tag{7.2}$$

where $d(G)$ is the minimum number of generators for G, p runs over all primes, M runs over all irreducible $\mathbb{F}_p[G]$-modules, and $\xi_M = 0$ if M is the trivial module and 1 if not. Note that $\dim H^1(G,M) \le d(G) \dim M$: every element of $Z^1(G,M)$ is a map $G \to M$ that is completely determined on a generating set of G. Hence, (7.2) implies that

$$h(G) \le \hat{r}(G) \le h(G) + d(G), \tag{7.3}$$

where

$$h(G) := \sup_p \sup_M \frac{\dim H^2(G,M)}{\dim M}.$$

Since $d(G) = 2$ for every finite simple group, (7.3) proves our assertion that Theorems B and B′ are equivalent.

8. Concluding remarks

1. One of the purposes of [BGKLP] was to provide presentations for use in Computational Group Theory. This has turned out to be essential, for example in [KS1, KS2]. When used in these references, those presentations led to efficient algorithms to test whether or not a given matrix group actually is isomorphic to a specific simple group. Short presentations are also essential in [LG, KS2] for gluing together presentations in a normal series (essentially a chief series) in order to obtain a presentation for a given matrix group.

We expect that versions of many of the presentations in the present paper or [GKKL2] will have both practical and theoretical significance in Computational Group Theory.

2. As mentioned in the Introduction, an elementary counting argument shows that our $O(\log n + \log q)$ bound in Theorem A is optimal in terms of n and q. However, in terms of n, p and e (where $q = p^e$) one might hope for presentations providing a slightly better bound: $O(\log n + \log p + \log e)$. We have no idea whether such presentations exist even for $\mathrm{PSL}(2, p^e)$.

3. A standard topological interpretation of Corollary A′ states that all finite simple groups (except perhaps ${}^2G_2(q)$) are fundamental groups of 2–dimensional CW-complexes having a bounded number of cells.

4. As suggested in the Introduction, it is not difficult to check that fewer than 1000 relations are required in Theorem A. It would be interesting to have far better constants in all of the theorems. On the other hand, it would also be interesting to have presentations in Theorem A having only 2 generators, even if the number of relations grew somewhat (while remaining bounded); compare [GKKL2].

Earlier we mentioned Wilson's conjecture that the universal central extension of every finite simple group has a presentation with 2 generators and 2 relations [Wi]. In view of a standard property of the Schur multiplier $M(G)$ [Sch], this would be optimal for a bounded presentation and would imply that every nonabelian finite simple group G has a presentation with 2 generators and $2 + d(M(G))$ relations.

The only infinite family for which the conjecture is presently proven consists of the groups $\mathrm{PSL}(2,p)$ [CR2]. Wilson essentially proved the conjecture for $\mathrm{Sz}(q)$, $q > 8$, in the category of profinite presentations.

5. Finally, in a similar vein, we note that there does not appear to be a known presentation of $\mathrm{SL}(2,p)$, when p is prime, having k generators and $k+1$ relations for some k, and also having length $O(\log p)$. In Section 3.1 we already observed that the presentation of Sunday [Sun] is not this short.

Acknowledgement: We are grateful to the referee for many helpful comments. We also thank Ravi Ramakrishna and Joel Rosenberg for assistance with Lemma 4.23, and James Wilson for checking our computations in Table B.1 using GAP.

Appendix A. Field lemma

Proof of Lemma 4.23. Assume that $ab \neq 0$. Then (4.24) is equivalent to

$$\bar{a}^2 b + \bar{b}^2 a = ab, \quad \bar{a}a + \bar{b}b = 1. \tag{A.1}$$

Taking norms of both sides of the first equation and inserting the second yields

$$\begin{aligned} a^2\bar{a}^2 b\bar{b} + \bar{a}^3 b^3 + a^3\bar{b}^3 + a\bar{a}b^2\bar{b}^2 &= a\bar{a}b\bar{b}(\bar{a}a + \bar{b}b) \\ &= a^2\bar{a}^2 b\bar{b} + a\bar{a}b^2\bar{b}^2. \end{aligned}$$

Therefore, $(a\bar{b})^3 \in \mathbb{F}_q\theta$ with $\theta^q = -\theta \neq 0$.

Consequently, $a\bar{b} = k\theta\omega$ with $k \in \mathbb{F}_q$, $\omega \in F$ and $\omega^3 = 1 = \omega\bar{\omega}$, where $\omega = 1$ if $q \not\equiv -1 \pmod 3$. Substituting $\bar{b} = k\theta\omega/a$ into (A.1) gives

$$-\bar{a}^2 k\theta\omega^{-1}/\bar{a} + a(k\theta\omega/a)^2 = -ak\theta\omega^{-1}/\bar{a}, \quad a\bar{a} - k^2\theta^2/(a\bar{a}) = 1,$$

which imply that

$$k\theta = a(\bar{a} - a/\bar{a}), \quad k^2\theta^2 = a\bar{a}(1 - a\bar{a}).$$

Then $-k^2\theta^2 = k\theta k\bar{\theta} = a\bar{a}(\bar{a} - a/\bar{a})(a - \bar{a}/a)$, which simplifies to

$$2(a\bar{a})^2 = a^3 + \bar{a}^3. \tag{A.2}$$

Any solution of (A.2) yields $(3, q+1)$ solutions of (A.1) corresponding to the different possibilities for ω. For example, using $k := a(\bar{a} - a/\bar{a})/\theta$ and $\bar{b} := k\theta\omega/a$ with $\omega^3 = 1$, we find that $k \in \mathbb{F}_q$ since

$$\bar{k} = -\bar{a}(a - \bar{a}/a)/\theta = a(\bar{a} - a/\bar{a})/\theta = k$$

by (A.2); and hence

$$\begin{aligned} \bar{a}^2 b + a\bar{b}^2 - ab &= b \cdot \bar{a}\bar{b}/\omega + a\bar{b}^2 \\ &= \bar{b}[\bar{a}\bar{\omega} \cdot (a - \bar{a}/a)\bar{\omega} + a(\bar{a} - a/\bar{a})\omega] \\ &= \bar{b}\omega[2a^2\bar{a}^2 - \bar{a}^3 - a^3]/a\bar{a} = 0. \end{aligned}$$

If q is odd then $\mathbb{F}_{q^2} = \mathbb{F}_q[\theta]$, so we can write $a = x + y\theta$ with $x, y \in \mathbb{F}_q$ and (A.2) becomes

$$2(x^2 - y^2\theta^2)^2 = 2(x^3 + 3xy^2\theta^2).$$

The left side is homogeneous of degree 4 and the right side is homogeneous of degree 3. This implies that this equation defines a curve is of genus 0, and a parametrization can be built using the fact that each line $x = ty$ intersects the

curve at $(0,0)$ and at most one other point. Letting $x = tf(t)$ and $y = f(t)$ we obtain the equation

$$2f(t)^4(t^2-\theta^2)^2 = 2f(t)^3t(t^2+3\theta^2),$$

which has solutions

$$f(t) = t\frac{(t^2+3\theta^2)}{(t^2-\theta^2)^2} \quad \text{and} \quad f(t) = 0.$$

Therefore

$$a = t\frac{(t^2+3\theta^2)}{(t^2-\theta^2)^2}(t+\theta)$$

(with a related formula for b) is a parametrization of the curve.

The map $t \mapsto a\bar{a}$ is a rational function over $\mathbb{F}_q$ of degree 6, therefore $a\bar{a}$ does not generate $\mathbb{F}_q$ for at most $\sum_{k|e,k<e} 6p^k$ values of t. Moreover, a^3 (or, equivalently, a^{2q-1}) is in $\mathbb{F}_q$ for at most 5 choices t, because

$$a^3 = t^3\frac{(t^2+3\theta^2)^3}{(t^2-\theta^2)^6}\left[(t^2+3\theta^2)+\theta(3t^2+\theta^2)\right],$$

and the equation $t^3(t^2+3\theta^2)^3(3t^2+\theta^2) = 0$ has at most 5 distinct solutions. Hence, the number of t such that $a^3 \notin \mathbb{F}_q$ and $a\bar{a}$ generates $\mathbb{F}_q$ is at least

$$q - 5 - 6\sum_{k|e,k<e} p^k \geq q - 5 - 6(e-1)q^{1/2}.$$

Hence there exist at least 4 such a whenever $q \neq 3,5,9,25$. Since $a\bar{a} = a^{2q-1}\bar{a}^{2q-1}$ generates $\mathbb{F}_q$ over $\mathbb{F}_p$, it follows that a^{2q-1} generates F over $\mathbb{F}_p$. Direct verification shows that (A.1) has a solution with the desired properties for $q = 9$ and $q = 25$.

If q is even then (A.2) is a union of $(3, q+1)$ lines: $\bar{a} = a$, together with $\bar{a} = \omega a$ and $\bar{a} = \bar{\omega}a$ if some $\omega \notin \mathbb{F}_q$ satisfies $\omega^3 = 1$. If a is any generator of $\mathbb{F}_q^*$ then the pair $(a, 1-a)$ is solution to (A.1) such that $\mathbb{F}_q = \mathbb{F}_p[a\bar{a}] = \mathbb{F}_p[a^{2q-1}]$. $\square$

Our original approach to Lemma 4.23 was very different. First we observed that if $(3, q+1) = 1$ then (4.24) is equivalent to

$$A + B = 1, \quad A^{q+1} + B^{q+1} = 1$$

and hence, after substitution of $B = 1 - A$ in the second equation, to

$$A + \bar{A} = 2A\bar{A},$$

and this has q solutions in F, because it defines a curve of genus 0 over $\mathbb{F}_q$ (if we view $A = x + y\theta$ with $x, y \in \mathbb{F}_q$). An easy counting argument shows that unless q is even, there is always a solution which generates F over $\mathbb{F}_p$.

If $3|q+1$ then $F = \mathbb{F}_q[\omega]$, where $\omega^2+\omega+1 = 0$. This allows us to write $a = x+y\omega$, $b = u + v\omega$ where $x, y, u, v \in \mathbb{F}_q$. Substituting in (4.24), we obtain a system of 3 equations in 4 variables over $\mathbb{F}_q$, of degrees 3, 3 and 2. Since the coefficients in these equations turn out to be integers these equations define a curve C (or several curves) over $\mathbb{Q}$. Using the bounds on the degrees of the equations one can obtain an upper bound for the genus g of the curve. It is known that the number of points of C over $\mathbb{F}_q$ is at least $q - 2g\sqrt{q}$, which for large enough q is more than the number of elements in $\mathbb{F}_q$ which do not generate this field.

We started studying the above system in order to obtain a nice bound for the genus g and to find how large q has to be for the above argument to work. We

expected to get a curve of genus 9 or 10 but after doing computations we were surprised to find that actually we had 3 curves of genus 0, which allowed us to parametrize all solutions of the system. (N.B. If we change one of the constants on the right in (4.24) to a generic element in $\mathbb{F}_p$, we indeed obtain a curve of genus 9.)

Appendix B. Suzuki triples

Proof of Lemma 4.29. Every vector in $\mathbb{Z}^4$ that cannot be reduced using (a) and (b) must have one of the following 25 types for some integers $a, b \geq 0$. (For, by Example 4.11, the first and third coordinates must be one of $(-a, 0)$, $(0, -a)$, (a, a), $(a, a+1)$, $(a+1, a)$, and similarly for the second and fourth.)

Case 1	Case 2	Case 3	Case 4	Case 5
$(-a,-b,0,0)$	$(-a,0,0,-b)$	$(a,-b,a,0)$	$(0,b,-a,b)$	(a,b,a,b)
		$(a,-b,a+1,0)$	$(0,b+1,-a,b)$	$(a,b,a+1,b)$
$(0,0,-a,-b)$	$(0,-b,-a,0)$	$(a+1,-b,a,0)$	$(0,b,-a,b+1)$	$(a+1,b,a,b)$
				$(a,b,a,b+1)$
		$(a,0,a,-b)$	$(-a,b,0,b)$	$(a,b,a+1,b+1)$
		$(a+1,0,a,-b)$	$(-a,b,0,b+1)$	$(a+1,b,a,b+1)$
		$(a,0,a+1,-b)$	$(-a,b+1,0,b)$	$(a,b+1,a,b)$
				$(a,b+1,a+1,b)$
				$(a+1,b+1,a,b)$

Since $\mathcal{M}_{\mathrm{Sz}}$ is invariant under the permutation $(1,3)(2,4)$ of coordinates, there are only 5 cases that need to be considered on this list: the ones on the second row above, in the last 3 columns decorated with some $+1$'s as seen on the remainder of the last three columns. That is, we only need to consider the five cases $v = (-a,-b,0,0), (-a,0,0,-b), (a+\alpha, b+\beta, a+\gamma, b+\delta)$, $(a+\alpha, -b, a+\gamma, 0)$, $(0, b+\beta, -a, b+\delta)$ with $a, b \geq 0$ and $\alpha, \beta, \gamma, \delta \in \{0, 1\}$. We will present the second and third of these cases, leaving the remaining high school algebra to the reader.

Case 2: Let $v = (-a, 0, 0, -b)$ with $a, b \geq 0$. Triple (c) has three translates containing v:

1	$(-a,0,0,-b)$	$(1-a,-1,0,-b)$	$(-a,0,1,-1-b)$
2	$(-1-a,1,0,-b)$	$(-a,0,0,-b)$	$(-1-a,1,-1-b)$
3	$(-a,0,-1,1-b)$	$(1-a,-1,-1,1-b)$	$(-a,0,0,-b)$

Triple (d) has three translates containing v:

4	$(-a,0,0,-b)$	$(-2-a,1,0,-b)$	$(-a,0,-2,1-b)$
5	$(2-a,-1,0,-b)$	$(-a,0,0,-b)$	$(2-a,-1,-2,1-b)$
6	$(-a,0,2,-1-b)$	$(-2-a,1,2,-1-b)$	$(-a,0,0,-b)$

Here v is (c)-reducible via 1 iff
$a^2 + 2b^2 > (1-a)^2 + 2 + 2b^2$ and
$a^2 + 2b^2 > a^2 + 1 + 2(1-b)^2$,
hence iff $2a > 2$ and $0 > 2 + 4b$, which never occurs.

Also v is (c)-reducible via 2 iff
$a^2 + 2b^2 > (-1-a)^2 + 2 + 2b^2$ and
$a^2 + 2b^2 > (-1-a)^2 + 2 + 1 + 2(-1-b)^2$,
hence iff $0 > 2a + 3$ and $0 > 2a + 4b + 6$, which never occurs.

Also v is (c)-reducible via 3 iff

$a^2+2b^2 > a^2+1+2(1-b)^2$ and
$a^2+2b^2 > (1-a)^2+2+1+2(1-b)^2$,
hence iff $4b>3$ and $2a+4b>6$.

Thus, *the only way v can be* (c)*-irreducible is to have* $4b \le 3$ *or* $2a+4b \le 4$. *Hence* $b \le 1$.

Next, x is (d)-reducible via 4 iff
$a^2+2b^2 > (-2-a)^2+2+2b^2$ and
$a^2+2b^2 > a^2+4+2(1-b)^2$,
hence iff $0>4a+6$ and $4b>6$, which never occurs.

Also v is (d)-reducible via 5 iff
$a^2+2b^2 > (2-a)^2+2+2b^2$ and
$a^2+2b^2 > (2-a)^2+2+4+2(1-b)^2$,
hence iff $4a>6$ and $4a+4b>12$.

Also v is (d)-reducible via 6 iff
$a^2+2b^2 > a^2+4+2(-1-b)^2$ and
$a^2+2b^2 > (-2-a)^2+2+4+2(-1-b)^2$,
hence iff $0>4b+6$ and $0>4a+4b+12$, which never occurs.

Thus, *the only way v can be* (d)*-irreducible is to have* $4a \le 6$ *or* $4a+4b \le 12$. *Hence* $a \le 3$.

Consequently, all irreducible v of the specified sort have $a \le 3$, $b \le 1$ and hence $|v|^2 \le 3^2+2$.

Case 3: The 9 types in the last column have the form $v=(a+\alpha, b+\beta, a+\gamma, b+\delta)$ with $a,b \ge 0$ and $\alpha,\beta,\gamma,\delta \in \{0,1\}$. Triple (c) has three translates containing v:

1	$(a+\alpha,b+\beta,a+\gamma,b+\delta)$	$(1+a+\alpha,-1+b+\beta,a+\gamma,b+\delta)$	$(a+\alpha,b+\beta,1+a+\gamma,-1+b+\delta)$
2	$(-1+a+\alpha,1+b+\beta,a+\gamma,b+\delta)$	$(a+\alpha,b+\beta,a+\gamma,b+\delta)$	$(-1+a+\alpha,1+b+\beta,1+a+\gamma,-1+b+\delta)$
3	$(a+\alpha,b+\beta,-1+a+\gamma,1+b+\delta)$	$(1+a+\alpha,-1+b+\beta,-1+a+\gamma,1+b+\delta)$	$(a+\alpha,b+\beta,a+\gamma,b+\delta)$

Triple (d) has three translates containing v:

4	$(a+\alpha,b+\beta,a+\gamma,b+\delta)$	$(-2+a+\alpha,1+b+\beta,a+\gamma,b+\delta)$	$(a+\alpha,b+\beta,-2+a+\gamma,1+b+\delta)$
5	$(2+a+\alpha,-1+b+\beta,a+\gamma,b+\delta)$	$(a+\alpha,b+\beta,a+\gamma,b+\delta)$	$(2+a+\alpha,-1+b+\beta,-2+a+\gamma,1+b+\delta)$
6	$(a+\alpha,b+\beta,2+a+\gamma,-1+b+\delta)$	$(-2+a+\alpha,1+b+\beta,2+a+\gamma,-1+b+\delta)$	$(a+\alpha,b+\beta,a+\gamma,b+\delta)$

Here v is (c)-reducible via 1 iff

$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (1+a+\alpha)^2+2(-1+b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 \quad \text{and}$$
$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (a+\alpha)^2+2(b+\beta)^2+(1+a+\gamma)^2+2(-1+b+\delta)^2,$$

hence iff $4(b+\beta) > 2(a+\alpha)+3$ and $4(b+\delta) > 2(a+\gamma)+3$.

Also v is (c)-reducible via 2 iff

$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (-1+a+\alpha)^2+2(1+b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 \quad \text{and}$$
$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (-1+a+\alpha)^2+2(1+b+\beta)^2+(1+a+\gamma)^2+2(-1+b+\delta)^2,$$

hence iff $2(a+\alpha) > 4(b+\beta)+3$ and $0 > 2\gamma-2\alpha+4\beta-4\delta+6$, which never occurs.

Also v is (c)-reducible via 3 iff

$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (a+\alpha)^2+2(b+\beta)^2+(-1+a+\gamma)^2+2(1+b+\delta)^2 \quad \text{and}$$
$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (1+a+\alpha)^2+2(-1+b+\beta)^2+(-1+a+\gamma)^2+2(1+b+\delta)^2,$$

hence iff $2(a+\gamma) > 4(b+\beta)+3$ and $0 > 2\alpha - 2\gamma + 4\delta - 4\beta + 6$, which never occurs.

Thus, *the only way v can be* (c)*-irreducible is to have* $4(b+\beta) \le 2(a+\alpha)+3$ *or* $4(b+\delta) \le 2(a+\gamma)+3$, *and hence to have* $4b \le 2a+5$.

Next, v is (d)-reducible via 4 iff

$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (-2+a+\alpha)^2+2(1+b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 \quad \text{and}$$
$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (a+\alpha)^2+2(b+\beta)^2+(-2+a+\gamma)^2+2(1+b+\delta)^2,$$

hence iff $4(a+\alpha) > 4(b+\beta)+6$ and $4(a+\gamma) > 4(b+\delta)+6$.

Also v is (d)-reducible via 5 iff

$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (2+a+\alpha)^2+2(-1+b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 \quad \text{and}$$
$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (2+a+\alpha)^2+2(-1+b+\beta)^2+(-2+a+\gamma)^2+2(1+b+\delta)^2,$$

hence iff $4(b+\beta) > 4(a+\alpha)+6$ and $0 > 4\alpha - 4\gamma - 4\beta + 4\delta + 12$, which never occurs.

Also v is (d)-reducible via 6 iff

$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (a+\alpha)^2+2(b+\beta)^2+(2+a+\gamma)^2+2(-1+b+\delta)^2 \quad \text{and}$$
$$(a+\alpha)^2+2(b+\beta)^2+(a+\gamma)^2+2(b+\delta)^2 > (-2+a+\alpha)^2+2(1+b+\beta)^2+(2+a+\gamma)^2+2(-1+b+\delta)^2,$$

hence iff $4(b+\delta) > 4(a+\gamma)+6$ and $0 > -4\alpha + 4\gamma + 4\beta - 4\delta + 12$, but the latter never occurs.

Thus, *the only way v can be* (d)*-irreducible is to have* $4(a+\alpha) \le 4(b+\beta)+6$ *and* $4(a+\gamma) \le 4(b+\delta)+6$; *and hence to have* $4a \le 4b+10$.

Consequently, *all irreducible v of the specified sort have* $4b \le 2a+5$ *and* $4a \le 4b+10$, *and hence satisfy* $0 \le a \le 7, 0 \le b \le 4$ *and* $|v|^2 \le 2(8^2+2\cdot 5^2)$.

We have now bounded the length of each irreducible vector. The bounds are in the following table.

Case	$a \le$	$b \le$	$\lvert v\rvert^2 \le$
1	9	6	153
2	3	1	11
3	1	2	9
4	1	2	26
5	7	4	228

Instead of considering all cases even more tediously than above, we used a computer calculation to obtain all irreducible vectors. □

For completeness we list all 139 irreducible vectors in Table B.1.

A more conceptual way to prove Lemma 4.29 is following. The set $\mathcal{M}_{\mathcal{U}}$ is the product of two copies of $\mathcal{M}_{\mathrm{BCRW}}$. Therefore any irreducible vector for $\mathcal{M}_{\mathcal{U}}$ lies close to (at distance at most 1) one of 9 quarter planes A_i (each quarter plane arises as a Cartesian product of 2 rays). These quarter planes are listed in the cases in the beginning of this appendix

The system $\mathcal{M}_{-\mathcal{V}}$ also can be transformed to the product of two copies of $\mathcal{M}_{\mathrm{BCRW}}$. This implies that the irreducible vectors for $\mathcal{M}_{-\mathcal{V}}$ also are close to one of 9 quarter planes B_j. A relatively easy computation shows that the intersection of any two quarter planes $A_i \cap B_j$ consists only of the origin, which implies that the set of irreducible vectors is bounded.

In order to significantly decrease the size of the set of irreducible vectors we will enlarge $\mathcal{M}_{\mathrm{Sz}}$. This can be done using the following refinement of Section 4.2:

TABLE B.1. The irreducible vectors for $\mathcal{M}_{\mathrm{Sz}}$

(−9,−6,0,0)	(−8,−5,0,0)	(−7,−5,0,0)	(−7,−4,0,0)	(−6,−4,0,0)	(−6,−3,0,0)
(−5,−4,0,0)	(−5,−3,0,0)	(−5,−2,0,0)	(−4,−3,0,0)	(−4,−2,0,0)	(−4,−1,0,0)
(−3,−3,0,0)	(−3,−2,0,0)	(−3,−1,0,0)	(−3,0,0,0)	(−2,−2,0,0)	(−2,−1,0,0)
(−2,0,0,0)	(−2,1,0,0)	(−1,−2,0,0)	(−1,−1,0,0)	(−1,0,0,−1)	(−1,0,0,0)
(−1,0,0,1)	(−1,1,0,0)	(0,−2,1,0)	(0,−1,−1,0)	(0,−1,0,0)	(0,−1,1,0)
(0,0,−9,−6)	(0,0,−8,−5)	(0,0,−7,−5)	(0,0,−7,−4)	(0,0,−6,−4)	(0,0,−6,−3)
(0,0,−5,−4)	(0,0,−5,−3)	(0,0,−5,−2)	(0,0,−4,−3)	(0,0,−4,−2)	(0,0,−4,−1)
(0,0,−3,−3)	(0,0,−3,−2)	(0,0,−3,−1)	(0,0,−3,0)	(0,0,−2,−2)	(0,0,−2,−1)
(0,0,−2,0)	(0,0,−2,1)	(0,0,−1,−2)	(0,0,−1,−1)	(0,0,−1,0)	(0,0,−1,1)
(0,0,0,−1)	(0,0,0,0)	(0,0,0,1)	(0,0,1,−1)	(0,0,1,0)	(0,0,1,1)
(0,1,−1,0)	(0,1,0,0)	(0,1,1,0)	(0,1,1,1)	(0,2,1,1)	(1,−1,0,0)
(1,−1,1,0)	(1,0,0,−2)	(1,0,0,−1)	(1,0,0,0)	(1,0,0,1)	(1,0,1,−1)
(1,0,1,0)	(1,0,1,1)	(1,0,2,−1)	(1,0,2,0)	(1,0,2,1)	(1,1,0,0)
(1,1,0,1)	(1,1,0,2)	(1,1,1,0)	(1,1,1,1)	(1,1,1,2)	(1,1,2,0)
(1,1,2,1)	(1,1,2,2)	(1,2,1,1)	(1,2,2,1)	(2,−1,1,0)	(2,0,1,0)
(2,0,1,1)	(2,0,2,1)	(2,1,1,0)	(2,1,1,1)	(2,1,1,2)	(2,1,2,0)
(2,1,2,1)	(2,1,2,2)	(2,1,3,0)	(2,1,3,1)	(2,1,3,2)	(2,2,1,1)
(2,2,2,1)	(2,2,3,1)	(2,2,3,2)	(2,3,3,2)	(3,0,2,1)	(3,1,2,1)
(3,1,2,2)	(3,1,3,2)	(3,2,2,1)	(3,2,2,2)	(3,2,2,3)	(3,2,3,1)
(3,2,3,2)	(3,2,3,3)	(3,2,4,1)	(3,2,4,2)	(3,2,4,3)	(3,3,3,2)
(3,3,4,2)	(4,1,3,2)	(4,2,3,2)	(4,2,3,3)	(4,2,4,3)	(4,3,3,2)
(4,3,4,2)	(4,3,5,2)	(4,3,5,3)	(4,4,5,3)	(5,2,4,3)	(5,3,4,3)
(5,3,4,4)	(5,3,5,4)	(5,4,5,3)	(5,4,6,3)	(6,3,5,4)	(6,5,7,4)
(7,4,6,5)					

Lemma B.1. *Let $\mathcal{M}$ be as in* Definition 4.9. *Let $M_1, M_2 \in \mathcal{M}$ and $v \in \mathbb{Z}^k$ be such that $M_1 \cap (v + M_2)$ is a single point. Let $M = (M_1 \setminus (v + M_2)) \cup ((v + M_2) \setminus M_1)$ denote the symmetric difference of M_1 and $v + M_2$. Then any set $S \subseteq \mathbb{Z}^k$ that is $\mathcal{M}$-closed is also $\mathcal{M} \cup \{M\}$-closed.*

Proof. Assume that S is $\mathcal{M}$-closed. For any u such that $|S \cap (u+M)| \geq |M| - 1$ we have $|S \cap (u + M_1)| \geq |M_1| - 1$ or $|S \cap (u + v + M_2)| \geq |M_2| - 1$; we assume the first of these. Since S is $\mathcal{M}$-closed we have $u + M_1 \subseteq S$; in particular the intersection $(u+M_1) \cap (u+v+M_2)$ is in S. Since $|(u+M_1) \cap (u+v+M_2)| = |M_1 \cap (v+M_2)| = 1$ by hypothesis, it follows that $|S \cap (u + v + M_2)| \geq |M_2| - 1$. Since S is $\mathcal{M}$-closed, we see that $u + v + M_2 \subseteq S$ and hence that $u + M \subseteq S$, which shows that S is also closed under $\{M\}$. □

This lemma is natural from a group-theoretic point of view. For, in our uses of Definition 4.9, for each $M \in \mathcal{M}$ there is a relation of the form $\prod_{m \in M} (u^m)^{\pm 1} = 1$. In the preceding lemma we have two such relations $\prod_{m \in M_1} (u^m)^{\pm 1} = 1 = \prod_{m \in M_2} (u^{v+m})^{\pm 1}$. By eliminating the one common term in these products we obtain a third relation $\prod_{m \in M} (u^m)^{\pm 1} = 1$.

Let $\mathcal{M}_{22}$ be the system consisting of the 22 sets in the Table B.2. Each of the sets S_5–S_{22} in that table is the symmetric difference of the two sets in the last two columns, and the preceding lemma can be applied.

TABLE B.2. The set $\mathcal{M}_{22}$

S_1	{(0,0,0,0),(1,0,0,0),(0,0,1,0)}		
S_2	{(0,0,0,0),(0,1,0,0),(0,0,0,1)}		
S_3	{(0,0,0,0),(1,−1,0,0),(0,0,1,−1)}		
S_4	{(0,0,0,0),(−2,1,0,0),(0,0,−2,1)}		
S_5	{(0,0,0,0),(0,0,1,0),(0,1,0,0),(0,1,1,−1)}	S_1	(0,1,0,0)+S_3
S_6	{(0,0,0,0),(1,0,0,0),(0,0,0,1),(1,−1,0,1)}	S_1	(0,0,0,1)+S_3
S_7	{(1,0,0,0),(0,0,1,0),(2,−1,0,0),(2,−1,−2,1)}	S_1	(2,−1,0,0)+S_4
S_8	{(1,0,0,0),(0,0,1,0),(0,0,2,−1),(−2,1,2,−1)}	S_1	(0,0,2,−1)+S_4
S_9	{(0,0,0,0),(0,0,1,0),(1,0,2,−1),(−1,1,2,−1)}	S_1	(1,0,2,−1)+S_4
S_{10}	{(0,0,0,0),(1,0,0,0),(2,−1,1,0),(2,−1,−1,1)}	S_1	(2,−1,1,0)+S_4
S_{11}	{(0,0,0,0),(0,0,0,1),(0,1,−1,1),(1,0,−1,1)}	S_2	(0,1,−1,1)+S_3
S_{12}	{(0,0,0,0),(0,1,0,0),(−1,1,0,1),(−1,1,1,0)}	S_2	(−1,1,0,1)+S_3
S_{13}	{(0,0,0,0),(0,0,1,−1),(1,−1,2,−1),(−1,0,2,−1)}	S_3	(1,−1,2,−1)+S_4
S_{14}	{(0,0,0,0),(1,−1,0,0),(2,−1,1,−1),(2,−1,−1,0)}	S_3	(2,−1,1,−1)+S_4
S_{15}	{(0,0,0,0),(0,0,0,1),(1,0,−2,1),(1,0,−1,1),(2,0,0,0)}	S_2	(1,0,−2,1)+S_9
S_{16}	{(0,0,0,0),(0,1,0,0),(−2,1,1,0),(−1,1,1,0),(0,0,2,0)}	S_2	(−2,1,1,0)+S_{10}
S_{17}	{(0,0,0,0),(0,1,0,0),(1,−1,0,0),(1,0,0,0),(0,0,1,0)}	S_2	(1,−1,0,0)+S_{12}
S_{18}	{(1,0,0,0),(0,0,1,0),(0,0,2,−1), (−1,0,2,−1),(0,0,3,−2),(0,0,1,−1)}	S_8	(−2,1,2,−1)+S_{14}
S_{19}	{(1,0,0,0),(0,0,1,0),(2,−1,0,0), (2,−1,−1,0),(3,−2,0,0),(1,−1,0,0)}	S_7	(2,−1,−2,1)+S_{13}
S_{20}	{(0,0,0,0),(0,0,1,0),(1,−1,0,0), (−1,0,1,0),(0,0,1,0),(1,−1,2,0)}	S_1	(1,−1,0,0)+S_{16}
S_{21}	{(0,0,1,0),(1,0,2,−1),(−1,1,2,−1), (0,0,1,−1),(1,−1,2,−1),(−1,0,2,−1)}	S_9	S_{13}
S_{22}	{(1,0,0,0),(2,−1,1,0),(2,−1,−1,1), (1,−1,0,0),(2,−1,1,−1),(2,−1,−1,0)}	S_{10}	S_{14}

TABLE B.3. The irreducible vectors for $\mathcal{M}_{22}$

(−2,−1,0,0)	(−1,−1,0,0)	(−1,0,0,0)	(0,−1,0,0)	(0,−1,1,0)	(0,0,−2,−1)	(0,0,−1,−1)
(0,0,−1,0)	(0,0,−1,1)	(0,0,0,−1)	(0,0,0,0)	(0,0,0,1)	(0,0,1,0)	(0,0,1,1)
(0,1,0,0)	(0,1,1,0)	(1,0,0,−1)	(1,0,0,0)	(1,0,0,1)	(1,0,1,0)	(1,0,1,1)
(1,1,0,0)	(1,1,1,0)	(1,1,1,1)	(1,1,2,0)	(1,1,2,1)	(2,0,1,1)	(2,1,1,1)
(2,1,2,1)	(2,2,3,1)	(3,2,3,2)	(−2,−2,0,0)	(0,0,−2,−2)		

By Lemma B.1, any set which is closed under $\mathcal{M}_{\mathrm{Sz}}$ is also closed under $\mathcal{M}_{22}$. Consequently, the set of irreducible vectors for $\mathcal{M}_{22}$ is a subset of the set of irreducible vectors for $\mathcal{M}_{\mathrm{Sz}}$. A computer calculation gives that this system has only the 33 irreducible vectors listed in Table B.3 (this involves checking that 106 vectors in Table B.1 reduce under the new tuples in $\mathcal{M}_{22}$).

By slightly perturbing the length function we will further shrink the set of irreducible vectors without introducing any new irreducible vectors. With respect to the length given by $|(x_1, x_2, x_3, x_4)|^2 = 1.01x_1^2 + 2x_2^2 + 1.011x_3^2 + 2.002x_4^4$ there are only 16 irreducible vectors for $\mathcal{M}_{22}$, listed in Table B.4. (Even without leaving the original collection $\mathcal{M}_{\mathrm{Sz}}$, a simpler calculation shows that this new length decreases

TABLE B.4. The set S_0 of irreducible vectors for $\mathcal{M}_{22}$, with respect to the perturbed length

$(-1,-1,0,0)$	$(-1,0,0,0)$	$(0,-1,0,0)$	$(0,0,-1,-1)$	$(0,0,-1,0)$
$(0,0,0,-1)$	$(0,0,0,0)$	$(0,1,0,0)$	$(1,0,0,0)$	$(1,0,1,0)$
$(1,1,0,0)$	$(1,1,1,0)$	$(1,1,1,1)$	$(2,1,1,1)$	
$(-2,-2,0,0)$	$(-2,-1,0,0)$			

the number of irreducible vectors from 139 to 69.) Therefore, using this length we have:

Proposition B.2. *The only $\mathcal{M}_{\mathrm{Sz}}$-closed subset of $\mathbb{Z}^4$ which contains the set S_0 of* 16 *vectors in* Table B.4 *is $\mathbb{Z}^4$.*

APPENDIX C. REE TRIPLES AND QUADRUPLES

We will sketch the ideas involved in the proof of Proposition 4.33. We use the notation for U and h_ζ introduced before that proposition. Let V denote the subgroup $x(0,*,*)$ of U, and let W be the subgroup $x(0,0,*)$. As in (4.28), if $(i,j,k,l,m,n) \in \mathbb{Z}^6$ and $x \in U$ we write

$$x^{(i,j,k,l,m,n)} := x^h \text{ with } h = h_\zeta^i h_{\zeta^\theta}^j h_{\zeta+1}^k h_{\zeta^\theta+1}^l h_{\zeta+2}^m h_{\zeta^\theta+2}^n. \tag{C.1}$$

The following trivial identities in F

$$\begin{array}{ll} (\zeta)+1=(\zeta+1) & (\zeta)=1+(\zeta+2) \\ (\zeta+1)+1=(\zeta+2) & (\zeta^\theta)+1=(\zeta^\theta+1) \\ (\zeta^\theta)=1+(\zeta^\theta+2) & (\zeta^\theta+1)+1=(\zeta^\theta+2) \end{array}$$

become relations in U/V by (4.27):

$$\begin{array}{ll} u^{(0,0,0,0,0,0)}u^{(1,0,0,0,0,0)} \equiv u^{(0,0,1,0,0,0)} & u^{(1,0,0,0,0,0)} \equiv u^{(0,0,0,0,0,0)}u^{(0,0,0,0,1,0)} \\ u^{(0,0,0,0,0,0)}u^{(0,0,1,0,0,0)} \equiv u^{(0,0,0,0,1,0)} & u^{(0,0,0,0,0,0)}u^{(0,1,0,0,0,0)} \equiv u^{(0,0,0,1,0,0)} \\ u^{(0,1,0,0,0,0)} \equiv u^{(0,0,0,0,0,0)}u^{(0,0,0,0,0,1)} & u^{(0,0,0,0,0,0)}u^{(0,0,0,1,0,0)} \equiv u^{(0,0,0,0,0,1)} \end{array}$$

for any $u \in U$. These relations correspond to the following collection $\mathcal{M}_{\mathcal{U}}$ of triples from $\mathbb{Z}^6$:

$$\begin{array}{ll} \{(0,0,0,0,0,0),(1,0,0,0,0,0),(0,0,1,0,0,0)\} & \{(0,0,0,0,0,0),(1,0,0,0,0,0),(0,0,0,0,1,0)\} \\ \{(0,0,0,0,0,0),(0,0,1,0,0,0),(0,0,0,0,1,0)\} & \{(0,0,0,0,0,0),(0,1,0,0,0,0),(0,0,0,1,0,0)\} \\ \{(0,0,0,0,0,0),(0,1,0,0,0,0),(0,0,0,0,0,1)\} & \{(0,0,0,0,0,0),(0,0,0,1,0,0),(0,0,0,0,0,1)\}. \end{array}$$

Since $\theta^2=3$, we have

$$\begin{aligned} \left(\zeta^\theta\right)^{\theta-1}\zeta^{\theta-1} &= \zeta^{(\theta+1)(\theta-1)} = \zeta^{\theta^2-1} = \zeta^2 \\ \left(\zeta^\theta\right)^{\theta-1}\left(\zeta^{\theta-1}\right)^3 &= \zeta^{(\theta-1)(\theta+3)} = \zeta^{\theta^2+2\theta-3} = \left(\zeta^\theta\right)^2. \end{aligned}$$

The relations $1+\zeta^2+(\zeta+1)^2+(\zeta+2)^2=0$ and $\zeta^2(\zeta+1)^2+(\zeta+1)^2(\zeta+2)+(\zeta+2)^2\zeta^2=1$ imply

$$1+(\zeta)^{\theta+1}(\zeta^\theta)^{\theta+1}+(\zeta+1)^{\theta+1}(\zeta^\theta+1)^{\theta+1}+(\zeta+2)^{\theta+1}(\zeta^\theta+2)^{\theta+1}=0,$$

$$\begin{aligned} (\zeta)^{\theta+1}(\zeta^\theta)^{\theta+1}(\zeta+1)^{\theta+1}(\zeta^\theta+1)^{\theta+1}+(\zeta+1)^{\theta+1}(\zeta^\theta+1)^{\theta+1}(\zeta+2)^{\theta+1}(\zeta^\theta+2)^{\theta+1} \\ +\ (\zeta+2)^{\theta+1}(\zeta^\theta+2)^{\theta+1}(\zeta)^{\theta+1}(\zeta^\theta)^{\theta+1}=1 \end{aligned}$$

and two further identities obtained by applying θ, which by (4.32) imply relations in V/W:

$$\begin{aligned}
&v^{(0,0,0,0,0,0,0)}v^{(1,1,0,0,0,0)}v^{(0,0,1,1,0,0)}v^{(0,0,0,0,1,1)} \equiv 1\\
&v^{(0,0,0,0,0,0,0)} \equiv v^{(1,1,1,1,0,0)}v^{(0,0,1,1,1,1)}v^{(1,1,0,0,1,1)}\\
&v^{(0,0,0,0,0,0,0)}v^{(3,1,0,0,0,0)}v^{(0,0,3,1,0,0)}v^{(0,0,0,0,3,1)} \equiv 1\\
&v^{(0,0,0,0,0,0,0)} \equiv v^{(3,1,3,1,0,0)}v^{(0,0,3,1,3,1)}v^{(3,1,0,0,3,1)}
\end{aligned}$$

for any $v \in V$. These relations correspond to the following collection $\mathcal{M}_{\mathcal{V}}$ of quadruples from $\mathbb{Z}^6$:

$$\begin{aligned}
&\{(0,0,0,0,0,0),(1,1,0,0,0,0),(0,0,1,1,0,0),(0,0,0,0,1,1)\}\\
&\{(0,0,0,0,0,0),(1,1,1,1,0,0),(0,0,1,1,1,1),(1,1,0,0,1,1)\}\\
&\{(0,0,0,0,0,0),(3,1,0,0,0,0),(0,0,3,1,0,0),(0,0,0,0,3,1)\}\\
&\{(0,0,0,0,0,0),(3,1,3,1,0,0),(0,0,3,1,3,1),(3,1,0,0,3,1)\}.
\end{aligned}$$

Similarly we have

$$\left(\zeta^{\theta}\right)^{2-\theta}\left(\zeta^{2-\theta}\right)^{2} = \zeta^{(2-\theta)(\theta+2)} = \zeta^{4-\theta^2} = \zeta$$

and

$$\begin{aligned}
(\zeta^2)^{2-\theta}(\zeta^{\theta})^{2-\theta} + 1 &= ((\zeta+1)^2)^{2-\theta}(\zeta^{\theta}+1)^{2-\theta}\\
(\zeta^2)^{2-\theta}(\zeta^{\theta})^{2-\theta} &= 1 + ((\zeta+2)^2)^{2-\theta}(\zeta^{\theta}+2)^{2-\theta}\\
((\zeta+1)^2)^{2-\theta}(\zeta^{\theta}+1)^{2-\theta} + 1 &= ((\zeta+1)^2)^{2-\theta}(\zeta^{\theta}+1)^{2-\theta},
\end{aligned}$$

together with additional similar identities; by (4.32) these imply the following relations in W:

$$\begin{aligned}
&w^{(0,0,0,0,0,0,0)}w^{(2,1,0,0,0,0)} = w^{(0,0,2,1,0,0)} \qquad && w^{(2,1,0,0,0,0)} = w^{(0,0,0,0,0,0,0)}w^{(0,0,0,0,2,1)}\\
&w^{(0,0,0,0,0,0,0)}w^{(0,0,2,1,0,0)} = w^{(0,0,0,0,2,1)} && w^{(0,0,0,0,0,0,0)}w^{(3,2,0,0,0,0)} = w^{(0,0,3,2,0,0)}\\
&w^{(3,2,0,0,0,0)} = w^{(0,0,0,0,0,0,0)}w^{(0,0,0,0,3,2)} && w^{(0,0,0,0,0,0,0)}w^{(0,0,3,2,0,0)} = w^{(0,0,0,0,3,2)}
\end{aligned}$$

for any $w \in W$. These relations correspond to the following collection $\mathcal{M}_{\mathcal{W}}$ of triples from $\mathbb{Z}^6$:

$$\begin{aligned}
&\{(0,0,0,0,0,0),(2,1,0,0,0,0),(0,0,2,1,0,0)\} \quad \{(0,0,0,0,0,0),(2,1,0,0,0,0),(0,0,0,0,2,1)\}\\
&\{(0,0,0,0,0,0),(0,0,2,1,0,0),(0,0,0,0,2,1)\} \quad \{(0,0,0,0,0,0),(3,2,0,0,0,0),(0,0,3,2,0,0)\}\\
&\{(0,0,0,0,0,0),(3,2,0,0,0,0),(0,0,0,0,3,2)\} \quad \{(0,0,0,0,0,0),(0,0,3,2,0,0),(0,0,0,0,3,2)\}.
\end{aligned}$$

We will use the fact that all six systems

$$\mathcal{M}_{\mathcal{U}} \cup \mathcal{M}_{-\mathcal{U}},\ \mathcal{M}_{\mathcal{U}} \cup \mathcal{M}_{-\mathcal{V}},\ \mathcal{M}_{\mathcal{U}} \cup \mathcal{M}_{-\mathcal{W}},\ \mathcal{M}_{\mathcal{V}} \cup \mathcal{M}_{-\mathcal{V}},\ \mathcal{M}_{\mathcal{V}} \cup \mathcal{M}_{-\mathcal{W}},\ \mathcal{M}_{\mathcal{W}} \cup \mathcal{M}_{-\mathcal{W}}$$

are of finite type with respect to suitable length functions in order to show that $[U,U] \equiv 1 \pmod V$, $[U,V] \equiv 1 \pmod W$, $[U,W] = 1$, $[V,V] \equiv 1 \pmod W$, $[V,W] = 1$ and $[W,W] = 1$, respectively. The first system contains several copies of $\mathcal{M}_{\pm\mathrm{BCRW}}$, therefore it is of finite type. The last system can be transformed to the first one by a linear transformation, therefore is also of finite type for the transformed length function. The next lemma gives the same result for the remaining four systems; we sketch a proof in Appendix C. As in Section 4.3.3, we can avoid using the first and last sets because we can use Lemma 4.1 to show that the group $\langle u^{\langle h_\zeta\rangle}\rangle$ is abelian (mod V), and then impose additional relations to make it invariant under all $h_\star$.

Lemma C.2. *Each of the systems $\mathcal{M}_{\mathrm{Ree}_1} := \mathcal{M}_{\mathcal{U}} \cup \mathcal{M}_{-\mathcal{V}}$, $\mathcal{M}_{\mathrm{Ree}_2} := \mathcal{M}_{\mathcal{U}} \cup \mathcal{M}_{-\mathcal{W}}$, $\mathcal{M}_{\mathrm{Ree}_3} := \mathcal{M}_{\mathcal{V}} \cup \mathcal{M}_{-\mathcal{V}}$ and $\mathcal{M}_{\mathrm{Ree}_4} := \mathcal{M}_{\mathcal{W}} \cup \mathcal{M}_{-\mathcal{V}}$ in $\mathbb{Z}^6$ is of finite type using suitable length functions for the various systems.*

Proof. As in the case of Lemma 4.29, again the proof is long and tedious. We use two 3-dimensional analogues $\mathcal{M}_{3\mathrm{BCRW}}$ and $\mathcal{M}_{3\mathrm{BCRW}'}$ of BCRW: $\mathcal{M}_{3\mathrm{BCRW}}$ is

$$\{(0,0,0),(1,0,0),(0,1,0)\}$$
$$\{(0,0,0),(1,0,0),(0,0,1)\}$$
$$\{(0,0,0),(0,1,0),(0,0,1)\},$$

consisting of 3 copies of $\mathcal{M}_{\mathrm{BCRW}}$; and $\mathcal{M}_{3\mathrm{BCRW}'}$ is

$$\{(0,0,0),(1,0,0),(0,1,0),(0,0,1)\}$$
$$\{(0,0,0),(1,1,0),(0,1,1),(1,0,1)\}.$$

We use the Euclidean length defined by $|(x,y,z)|^2 = x^2+y^2+z^2$.

Lemma C.3. (1) *The irreducible vectors for* $\mathcal{M}_{3\mathrm{BCRW}}$ *are as follows* (*up to permutation of the coordinates*):

$$(-a,0,0),(a,a,a),(a,a,a+1),(a,a,a-1)$$

for integers $a \geq 0$.

(2) *The irreducible vectors for* $\mathcal{M}_{3\mathrm{BCRW}'}$ *are as follows* (*up to permutation of the coordinates*):

$$(-a,0,0),(-a,0,-1),(-a,0,1),(a,a,a),(a,a,a+1),(a,a,a-1),(a,a+1,a-1)$$

for integers $a \geq 0$.

Proof. (1) $\mathcal{M}_{3\mathrm{BCRW}}$ contains one copy of $\mathcal{M}_{\mathrm{BCRW}}$ for each pair of coordinates, therefore the projection of an irreducible vector to any coordinate plane is an irreducible vector for $\mathcal{M}_{\mathrm{BCRW}}$, i.e., is one of $(-a,0),(a,a),(a,a+1)$ for some nonnegative integer a. It is easy to see that the vectors with this property are exactly the ones in the statement of the lemma.

(2) Here we have to consider several different cases depending on the signs of the coordinates. We will do just one case and leave the remainder to the reader.

Let (x,y,z) be an irreducible vector with $x \geq y \geq z \geq 0$. This vector reduces via $(x-1,y,z)$, (x,y,z), $(x-1,y+1,z)$, $(x-1,y+1,z)$ unless $x < y+2$. Similarly it reduces via $(x-1,y-1,z)$, (x,y,z), $(x-1,y,z+1)$, $(x,y-1,z+1)$ unless $y < z+2$. Therefore the only irreducible vectors of this form are the ones with

$$x = y = z \text{ or } x = y = z+1 \text{ or } x = y+1 = z+1 \text{ or } x = y+1 = z+2.$$

The other three cases $x \geq y \geq 0 > z$; $x \geq 0 > y \geq z$ and $0 > x \geq y \geq z$ are similar. □

Proof of Lemma C.2. As in the preceding argument, $\mathcal{M}_{\mathcal{U}}$ can be transformed by a linear transformation to a product of two copies of $\mathcal{M}_{3\mathrm{BCRW}}$, and $\mathcal{M}_{\mathcal{V}}$ can be transformed to $\mathcal{M}_{3\mathrm{BCRW}'}$. We choose the length function so that our transformation turns it into the Euclidean length function.

In order to find all irreducible vectors for the system $\mathcal{M}_X \cup \mathcal{M}_{-Y}$, where $X, Y \in \{\mathcal{U},\mathcal{V},\mathcal{W}\}$, we can proceed as follows: First we apply a linear transformation and pick a length function such that $\mathcal{M}_X$ becomes a product of two copies of $\mathcal{M}_{3\mathrm{BCRW}}$ or $\mathcal{M}_{3\mathrm{BCRW}'}$ with the Euclidean norm. Any vector which is irreducible for $\mathcal{M}_X$ lies close (a bounded distance from) one of 16 quarter planes A_i, so we only have to check which of these vectors are irreducible for $\mathcal{M}_{-Y}$. After doing the computations in each case one obtains a bound for the length of any irreducible vector, which proves that the system is of finite type. Unfortunately, the computations are very long and tedious and (as in Appendix B) do not provide any insight.

Equivalently one can argue that a vector which is irreducible for $\mathcal{M}_{-Y}$ lies close to one of 16 quarter planes B_j, and check that the intersection $A_i \cap B_j$ is trivial in all cases. □

We are now ready to provide the presentation needed in Proposition 4.33.

Generators: $u, v, w, h_\star$ for $\star \in \{\zeta, \zeta+1, \zeta+2, \zeta^\theta, \zeta^\theta+1, \zeta^\theta+2\}$.
Define $x^{(i,j,k,l,m,n)}$ as in (C.1).

Relations:

(1) $[h_\star, h_\bullet] = 1$ for $\star \in \{\zeta, \zeta+1, \zeta+2, \zeta^\theta, \zeta^\theta+1, \zeta^\theta+2\}$.
(2) $w^3 = 1$.
(3) $w^{(0,0,1,0,0,0)} = ww^{(1,0,0,0,0,0)}$.
(4) $ww^{(0,0,0,0,1,0)} = w^{(1,0,0,0,0,0)}$.
(5) $[w, w^{(1,0,0,0,0,0)}] = 1$.
(6) $[[w^{m_\zeta(x)}]]_{h_\zeta} = 1$.
(7) $w^{h_{\zeta^\theta}} = [[w^{x^{3^{k+1}}}]]_{h_\zeta}$, where $x^{3^{k+1}}$ is reduced mod $m_\zeta(x)$ in order to obtain a short relation.
(8) $w^{(0,0,0,1,0,0)} = ww^{(0,1,0,0,0,0)}$.
(9) $ww^{(0,0,0,0,1,0)} = w^{(0,1,0,0,0,0)}$.
(10) $v^3 = w_1$.
(11) $vv^{(1,1,0,0,0,0)}v^{(0,0,1,1,0,0)}v^{(0,0,0,0,1,1)} = w_2$.
(12) $vv^{(3,1,0,0,0,0)}v^{(0,0,3,1,0,0)}v^{(0,0,0,0,3,1)} = w_3$.
(13) $v^{(1,1,1,1,0,0)}v^{(0,0,1,1,1,1)}v^{(1,1,0,0,1,1)} = vw_4$.
(14) $v^{(3,1,3,1,0,0)}v^{(0,0,3,1,3,1)}v^{(3,1,0,0,3,1)} = vw_5$.
(15) $[w^{S_1}, v] = 1$, where S_1 is the set of irreducible vectors for $\mathcal{M}_\mathcal{W} \cup \mathcal{M}_{-\mathcal{V}}$.
(16) $[v^{S_2}, v] = w_6$, where S_2 is the set of irreducible vectors for $\mathcal{M}_\mathcal{V} \cup \mathcal{M}_{-\mathcal{V}}$.
(17) $[[v^{m_{\zeta^2}(x)}]]_{h_\zeta h_{\zeta^\theta}} = w_7$.
(18) $v^{h_\star} = [[v^{f_{\star^\theta-2,\zeta^2}(x)}]]_{h_\zeta h_{\zeta^\theta}} w_\star$ for $\star \in \{\zeta, \zeta+1, \zeta+2, \zeta^\theta, \zeta^\theta+1, \zeta^\theta+2\}$.
(19) $u^3 = v_1$.
(20) $u^{(0,0,2,1,0,0)} = uu^{(2,1,0,0,0,0)}v_2$.
(21) $uu^{(0,0,0,0,2,1)} = u^{(2,1,0,0,0,0)}v_3$.
(22) $u^{(0,0,3,2,0,0)} = uu^{(3,2,0,0,0,0)}v_4$.
(23) $uu^{(0,0,0,0,3,2)} = u^{(3,2,0,0,0,0)}v_5$.
(24) $[u, u^{(2,1,0,0,0,0)}] = v_6$.
(25) $[u^{S_3}, w] = 1$, where S_3 is the set of irreducible vectors for $\mathcal{M}_\mathcal{U} \cup \mathcal{M}_{-\mathcal{W}}$.
(26) $[u^{S_4}, v] = w_9$, where S_4 is the set of irreducible vectors for $\mathcal{M}_\mathcal{U} \cup \mathcal{M}_{-\mathcal{V}}$.
(27) $[[u^{m_\zeta(x)}]]_{h_\zeta^2 h_{\zeta^\theta}} = v_8$.
(28) $u^{h_\star} = [[u^{f_{\star^\theta-2,\zeta}(x)}]]_{h_{\zeta^2} h_{\zeta^\theta}} v_\star$ for $\star \in \{\zeta, \zeta+1, \zeta+2, \zeta^\theta, \zeta^\theta+1, \zeta^\theta+2\}$.

Here, $w_1, \ldots$ and $v_1, \ldots$ are suitable elements of $W := \langle w^{\langle \text{all } h_\star \rangle} \rangle$ and $V := \langle v^{\langle \text{all } h_\star \rangle}, W \rangle$, respectively. These depend on our initial choices ζ, u, v, w as well as on the length functions used to determine the sets S_i.

The proof that the above is a presentation for an infinite central extension of B is similar to the Suzuki case. We omit the details.

There are $46 + \sum_i |S_i|$ relations in this presentation, which would probably be somewhat unmanageable in practice. □

References

[AG] M. Aschbacher and R. Guralnick, Some applications of the first cohomology group. J. Algebra 90 (1984) 446–460.

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

[BKL] L. Babai, W. M. Kantor and A. Lubotzky, Small diameter Cayley graphs for finite simple groups. Eur. J. Comb. 10 (1989) 507–522.

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

[BM] H. Behr and J. Mennicke, A presentation of the groups $\mathrm{PSL}(2,p)$. Canad. J. Math. 20 (1968) 1432–1438.

[BS] C. D. Bennett and S. Shpectorov, A new proof of Phan's theorem. J. Group Theory 7 (2004) 287–310.

[BCLO] J. Bray, M. D. E. Conder, C. R. Leedham-Green and E. A. O'Brien, Short presentations for alternating and symmetric groups (preprint).

[Bur] W. Burnside, Theory of Groups of Finite Order, 2nd ed. Cambridge Univ. Press, Cambridge 1911.

[CCHR] C. M. Campbell, P. P. Campbell, B. T. K. Hopson and E. F. Robertson, On the efficiency of direct powers of $\mathrm{PGL}(2,p)$, pp. 27–34 in: Recent advances in group theory and low-dimensional topology (Pusan, 2000; Eds. J. Rae Cho and J. Mennicke). Heldermann, Lemgo 2003.

[CHLR] C. M. Campbell, G. Havas, S. Linton and E. F. Robertson, Symmetric presentations and orthogonal groups, pp. 1–10 in: The atlas of finite groups: ten years on (Birmingham, 1995; Eds. R. Curtis and R. Wilson), Lond. Math. Soc. Lecture Note 249. Cambridge Univ. Press, Cambridge 1998.

[CHRR] C. M. Campbell, G. Havas, C. Ramsay and E. F. Robertson, Nice efficient presentations for all small simple groups and their covers. Lond. Math. Soc. J. Comput. Math. 7 (2004) 266–283.

[CR1] C. M. Campbell and E. F. Robertson, Classes of groups related to $F^{a,b,c}$. Proc. Roy. Soc. Edinburgh A78 (1977/78) 209–218.

[CR2] C. M. Campbell and E. F. Robertson, A deficiency zero presentation for $\mathrm{SL}(2,\ p)$. Bull. Lond. Math. Soc. 12 (1980) 17–20.

[CRW1] C. M. Campbell, E. F. Robertson and P. D. Williams, On presentations of $\mathrm{PSL}(2,p^n)$. J. Austral. Math. Soc. 48 (1990) 333–346.

[CRW2] C. M. Campbell, E. F. Robertson and P. D. Williams, Efficient presentations for finite simple groups and related groups, pp. 65–72 in: Groups–Korea 1988 (Pusan, 1988; Eds. Y. Gheel Baik et al.). Springer, Berlin 1989.

[Carm] R. D. Carmichael, Introduction to the theory of groups of finite order. Ginn, Boston 1937.

[Cox] H. S. M. Coxeter, Abstract groups of the form $V_1^k = V_j^3 = (V_iV_j)^2 = 1$. J. Lond. Math. Soc. 9 (1934) 213–219.

[CoMo] H. S. M. Coxeter and W. O. J. Moser, Generators and relations for discrete groups, 3rd ed. Ergebnisse der Mathematik und ihrer Grenzgebiete B. 14. Springer, New York-Heidelberg 1972.

[Cur] C. W. Curtis, Central extensions of groups of Lie type. J. reine angew. Math. 220 (1965) 174–185.

[Do] D. Z. Djokovic, Presentations of some finite simple groups. J. Austral. Math. Soc. 45 (1988) 143–168.

[GLS] D. Gorenstein, R. Lyons and R. Solomon, The classification of the finite simple groups. Number 3. Part I. Chapter A. Almost simple K-groups. Amer. Math. Soc., Providence 1998.

[GHNS] R. Gramlich, C. Hoffman, W. Nickel and S. Shpectorov, Even-dimensional orthogonal groups as amalgams of unitary groups. J. Algebra 284 (2005) 141–173.

[GHS] R. Gramlich, C. Hoffman and S. Shpectorov, A Phan-type theorem for $\mathrm{Sp}(2n,q)$. J. Algebra 264 (2003) 358–384.

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

[Gru] K. W. Gruenberg, Cohomological topics in group theory. Lecture Notes in Mathematics 143. Springer, Berlin-New York 1970

[GKKL1] R. M. Guralnick, W. M. Kantor, M. Kassabov and A. Lubotzky, Presentations of finite simple groups: a cohomological and profinite approach. Groups, Geometry and Dynamics (to appear).

[GKKL2] R. M. Guralnick, W. M. Kantor, M. Kassabov and A. Lubotzky, Presentations of finite simple groups: a computational approach (preprint).

[Ho] D. F. Holt, On the second cohomology group of a finite group. Proc. Lond. Math. Soc. (3) 55 (1987) 22–36.

[HEO] D. F. Holt, B. Eick and E. A. O'Brien, Handbook of computational group theory. Chapman & Hall, Boca Raton 2005.

[HS] A. Hulpke and Á. Seress, Short presentations for three-dimensional unitary groups. J. Algebra 245 (2001) 719–729.

[Ka] W. M. Kantor, Some topics in asymptotic group theory, pp. 403–421 in: Groups, Combinatorics and Geometry (Durham, 1990; Eds. M. W. Liebeck and J. Saxl), Lond. Math. Soc. Lecture Note 165. Cambridge Univ. Press, Cambridge 1992.

[KS1] W. M. Kantor and Á. Seress, Black box classical groups. Memoirs Amer. Math. Soc. 708 (2001).

[KS2] W. M. Kantor and Á. Seress, Computing with matrix groups, pp. 123–137 in: Groups, combinatorics and geometry (Durham, 2001; Eds. A. A. Ivanov et al.). World Sci. Publ., River Edge, NJ 2003.

[KLM] G. Kemper, F. Lübeck and K. Magaard, Matrix generators for the Ree groups ${}^2G_2(q)$. Comm. Alg. 29 407–413 (2001).

[KoLu] I. Korchagina and A. Lubotzky, On presentations and second cohomology of some finite simple groups. Publ. Math. Debrecen 69 (2006) 341–352.

[KM] S. Krstic and J. McCool, Presenting $\mathrm{GL}_n(k\langle T\rangle)$. J. Pure Appl. Algebra 141 (1999) 175–183.

[LG] C. R. Leedham-Green, The computational matrix group project, pp. 229–247 in: Groups and Computation III (Eds. W. M. Kantor and Á. Seress). deGruyter, Berlin-New York 2001.

[Lub1] A. Lubotzky, Enumerating boundedly generated finite groups. J. Algebra 238 (2001) 194–199.

[Lub2] A. Lubotzky, Pro-finite presentations. J. Algebra 242 (2001) 672–690.

[Lub3] A. Lubotzky, Finite presentations of adelic groups, the congruence kernel and cohomology of finite simple groups. Pure Appl. Math. Q. 1 (2005) 241–256.

[LS] A. Lubotzky and D. Segal, Subgroup growth. Birkhäuser, Basel 2003.

[MKS] W. Magnus, A. Karrass and D. Solitar, Combinatorial group theory; presentations of groups in terms of generators and relations. Interscience, New York 1966.

[Man] A. Mann, Enumerating finite groups and their defining relations. J. Group Theory 1 (1998) 59–64.

[Mar] G. A. Margulis, Discrete subgroups of semisimple Lie groups. Springer, Berlin 1991.

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

[Mo] E. H. Moore, 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. Lond. Math. Soc. 28 (1897) 357–366.

[Ph] K. W. Phan, On groups generated by three-dimensional special unitary groups I. J. Austral. Math. Soc. Ser. A23 (1977) 67–77.

[Py] L. Pyber, Enumerating finite groups of given order. Ann. of Math. 137 (1993) 203–220.

[Sch] I. Schur, Untersuchungen über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen. J. reine angew. Math. 132 (1907) 85–137.

[Ser] Á. Seress, Permutation group algorithms. Cambridge Univ. Press, Cambridge 2003.

[Ser1] J.-P. Serre, Le problème des groupes de congruence pour SL2. Ann. of Math. 92 (1970) 489–527.

[Ser2] J.-P. Serre, Trees. Springer, Berlin 2003.

[Shi] K. Shinoda, The conjugacy classes of the finite Ree groups of type F_4. J. Fac. Sci. Univ. Tokyo 22 (1975) 1-15.

[Sim] C. C. Sims, Computation with finitely presented groups. Cambridge Univ. Press, Cambridge 1994.
[St1] R. Steinberg, Lectures on Chevalley groups (mimeographed notes). Yale Univ. 1967.
[St2] R. Steinberg. Generators, relations and coverings of algebraic groups, II. J. Algebra 71 (1981) 527–543.
[Sun] J. G. Sunday, Presentations of the groups SL(2, m) and PSL(2, m). Canad. J. Math. 24 (1972) 1129–1131.
[Suz] M. Suzuki, On a class of doubly transitive groups. Ann. of Math. 75 (1962) 105–145.
[Ti1] J. Tits, Les groupes de Lie exceptionnels et leur interprétation géométrique. Bull. Soc. Math. Belg. 8 (1956) 48–81.
[Ti2] J. Tits, Buildings of spherical type and finite BN-pairs. Springer, Berlin-New York 1974.
[Wi] J. S. Wilson, Finite axiomatization of finite soluble groups. JLMS 74 (2006) 566-582.

Department of Mathematics, University of Southern California, Los Angeles, CA 90089-2532, USA
E-mail address: guralnic@usc.edu

Department of Mathematics, University of Oregon, Eugene, OR 97403, USA
E-mail address: kantor@math.uoregon.edu

Department of Mathematics, Cornell University, Ithaca, NY 14853-4201, USA
E-mail address: kassabov@math.cornell.edu

Department of Mathematics, Hebrew University, Givat Ram, Jerusalem 91904, Israel
E-mail address: alexlub@math.huji.ac.il