Short presentations for classical groups

M.D.E. Conder and C.R. Leedham-Green and E.A. O'Brien

Abstract

We prove that the finite classical groups of rank r defined over fields of size q have presentations of length $O(\log r \log\log r + \log q)$. We exhibit such presentations explicitly for three of the infinite families of classical groups. We also prove that both the special linear and symplectic groups have presentations where the number of relations is bounded, independent of the rank and field. This represents major improvements to existing results.

1 Introduction

The *length* of a presentation may be defined the number of symbols required to write down the presentation.

Perhaps the best known presentations for the classical groups are those of Steinberg [12, 13]. We describe these in more detail in Section 3. The resulting presentation for $\mathrm{SL}(n, q)$ has $O(n^2 q)$ generators, $O(n^4 q^2)$ relations, and its length is polynomial in q.

As shown in [5], subsets of these suffice to define a presentation for a classical group. Other refinements were developed by Babai *et al.* [2]. In particular, they prove that, with the possible exception of the rank 1 twisted groups of Lie type, every finite simple group G has a presentation of length $O(\log^2 |G|)$.

Two of the three exceptional cases have been resolved: Hulpke & Seress [7] exhibit a presentation of this length for the unitary groups $\mathrm{PSU}(3, q)$; a similar presentation for ${}^2B_2(2^{2m+1})$ appears in Suzuki's original paper [15]; no such presentation is currently known for ${}^2G_2(3^{2m+1})$.

Our principal result is the following.

Theorem 1.1 *The finite classical groups of Lie rank r defined over finite fields of size q have presentations of length $O(\log r \log\log r + \log q)$.*

We prove this theorem by exhibiting explicitly such presentations for three of the four families of classical groups. Similar results can be obtained for the orthogonal groups.

This work was supported in part by the Marsden Fund of New Zealand via grant UOA 0412. 2000 *Mathematics Subject Classification.* Primary 20C20.

Critical to this work is our recent construction [6] of presentations having length $O(\log n \log\log n)$ for the alternating and symmetric groups of degree n. We use these to obtain short presentations for the Weyl groups for the classical groups and so produce short presentations for the classical groups.

In the same paper [6], we prove that the symmetric and alternating groups have presentations with two generators and a bounded number of relations. Campbell, Robertson and Williams [4] prove that $\mathrm{PSL}(2, q)$ has a presentation on a fixed number of relations independent of q. We combine these results to obtain the following.

Theorem 1.2 *The classical groups have presentations where the number of relations is bounded by an absolute constant.*

Again we exhibit these presentations explicitly.

Some of the relations in the presentations we construct involve terms of the form x^m, where the modulus of the integer m is potentially large. What contribution should such a term be considered to make to the length of the presentation? A case can be made out for taking this contribution to be 1, or m, or $1 + \lceil \log_2 m \rceil$. We choose the last of these, this being a measure the number of characters required to write this term.

Presentations can generally be reduced in length by adding new generators defined in terms of the given generators (or recursively in terms of earlier defined generators). As Babai *et al.* [2, Remark 1.3] comment, if a group has a presentation $\mathcal{P}$, then it has a presentation with no exponents and length asymptotically at most 4 times that of $\mathcal{P}$.

The content of the paper is the following. In Section 2 we motivate this work, which has important algorithmic applications. In Section 3 we introduce the Steinberg presentations for the classical groups. In Sections 4 and 5 we produce explicit presentations for each of $\mathrm{Sp}(d, q)$ and $\mathrm{SL}(d, q)$ which satisfy Theorem 1.1.

2 Background and motivation

Part of the motivation for this work comes from potential algorithmic applications. A major focus of research activity over the past 15 years has been the development of effective algorithms for the study of subgroups of $\mathrm{GL}(d, \mathrm{GF}(q))$. A particular goal is to construct the composition factors of the linear group. We refer the reader to [10] for background and concepts related to this work.

Many of the algorithms developed are randomised, relying on random selections of elements. A *Monte Carlo* algorithm is a randomised algorithm which may return an incorrect answer to a decision question, and the probability of this event is less than some specified value. A *Las Vegas* algorithm is one which never returns an incorrect answer, but may report failure with probability less than some specified value.

A *constructive recognition algorithm* for a group G of Lie type constructs an explicit isomorphism between G and a "standard" (or natural) representation H of G and produces a new generating set $\mathcal{S}$ for G. By definition, $\mathcal{S}$ has the property that an element of G can be readily written as a word in $\mathcal{S}$. Kantor & Seress [8] provide such constructive recognition algorithms for the classical groups. Others are under development: see, for example, [9].

Babai *et al.* [1] present a Monte Carlo algorithm to determine the name of a non-abelian composition factor of a quasi-simple group G of Lie type in known defining characteristic p. If we believe, by applying this algorithm or otherwise, that G is a group of Lie type, then we can apply a constructive recognition algorithm to G and construct $\mathcal{S}$. If there exists a presentation $\mathcal{P}$ for G on $\mathcal{S}$, then we can decide whether or not G satisfies $\mathcal{P}$. If so, we can confirm that G is indeed as claimed.

If the presentation is "short", then this proposal is practical. Our desire for explicit evaluation in this manner motivates our decision to write down explicit presentations – and not just obtain estimates of their length. In particular, the chosen generating sets for certain of the classical groups may be constructed explicitly using the algorithms of [9].

3 The Steinberg presentation

The foundation of our work is the following theorem of Steinberg [12].

Theorem 3.1 *Let G be a universal Chevalley group of rank greater than 1 over the field* $\mathrm{GF}(q)$ *where $q = p^e$, and let Φ be the set of roots of the underlying Lie algebra. Then G has a presentation as follows. The generators are symbols $x_\alpha(s)$, for $\alpha \in \Phi$ and $s \in \mathrm{GF}(q)$, and the relations are*

$$\begin{aligned} x_\alpha(s+t) &= x_\alpha(s)x_\alpha(t) \\ [x_\alpha(s), x_\beta(t)] &= \prod_{a,b>0} x_{a\alpha+b\beta}(C_{ab\alpha\beta}s^a t^b) \end{aligned}$$

Here α and β range over all roots, subject to the condition $\alpha + \beta \neq 0$, and s and t range over $\mathrm{GF}(q)$. In the second relation, a and b range over all positive integers for which $a\alpha + b\beta$ is a root. This set of pairs (a, b) is very small for fixed α and β. It has cardinality at most four, and is frequently empty, giving the identity for the right hand side of the corresponding relation. Since the terms commute for the classical groups, their order does not matter. For $\mathrm{Sp}(2n, q)$, the set of pairs has cardinality at most two, and the coefficients $C_{ab\alpha\beta}$ are integers in the range $[-2, 2]$; for other Chevalley groups the range may extend to $[-3, 3]$.

The difficulty with defining these coefficients explicitly is that they depend on the parametrisation of the root groups, and there is no agreed uniform way of carrying out this parametrisation. If the parametrisation is changed in a way that changes the sign of the field elements, this changes the signs of the exponents, as $x_\alpha(-s) = x_\alpha(s)^{-1}$. This explains, for example, the difference between the presentation in Theorem 3.1 and that given by Carter [**?**].

For each of $\mathrm{Sp}(d, q)$ and $\mathrm{SL}(d, q)$ we first identify explicitly its Steinberg presentation and then construct a new shorter presentation for this group.

4 The symplectic group

Let $G = \mathrm{Sp}(2n, q)$ act on the vector space V of dimension $2n$ over $\mathrm{GF}(q)$. Then G preserves a non-degenerate symplectic form. We choose a hyperbolic basis $(e_1, e_2, \ldots, e_{2n})$ for V, where $e_{2u-1}.e_{2u} = -e_{2u}.e_{2u-1} = 1$ for $1 \leq u \leq n$, and all other pairs of basis vectors are mutually orthogonal. Thus the pairs (e_{2u-1}, e_{2u}) form hyperbolic bases of mutually orthogonal hyperbolic planes. We assume throughout that $n > 1$.

There is a natural maximal split torus for G, namely the group of diagonal elements of the form $e_{2u-1} \mapsto \alpha_u e_{2u-1}$, $e_{2u} \mapsto \alpha_u^{-1} e_{2u}$ for $1 \leq u \leq n$ and $\alpha_u \in \mathrm{GF}(q)^\times$.

We want to describe the root groups of G. To this end it is convenient to define $e_{-i} = -e_i$ for $i \in [1, 2n]$.

For i and j signed integers, with distinct moduli in the range $[1, 2n]$, and $t \in \mathrm{GF}(q)$, define $\overline{\tau}_{ij}(t)$ by $e_i \mapsto e_i + te_j$, and $e_k \mapsto e_k$ for $|k| \neq |i|$. Thus $\overline{\tau}_{ij}(t) = \overline{\tau}_{-i,-j}(t) = \overline{\tau}_{i,-j}(-t)$.

The long roots correspond naturally with the set of ordered pairs $\alpha = (i, j)$ where $\{i, j\} = \{2u - 1, 2u\}$ for some $u \in [1, n]$.

The root group corresponding to $\alpha = (2u - 1, 2u)$ is the set of elements $x_\alpha(t) = \overline{\tau}_{2u-1,2u}(t)$ for $t \in \mathrm{GF}(q)$, and the root group corresponding to $\alpha = (2u, 2u - 1)$ is the set of elements $x_\alpha(t) = \overline{\tau}_{2u,-2u+1}(t)$ for $t \in \mathrm{GF}(q)$. Thus these elements are symplectic transvections.

For i, j, k, ℓ signed integers, with distinct moduli in the range $[1, 2n]$, define $\overline{\sigma}_{ijk\ell}(t)$ by $e_i \mapsto e_i + te_\ell$, and $e_k \mapsto e_k + te_j$, and $e_w \mapsto e_w$ for $|i| \neq |w| \neq |k|$. Thus $\overline{\sigma}_{ijk\ell}(t) = \overline{\sigma}_{k\ell ij}(t)$.

The short roots correspond naturally with the set of ordered 4-tuples $\alpha = (i, j, k, \ell)$, where $\{i, j\} = \{2u - 1, 2u\}$ and $\{k, \ell\} = \{2v - 1, 2v\}$ and $1 \leq u < v \leq n$.

The root group corresponding to such an ordered 4-tuple $\alpha = (i, j, k, \ell)$ is the set of elements $x_\alpha(t) = \overline{\sigma}_{\pm i,j,\pm k,\ell}(t)$ for $t \in \mathrm{GF}(q)$, where the sign attached to i is $+$ if $i > j$, and is $-$ if $j > i$; and similarly the sign attached to k is $+$ if $k > \ell$, and is $-$ if $\ell < k$.

The Weyl group W of G corresponding to the above torus may be defined as the action on the frame consisting of the set of subspaces $\langle e_i \rangle$, for $1 \leq i \leq 2n$, of the subgroup of G that normalises the frame. This group is isomorphic to $C_2 \wr S_n$, and is the group of those permutations of the frame that preserve the system of imprimitivity whose blocks are the pairs $\{\langle e_{2u-1} \rangle, \langle e_{2u} \rangle\}$ for $1 \leq u \leq n$. Observe that G contains a subgroup isomorphic to $C_4 \wr S_n$ for odd p, and to $C_2 \wr S_n$ if $p = 2$, that maps homomorphically onto W, namely the subgroup $\langle Z, U, V \rangle$, defined as follows.

- $e_1 Z = e_2$, $e_2 Z = -e_1 = e_{-1}$, $e_i Z = 1$ for $i > 2$.
- $e_1 U = e_3$, $e_2 U = e_4$, $e_3 U = e_1$, $e_4 U = e_2$, $e_i U = e_i$ for $i > 4$.
- $e_i V = e_{i+2}$ for all i.

Here suffices are interpreted modulo $2n$.

Observe the minus sign in the definition of Z. This is enforced by the condition that G preserves the underlying symplectic form, and reflects the fact that all elements of G have determinant 1.

Note that $\langle Z, U, V\rangle$ acts, with respect to the given basis, as signed permutation matrices, and permutes the linear transformations $\overline{\tau}_{ij}(t)$ and $\overline{\sigma}_{ijk\ell}(t)$ by conjugation by applying the corresponding signed permutations to the suffices.

4.1 A Steinberg presentation for $\mathrm{Sp}(2n, q)$

We exploit Theorem 3.1 to obtain a presentation for $\mathrm{Sp}(2n, q)$ as follows. The generators are symbols $\tau_{ij}(s)$ and $\sigma_{ijk\ell}(s)$, where $\tau_{ij}(s)$ is a symbol for $(i, j) = (2u - 1, 2u)$ or $(i, j) = (-2u, 2u - 1)$, for any $u \in [1, n]$ and $s \in \mathrm{GF}(q)$; and $\sigma_{ijk\ell}(s)$ is a symbol for $(i, j) = (2u, 2u - 1)$ or $(i, j) = (-2u + 1, 2u)$, and $(k, \ell) = (2v, 2v - 1)$ or $(k, \ell) = (-2v + 1, 2v)$, for any $u < v$ in $[1, n]$ and $s \in \mathrm{GF}(q)$. We extend the set of symbols by allowing the following operations on the suffices:

(a) Multiply both suffices of $\tau_{ij}(s)$ by -1.

(b) Multiply the first two suffices of $\sigma_{ijk\ell}(s)$ by -1.

(c) Multiply the last two suffices of $\sigma_{ijk\ell}(s)$ by -1.

(d) Replace $\sigma_{ijk\ell}(s)$ by $\sigma_{k\ell ij}(s)$.

If $p = 2$ we regard two symbols that differ only in the signs of their suffices as being equal.

An effect of allowing these extra symbols is that the set of suffices, and hence the set of symbols, is permuted by $\langle Z, U, V\rangle$, regarded as a group of signed permutations. Note that the operations (a) to (d), when applied to the corresponding matrices $\overline{\tau}_{ij}(s)$ and $\overline{\sigma}_{ijk\ell}(s)$, replace a matrix by the identical matrix in cases (a) and (d), and by the inverse matrix in cases (b) and (c). We are only concerned with $\tau_{ij}(s)$ when (i, j) is in the orbit of $(1, 2)$ under the action of $\langle Z, U, V\rangle$, and with $\sigma_{ijk\ell}(s)$ when (i, j, k, ℓ) is in the orbit of $(2, 1, 4, 3)$. In this case we shall say that the suffices are *suitable*.

The Steinberg relations now assume the following form:

S(i) $\tau_{ij}(s)\tau_{ij}(t) = \tau_{ij}(s + t)$.

S(ii) $\sigma_{ijk\ell}(s)\sigma_{ijk\ell}(t) = \sigma_{ijk\ell}(s + t)$.

S(iii) $[\tau_{ij}(s), \tau_{k\ell}(t)] = 1$.

S(iv) $[\sigma_{ijk\ell}(s), \tau_{-i,j}(t)] = 1$.

S(v) $[\sigma_{ijk\ell}(s), \tau_{ji}(t)] = \sigma_{-j,i,k,\ell}(st)\tau_{-k,\ell}(-s^2t)$.

S(vi) $[\sigma_{ijk\ell}(s), \sigma_{-j,i,k,\ell}(t)] = \tau_{-k,\ell}(s^2t^2)$.

S(vii) $[\sigma_{ijk\ell}(s), \sigma_{k\ell xy}(t)] = 1$.

S(viii) $[\sigma_{ijk\ell}(s), \sigma_{-\ell,k,x,y}(t)] = \sigma_{ijxy}(-st)$.

S(ix) $[\sigma_{ijk\ell}(s), \sigma_{xyab}(t)] = 1$.

S(x) $[\tau_{ij}(s), \sigma_{klxy}(t)] = 1$.

S(xi) $\tau_{ij}(s) = \tau_{-i,-j}(s)$.

S(xii) $\sigma_{ijk\ell}(s) = \sigma_{-i,-j,k,\ell}(-s) = \sigma_{i,j,-k,-\ell}(-s) = \sigma_{k\ell ij}(s)$.

Here we assume that all suffices are suitable, and that suffices with different names have different moduli. Note that all suitable suffices are covered with the exception of the cases when the suffices correspond to two roots whose sum is zero; that is to say, we do not consider $[\tau_{ij}, \tau_{-j,i}]$, or $[\sigma_{ijk\ell}, \sigma_{-j,i,-\ell,k}]$.

4.2 A short presentation for $\mathrm{Sp}(2n, q)$

In general we take as generators for $\mathrm{Sp}(2n, q)$ the set

$$\{\tau = \overline{\tau}_{12}(1), \sigma = \overline{\sigma}_{2143}(1), \delta, Z, U, V\}$$

where δ is defined by $e_1 \mapsto \omega^{-1}e_1$, $e_2 \mapsto \omega e_2$, $e_i \mapsto e_i$ for $i > 2$, and the other generators are as defined earlier in this section.

The qualifier "in general" in the definition of this generating set reflects the fact that δ is omitted if q is prime, and V is omitted if $n = 2$.

To construct our short presentation for $\mathrm{Sp}(2n, q)$, we define $x_\alpha(t)$ to be a specific word in our generators – this definition will not be part of the presentation. We then prove that our relations hold when the generating symbols are mapped as above to the appropriate elements of $\mathrm{Sp}(2n, q)$, that the above Steinberg relations are consequences of our relations, and that our relations enable each of our generators to be expressed as a word in the $x_\alpha(t)$. This will prove the correctness of our presentation.

The generating symbols U and V are required to generate a copy of the symmetric group S_n, and our relations will include defining relations for S_n on these generators. The choice of presentation for S_n does not impact on the rest of the presentation. However our presentation of $\mathrm{Sp}(2n, q)$ is obtained in principle from a presentation for S_n, a presentation of a long and a short root group, following [4], and a bounded number of additional short relations. Thus the presentation for $\mathrm{Sp}(2n, q)$ is dominated by the choice of presentation for S_n, and inherits any good or bad features of that choice.

A presentation for $\mathrm{PSp}(2n, q)$ follows at once from our presentation for $\mathrm{Sp}(2n, q)$.

We shall on occasion simplify notation by equating words in our generating symbols with their natural images in $\mathrm{Sp}(2n, q)$, and in the quotient of the free group by the relations that we impose.

The following result can be established by direct calculation.

Lemma 4.1 *In* $\mathrm{Sp}(2n, q)$ *the following identities hold for all suitable suffices* (i, j) *and* (i, j, k, ℓ)*, and all* $s \in \mathrm{GF}(q)$*.*

$$\sigma_{ijk\ell}(s)^{\delta} = \sigma_{ijk\ell}(s\omega^{m})$$

where m *is defined as follows. If* $|i| = 1$ *and* $|j| = 2$*, or if* $|k| = 1$ *and* $|\ell| = 2$*, then* $m = 1$*. If* $|i| = 2$ *and* $|j| = 1$*, or if* $|k| = 2$ *and* $|\ell| = 1$*, then* $m = -1$*. If* $|x|$ (*and hence also* $|y|$) *is not in* $\{|i|, |j|, |k|, |\ell|\}$ *then* $m = 0$*. Further*

$$\tau_{ij}(s)^{\delta} = \tau_{ij}(s\omega^{m})$$

where m *is defined as follows. If* $|i| = 1$ *and* $|j| = 2$ *then* $m = 2$*. If* $|i| = 2$ *and* $|j| = 1$ *then* $m = -2$*. In all other cases* $m = 0$*.*

We can now describe our presentation for $\mathrm{Sp}(2n, q)$.

Theorem 4.2 *Let* $n > 1$ *and* $q = p^e$ *for some prime* p*. Let* H *be the group defined by* $\{X \mid \mathcal{R}\}$ *where* $X = \{\tau, \sigma, \delta, Z, U, V\}$*, but with* δ *omitted if* $e = 1$*, and* V *omitted if* $n = 2$*, and* $\mathcal{R}$ *is the union of the following sets of relations.*

R1. If $n > 2$ *a set of defining relations for* S_n *on* $\{U, V\}$*, where* U *maps to the transposition* $(1, 2)$*, and* V *to the cycle* $(1, 2, \ldots, n)$*.*

R2. A set of relations that, with R1, express the fact that $\langle U, V, Z\rangle$ (*omit* V *if* $n = 2$) *is isomorphic to* $C_4 \wr S_n$ *if* p *is odd, and to* $C_2 \wr S_n$ *if* $p = 2$*. These are the following:*
$[Z, Z^U] = 1$*. If* p *is odd then* $Z^4 = 1$ *else* $Z^2 = 1$*. If* $n > 2$ *then* $[Z, VU] = [Z, U^V] = 1$*.*

**** Omit* $[Z, VU] = 1$ *if* $n = 3$*. If* $n = 3$ *then* $VU = U^V$ *Applies also to R3****

R3. A set of relations that express the fact that the subgroup of $\langle U, V, Z\rangle$ (*omit* V *if* $n = 2$) *that centralises* e_1 *and* e_2*, and that is isomorphic to* $C_4 \wr S_{n-1}$*, centralises* τ*, and also* δ *if* $e > 1$*. These are the following:*
If $n = 2$ *then* $[\tau, Z^U] = 1$*; if also* $e > 1$ *then* $[\delta, Z^U] = 1$*. If* $n > 2$ *then* $[\tau, Z^V] = [\tau, U^V] = [\tau, VU] = 1$*; if also* $e > 1$ *then* $[\delta, Z^V] = [\delta, U^V] = [\delta, VU] = 1$*.*

R4. $\sigma^{Z^2} = \sigma^{-1}$ *if* $p > 2$*;* $\sigma^U = \sigma$*;* $\sigma^{U^\delta} = \sigma$ *if* $e > 1$*;* $[\sigma, \delta^{V^2}] = 1$ *if* $n > 2$ *and* $e > 1$*.*

R5. If $n > 2$ *set of relations that express the fact that the subgroup of* $\langle U, V, Z\rangle$ *that centralises* e_i *for* $1 \le i \le 4$*, and that is isomorphic to* $C_4 \wr S_{n-2}$*, centralises* σ*. These are the following:*
If $n > 2$ *then* $[\sigma, Z^{V^2}] = 1$*. If* $n > 3$ *then* $[\sigma, U^{V^2}] = [\sigma, VUU^V] = 1$*.*

**** Omit last if* $n = 4$*, as then* $U^{V^2} = VUU^V$ ******

R6. A presentation for $\mathrm{SL}(2, q)$ *on the generating set* $\{\tau, \delta, Z\}$ (*on* $\{\tau, Z\}$ *if* $e = 1$) *as in Section 8.*

R7. If $e > 1$, relations that express the fact that $\langle \delta, \sigma \rangle$ is isomorphic to $\mathrm{GF}(p^e) : \langle a \rangle$, *where a (as δ) acts by multiplication by ω^{-1}, as in Theorem 8.2 or Theorem 8.6.*

R8. Relations that express instances of the Steinberg relations listed above. The instances are labelled to correspond to the above listing. We state the relation as a relation in the presentation, and in parentheses the relation as an instance of a Steinberg relation.

S3. If $e > 1$ then $[\delta, \tau^U] = 1$. $\quad ([\tau_{12}(\omega^2), \tau_{34}(1)] = 1)$

S4. $[\sigma, \tau^Z] = 1$. $\quad ([\sigma_{2143}(1), \tau_{-2,1}(1)] = 1)$

S5. $[\sigma, \tau] = \sigma^Z \tau^{Z^{-1}U}$. $\quad ([\sigma_{2143}(1), \tau_{12}(1)] = \sigma_{-1,2,4,3}(1)\tau_{-4,3}(-1))$

S6. $[\sigma, \sigma^Z] = \tau^{ZU}$. $\quad ([\sigma_{2143}(1), \sigma_{-1,2,4,3}(1)] = \tau_{-4,3}(1))$

S7. $[\sigma, \sigma^V] = 1$. $\quad ([\sigma_{2143}(1), \sigma_{4365}(1)] = 1)$

S8. $[\sigma, \sigma^{ZV}] = (\sigma^{VU})^{-1}$. $\quad ([\sigma_{2143}(1), \sigma_{-3,4,6,5}(1)] = \sigma_{2165}^{-1}(1))$

S9. $[\sigma, \sigma^{V^2}] = 1$. $\quad ([\sigma_{2143}(1), \sigma_{6587}(1)] = 1)$

S10. $[\tau, \sigma^V] = 1$. $\quad ([\tau_{12}(1), \sigma_{4365}(1)] = 1)$

Relations (vii), (viii), and (x) are omitted if $n < 3$; relation (ix) is omitted if $n < 4$.

R9. If $e > 1$ relations that give the structure of $\langle \sigma, \delta, U \rangle$ as $\mathrm{GF}(q)^+ : (C_{q-1} \wr C_2)$*. With R4 and R7, these are $[\delta, \delta^U] = 1$ and $\sigma^{[U, \delta^{-1}]} = \sigma$.*

R10. A relation that expresses U as word in the Steinberg generators:

if p is odd then $U = Z^2 \sigma^{Z^U} \sigma^{-Z} \sigma^{Z^U}$, else $U = \sigma^{Z^U} \sigma^{-Z} \sigma^{Z^U}$.

R11. $[\tau^U, \delta] = 1$.

Then $H = \langle X \mid \mathcal{R} \rangle$ is isomorphic to $\mathrm{Sp}(2n, q)$*. If $p > 2$ a presentation for* $\mathrm{PSp}(2n, q)$ *is obtained by adding the relation $(ZV)^{2n} = 1$ if $n > 2$, and $[Z^2, U] = 1$ if $n = 2$.*

PROOF: We define elements $\tau_{ij}(s)$ and $\sigma_{ijk\ell}(s)$ of H for suitable values of (i, j, k, ℓ) and for $s \in \mathrm{GF}(q)$ as follows.

If $e = 1$ then we may take $s \in [0, p-1]$, and define $\tau_{12}(s) = \tau^s$. If $e > 1$ then let $s = \sum_i c_i \omega^{2i}$ for $c_i \in [0, p-1]$, and define $\tau_{12}(s)$ to be $\prod_i \tau^{c_i \delta^i}$. This is well defined by R6.

If (i, j) are any other suitable suffices define $\tau_{ij}(s) = \tau_{12}(s)^g$, where g is any element of $\langle Z, U, V \rangle$ that takes $(1, 2)$ to (i, j), ignoring signs if $p = 2$. This is well defined by R3.

Now define $\sigma_{2143}(s)$. If $e = 1$ then we may again take $s \in [0, p-1]$, and define $\sigma_{2143}(s) = \sigma^s$. If $e > 1$ then let $s = \sum_i c_i \omega^{-i}$ for $c_i \in [0, p-1]$, and define $\sigma_{2143}(s)$ to be $\prod_i \sigma^{c_i \delta^i}$, which is well defined by R7.

If (i, j, k, ℓ) are any other suitable suffices, define $\sigma_{ijk\ell}(s)$ to be $\sigma(s)^g$, where g is any element of $\langle Z, U, V\rangle$ that takes $(2, 1, 4, 3)$ to (i, j, k, ℓ), ignoring signs if $p = 2$. This is well defined by R3 and R5.

We now check the Steinberg relations.

S(xi) $\tau_{ij}(s) = \tau_{-i,-j}(s)$.

By conjugation this reduces to the case $(i, j) = (1, 2)$, and then follows from the relations $\tau = \tau^{Z^2}$ and $\delta = \delta^{Z^2}$ implied by R6.

S(xii) $\sigma_{ijk\ell}(s) = \sigma_{-i,-j,k,\ell}(-s) = \sigma_{i,j,-k,-\ell}(-s) = \sigma_{k\ell ij}(s)$.

This reduces to the case $(i, j, k, \ell) = (2, 1, 4, 3)$. The first equality then follows from $\sigma^{Z^2} = \sigma^{-1}$ in R4, and $\delta^{Z^2} = \delta$; and $\sigma_{ijk\ell}(s) = \sigma_{k\ell ij}(s)$ follows from $\sigma^U = \sigma$, and $\delta^U = \delta$, as in R3 and R4. The final equality follows from these two.

S(i) $\tau_{ij}(s)\tau_{ij}(t) = \tau_{ij}(s + t)$.

This reduces to the case $(i, j) = (1, 2)$, and then follows from R6.

S(ii) $\sigma_{ijk\ell}(s)\sigma_{ijk\ell}(t) = \sigma_{ijk\ell}(s + t)$.

We reduce this and all subsequent Steinberg relations to the case $(i, j, k, \ell) = (2, 1, 4, 3)$. This relation now follows from R7.

S(iii) $[\tau_{ij}(s), \tau_{k\ell}(t)] = 1$.

It suffices to prove that τ commutes with τ^U, and if $e > 1$ then τ^U and δ^U commute with δ. But τ and δ commute with U by R3, and τ^U commutes with δ by R11.

S(iv)$[\sigma_{ijk\ell}(s), \tau_{-i,j}(t)] = 1$.

The case $s = t = 1$ is S4. If the equation holds for some value of (s, t) then by conjugation by δ it also holds for $(s\omega^{-1}, t\omega^2)$; and by conjugation by δ^U it holds for $(s\omega^{-1}, t)$ since δ^U centralises τ by S3, and acts as δ on σ by R4.

S(v) $[\sigma_{ijk\ell}(s), \tau_{ji}(t)] = \sigma_{-j,i,k,\ell}(st)\tau_{-k,\ell}(-s^2t)$.

This translates to $[\sigma^{\delta^i}, \tau^{\delta^j}] = \sigma^{\delta^{i-2j}Z}\tau\delta^j ZU$. If the relation holds for some value of (s, t) then conjugation by δ implies that it holds for $(s\omega^{-1}, t\omega^2)$, and conjugating by δ^U implies that it holds for $(s\omega^{-1}, t)$. So if it holds for one value of (s, t), as in S5, it holds for all values.

S(vi) $[\sigma_{ijk\ell}(s), \sigma_{-j,i,k,\ell}(t)] = \tau_{-k,\ell}(s^2t^2)$.

If this relation holds for some value of (s, t) then conjugation by δ implies that it holds for $(s\omega^{-1}, t\omega)$, and conjugating by δ^U implies that it holds for $(s\omega^{-1}, t\omega^{-1})$. So it suffices to impose one case of this relation, such as S6. The remaining cases are as follows, and all follow easily from a single example by the same conjugation argument.

S(vii) $[\sigma_{ijk\ell}(s), \sigma_{k\ell xy}(t)] = 1$.

S(viii) $[\sigma_{ijk\ell}(s), \sigma_{-\ell,k,x,y}(t)] = \sigma_{ijxy}(-st)$.

S(ix) $[\sigma_{ijk\ell}(s), \sigma_{xyab}(t)] = 1$.

S(x) $[\tau_{ij}(s), \sigma_{klxy}(t)] = 1$.

This proves that the subgroup K of H generated by σ and τ, and their conjugates under elements of $\langle\delta, Z, U, V\rangle$, is isomorphic to $\mathrm{Sp}(2n, q)$. It remains to prove that δ, Z, U and V lie in K. Since $\mathrm{SL}(2, q)$ is generated by the conjugates of τ under $\langle\delta, Z\rangle$, and hence lies in K, it follows that δ and Z lie in K. Since $\langle U, V\rangle$ is isomorphic to S_n, and acts on H as the corresponding subgroup of the Weyl group, it follows that

H is either isomorphic to $\mathrm{Sp}(2n, q)$ or to $\mathrm{Sp}(2n, q) \times S_n$, where $\langle U, V\rangle$ is diagonally embedded in this group. But, by R10, U lies in H, so $H = K$ as required.

Now $\langle U, V\rangle$ is isomorphic to S_n, and corresponds to a subgroup of K isomorphic to S_n. Hence U and V lie in K. Similarly since $\langle \delta, \tau, Z\rangle$ is isomorphic to $\mathrm{SL}(2, q)$ it follows that δ (if $e > 1$) and Z lie in K. This completes the proof that our presentation of $\mathrm{Sp}(2n, q)$ is correct.

The additional relation required for $\mathrm{PSp}(2n, q)$ is clearly correct. □

5 A short presentation for $\mathrm{SL}(d, q)$

Let $G = \mathrm{SL}(d, q)$ where $q = p^e$, for prime p, and $d > 2$. We take matrices with respect to an ordered basis $(e_1, e_2, \ldots, e_d)$. As in the case of the symplectic groups we start with a Steinberg presentation as follows. We have symbols $\tau_{\pm i,\pm j}(s)$ where $s \in \mathrm{GF}(q)$, and i and j are distinct integers in the range 1 to d.

The Steinberg relations for G are as follows:

S(i) $\tau_{ij}(s)\tau_{ij}(t) = \tau_{ij}(s + t)$.

S(ii) $[\tau_{ij}(s), \tau_{jk}(t)] = \tau_{ik}(st)$.

S(iii) $[\tau_{ij}(s), \tau_{ik}(t)] = 1$.

S(iv) If $d > 3$ then $[\tau_{ij}(s), \tau_{k\ell}(t)] = 1$.

S(v) $\tau_{ij}(s) = \tau_{-i,j}(-s) = \tau_{i,-j}(-s)$.

Here s and t take all values in $\mathrm{GF}(q)$, and i, j, k, ℓ are integers (positive if $p = 2$) with moduli in $[1, d]$, suffices with different names having different moduli.

The generating set $\{\delta, \tau, U, V\}$ for the presentation corresponds to the following elements of G:

- δ corresponds to the diagonal matrix with diagonal entries $(\omega^{-1}, \omega, 1, 1, 1, \ldots, 1)$, where ω is a primitive element of $\mathrm{GF}(q)$. Note that ω^2 is also a generator for $\mathrm{GF}(q)$.
- $\tau = \tau_{12}$ corresponds to the transvection that maps e_1 to $e_1 + e_2$, and fixes the other basis vectors.
- U and V correspond to the following signed permutation matrices. U takes e_1 to e_2, and e_2 to $-e_1$, fixing the other basis vectors. V takes e_i to e_{i+1} for $1 \leq i < d$, and takes e_d to $\epsilon(-1)^{d+1}e_1$.

The generator δ is omitted if $e = 1$.

Note that the above signed matrices generate the group of all signed permutation matrices of determinant 1, this being isomorphic to a subgroup of index 2 in $C_2 \wr S_d$ if p is odd, and to S_d if p is even. We denote this group by $I = I(p, d)$

Theorem 5.1 *Let $d > 1$ and $q = p^e$ for some prime p. Let H be the group defined by $\{X \mid \mathcal{R}\}$ where $X = \{\delta, \tau, U, V\}$, with δ omitted if $e = 1$, and $\mathcal{R}$ is the union of the following sets of relations.*

R1. Relations for $I(p, d)$ on U and V. See Eamonn-Marston section.

R2. Relations expressing the fact that the copy of $I(p, d-2)$ in $I(p, d)$ that fixes e_1 and e_2 centralises τ and δ. These are the following:
$[\tau, VUU^V] = [\tau, U^{V^2}] = [\delta, VUU^V] = [\delta, U^{V^2}] = 1.$

R3. A presentation for $\mathrm{SL}(2, q)$ *on $\{\delta, \tau, U\}$, omitting δ if $e = 1$.*

R4. (a) $[\delta, \delta^V] = [\delta, (U^2)^V] = 1.$

(b) $\tau^{(U^2)^{VU}} = \tau^{-1}.$ $([\tau_{1,-2}(1) = \tau_{12}(-1))$

(c) $[\tau, \tau^V] = \left(\tau^{VU}\right)^{-1}.$ $([\tau_{12}(1), \tau_{23}(1)] = \tau_{-1,3}(-1))$

(d) $[\tau, \tau^{VU}] = 1.$ $([\tau_{12}(1), \tau_{13}(-1)] = 1)$

(e) $[\tau, \tau^{V^2}] = 1$ *if* $d > 3.$ $([\tau_{12}(1), \tau_{34}(1)] = 1)$

(f) $[\tau, \left(\delta^{V^{-1}}\right)^2 \delta] = 1.$

(g) $[\tau, \delta^V \left(\delta^{V^{-1}}\right)^{-1}] = 1.$

(h) $[\tau, \left(\delta^{U^{-V}}\right)^2 \delta^{-1}] = 1.$

(i) $\tau^{\delta^V} = \tau^a(\tau^\delta)^b$ *if* $q = 25$ *and* $\omega^{-1} = a + b\omega^2.$

(j) $[\tau, \tau^{\delta V}] = \tau^{\delta U^V}$ *if* $q = 4.$ $([\tau_{12}(1), \tau_{23}(\omega^2)] = \tau_{13}(\omega^2))$

If $e = 1$ then δ and all relations involving δ are omitted.

Then $H = \langle X \mid \mathcal{R}\rangle$ is isomorphic to $\mathrm{SL}(d, 2^e)$.

Proof: We define elements $\tau_{ij}(s)$ of H for suitable values of (i, j) and $s \in \mathrm{GF}(q)$ as follows.

If $e = 1$ then we may take $s \in [0, p-1]$, and define $\tau_{12}(s) = \tau^s$. If $e > 1$ then let $s = \sum_i c_i w^{2i}$ for $c_i \in [0, p-1]$, and define $\tau_{12}(s)$ to be $\prod_i \tau^{c_i \delta^i}$. This is well defined by R2.

If (i, j) are any other suitable suffices define $\tau_{ij}(s) = \tau_{12}(s)^g$, where g is any element of $\langle U, V\rangle$ that takes $(1, 2)$ to (i, j), ignoring signs if $p = 2$. This is well defined by R2.

We now check the Steinberg relations.

Note that R2 and R3 and R4(a) together imply that δ commutes with its conjugate under any element of $\langle U, V\rangle$, and with the conjugate of U^2 under any element of $\langle U, V\rangle$.

By conjugation we may assume in all cases that $(i, j, k) = (1, 2, 3)$.

S(v) $\tau_{ij}(s) = \tau_{-i,j}(-s) = \tau_{i,-j}(-s)$.

This is vacuous for $p = 2$, so assume $p > 2$. Note that $(U^2)^{VU}$, as a signed permutation, interchanges 1 and -1, and fixes 2, so $\tau_{-1,2}(s) = \tau_{12}(s)^{(U^2)^{VU}}$. Hence the case $s = 1$ is given by R4(b). Also $\tau_{-1,2}(\omega^{2i}) = \tau^{\delta^i (U^2)^{VU}} = \tau^{(U^2)^{VU} \delta^i} = \tau_{12}(\omega^{-2i})$, by R(b). Since the even powers of ω span GF(q) the first part of S(v) follows. The second part then follows provided that $\tau_{ij}(s) = \tau_{-i,-j}(s)$, which is the case, since δ and τ commute with U^2 by R3.

S(i) $\tau_{ij}(s)\tau_{ij}(t) = \tau_{ij}(s+t)$.

This is implied by R3.

S(ii) $[\tau_{ij}(s), \tau_{jk}(t)] = \tau_{ik}(st)$.

$[\tau_{12}(1), \tau_{23}(1)] = \tau_{13}(1)$ is given by R4(c). For general s and t it suffices, as we demonstrate below, to prove the following for $e > 1$, provided that $q \notin \{4, 25\}$.

$\tau_{12}(s)^{\delta^2} = \tau_{12}(s\omega^4), \tau_{23}(s)^{\delta^2} = \tau_{23}(s\omega^{-2}), \tau_{12}(s)^{\delta^{2V}} = \tau_{12}(s\omega^{-2}), \tau_{23}(s)^{\delta^{2V}} = \tau_{23}(s\omega^4)$.

The first and last of these follow directly from the definitions. The second is proved as follows.

$\tau_{23}(\omega^{2i})^{\delta^2} = \tau_{12}(\omega^{2i})^{V\delta^2}$ by definition, and this is $\tau^{\delta^i V \delta^2}$. Now $\delta^i V \delta^2 = \delta^i \delta^{2V^{-1}} V = \delta^{2V^{-1}} \delta^i V$ since δ commutes with δ^V by R4(a). But τ conjugated by $\delta^{2V^{-1}}$ is $\tau^{\delta^{-1}}$ as in R4(f). So $\tau^{\delta^i V \delta} = \tau_{12}(\omega^{-2})^{\delta^i V} = \tau_{23}(\omega^{2i-2})$. Thus $\tau_{23}(\omega^{2i})^{\delta^2} = \tau_{23}(\omega^{2i-2})$, and since the even powers of ω span GF(q) it follows that $\tau_{23}(s)^{\delta^2} = \tau_{23}(s\omega^{-2})$ for all s.

The proof of the third is similar, since $\tau^{\delta^V} = \tau^{\delta^{V^{-1}}}$ by R4(g).

Thus if S(ii) holds for a given value of (s,t) it also holds for $(s\omega^4, t\omega^{-2})$, and for $(s\omega^{-2}, t\omega^4)$, and hence for $(s, t\omega^6)$ and for $(s\omega^6, t)$. But if it holds for (s_1, t) and (s_2, t) it holds for $(s_1 + s_2, t)$, and since $q \notin \{4, 25\}$ it follows that ω^6 generates GF(q), as a field, and so, since S(ii) is true for $(s,t) = (1,1)$ it holds for all (s,t).

If $q = 25$ then GF(q) is generated by ω^3, so the same argument carries through if we prove the following:

$\tau_{12}(s)^{\delta} = \tau_{12}(s\omega^2)$, $\tau_{23}(s)^{\delta} = \tau_{23}(s\omega^{-1})$, $\tau_{12}(s)^{\delta^V} = \tau_{12}(s\omega^{-1})$, $\tau_{23}(s)^{\delta^{1V}} = \tau_{23}(s\omega^2)$.

Thus we need relations to evaluate $\tau^{\delta^{V^{-1}}}$ and τ^{δ^V}. If $\omega^{-1} = a + b\omega^2$ these relations are clearly $\tau^{\delta^{V^{-1}}} = \tau^{\delta^V} = \tau^{a+b\delta}$ as in R4(g) and R4(i).

If $q = 4$ then we prove that $\tau_{12}(s)^{\delta^2} = \tau_{12}(s\omega^4)$, $\tau_{23}(s)^{\delta^2} = \tau_{23}(s\omega^{-2})$ as before and include the relation $[\tau_{12}(1), \tau_{23}(\omega^2)] = \tau_{13}(\omega^2)$.

S(iii) $[\tau_{ij}(s), \tau_{ik}(t)] = 1$.

$[\tau_{12}(1), \tau_{13}(1)] = 1$ is implied by R4(b) and R4(c) . For general s and t it suffices to prove the following for $e > 1$:

$\tau_{12}(s)^{\delta^2} = \tau_{12}(s\omega^4)$, $\tau_{13}(s)^{\delta^2} = \tau_{13}(s\omega^2)$, $\tau_{12}(s)^{\delta^{2V}} = \tau_{12}(s\omega^{-2})$, $\tau_{13}(s)^{\delta^{2V}} = \tau_{13}(s\omega^{-2})$.

The first and third of these equations have been dealt with above. The second is proved as follows.

$\tau_{13}(\omega^{2i})^{\delta^2} = \tau^{\delta^i U^V \delta^2} = \tau^{\delta^i \delta^{(U^{-V})^2} U^V} = \tau^{\delta^{i+1} U^V}$ by Rg$= \tau_{13}(\omega^{2i+2})$. This implies, as above, that $\tau_{13}(s)^{\delta^2} = \tau_{13}(s\omega^2)$ for all $s \in$ GF(q).

Thus if S(iii) holds for a given value of (s,t) it also holds for $(s\omega^4, t\omega^2)$ and for $(s\omega^{-2}, t\omega^{-2})$, and hence for $(s, t\omega^{-2})$ and $(s\omega^2, t)$. Since ω^2 generates GF(q) S(iii) holds in all cases.

S(iv) $[\tau_{ij}(s), \tau_{k\ell}(t)] = 1$.

This obviously follows from the fact that $[\tau_{12}(1), \tau_{34}(1)] = 1$ (R4(e)) by conjugation.

It follows that the normal closure of $\langle \delta, \tau \rangle$ in H is isomorphic to SL(d,q). Also, U

lies in this subgroup by R3. The fact that V also lies in this subgroup now follows as in the case of the symplectic groups, and the theorem is proved.

□

If p is odd then $\langle U, V\rangle$ is a subgroup $\overline{S}_d$ of index 2 in the group $C_2 \wr S_d$ of signed permutation matrices. Let $\{u, v|R\}$ be a presentation for S_d, where u and v stand for $(1, 2)$ and $(1, \ldots, d)$ respectively. The presentation for $\overline{S_d}$ will be $H = \{u, v|T_1 \cup T_2\}$, where $T_1 = \{u^4 = [u^2, (u^2)^v] = 1, (u^2)^v = [u^2, [u^2, u^v]]\}$, and T_2 is a set of relations in bijective correspondence with R. For each relator r in R the corresponding relator in T_2 is obtained by multiplying it by a certain product of conjugates of u^2. The first two relations in T_1 imply that $\langle u^2\rangle$ and its conjugates generate an elementary abelian group, and the quotient of H by this elementary abelian group is isomorphic to S_d, the centraliser of the image of u in this quotient being the direct copy of the image of u with a copy of S_{d-2}. Thus u^2 has $d(d-1)/2$ conjugates that represent the signed diagonal matrices with exactly two entries equal to -1. Writing $(\underline{i}, \underline{j})$ for the diagonal matrix with -1 in the (i, i) and (j, j) places and with the other diagonal entries set to 1, the third relation in T_1 is equivalent to the relation $(\underline{2}, \underline{3}) = (\underline{1}, \underline{2})(\underline{1}, \underline{3})$. Conjugating by elements of H that map to suitable elements of the copy of S_{d-1} that fix 1 we see that the normal closure of u^2 in H is of order 2^{d-1}, having a basis $\{(\underline{1}, \underline{i}) : i > 1\}$; and the presentation is correct, provided that the relators in T_2 are correctly calculated from the relators in R. These relators are calculated using Fox derivatives.

Definition 5.2 *Let F be the free group on $(x_i : i \in I)$. For each $i \in I$ the map $w \mapsto \partial w/\partial x_i$ (the Fox derivative of w with respect to x_i) is defined by the rules*

(i) $\partial x_i/\partial x_i = 1$;

(ii) $\partial x_j/\partial x_i = 0$ *if* $i \neq j$;

(iii) $\partial(uv)/\partial x_i = (\partial u/\partial x_i)v + \partial v/\partial x_i$ *for any* $u, v \in F$.

Lemma 5.3 *Let F be the free group on $(x_i) : i \in I$, let $w \in F$, and for all $i \in I$ let $a_i \in \mathbb{Z}F$. Then $w((x_i a_i)) = w((x_i)) \sum_i a_i \partial w/\partial x_i$.*

Proof: This basic result is an easy calculation. □

Lemma 5.4 *With the above notation $\partial w^n/\partial x_i = (\partial w/\partial x_i)(1 + w + \cdots + w^{n-1})$ if $n > 0$, and $\partial w^n/\partial x_i = -(\partial w/\partial x_i)(w^{-1} + w^{-2} + \cdots + w^n)$ if $n < 0$.*

Proof: This is an immediate consequence of the definition of the Fox derivavtive. □

If $r = r(u, v) \in R$, and if s is the corresponding relator in T_2, and if $(\underline{i})$ is the diagonal matrix with -1 in the (i, i) place and all other diagonal entries set to 1, then s evaluates to $r(u_0(\underline{1}), v_0)$ if d is odd, and to $r(u_0(\underline{1}), v_0(\underline{1}))$ if d is even, where u_0 and v_0

are the unsigned permutation matrices corresponding to U and V respectively. Thus the relator s corresponds to rc, where c is $(\underline{1})^C$, and C is $\partial r/\partial u$ if d is odd, and is $\partial r/\partial u + \partial r/\partial v$ if d is even. Since the relator evaluates to a matrix with determinant $+1$ it follows that the image of C under the augmentation map is even. Write the relator r as a straight line program on u and v, the allowable operations being multiplication and raising to a (possibly negative) integer exponent. Each cell in the straight line program then corresponds to a word w in u and v, and we construct, for each such cell, an element $C = C_w$ of $\mathrm{GF}(2)(F(u,v))$, such that if U and V are substituted for u and v then $(U^2)^C$ evaluates, in $\mathrm{GF}(2)(S_d)$, to $(U^2)^{\partial w/\partial u}(\underline{1})^{\epsilon}$ if d is odd, and to $(U^2)^{\partial w/\partial u + \partial w/\partial v}(\underline{1})^{\epsilon}$ if d is even, where $\epsilon = \epsilon_w \in \{0,1\}$. The value of ϵ is also noted.

A cell corresponding to u has $\epsilon = 1$, and a cell corresponding to v has $\epsilon = d + 1 \bmod 2$. A cell corresponding to the product of two previous cells has ϵ set to the sum (modulo 2) of the values of ϵ for those cells, and a cell corresponding to a power of a previous cell has epsilon set to the value of ϵ for that cell multiplied (modulo 2) by that exponent.

A cell corresponding to u has $C = 1$, and a cell corresponding to v has $C = 1$ if d is even, and $C = 0$ if d is odd.

A cell corresponding to gh, where previous cells correspond to g and h, has C set to $C_g^h + C_h$ if $\epsilon_g = 0$, and to $C_g^h + C_h + (vu)^{h(1)-1}$ if $\epsilon_g = 1$. Here g and h are elements of the free group on u and v; we define $h(1)$ and $g(1)$ to be the images of 1 under the homomorphism of the free group onto $S(d)$ that takes u to $(1,2)$ and v to $(1,2,\ldots,d)$.

A cell corresponding to g^n has C set to

- $C_g(1 + g + \cdots + g^{n-1})$ if $n > 0$ and $\epsilon_g = 0$,
- $C_g(1 + g + \cdots + g^{n-1}) + (vu)^{g(1)-1}(1 + g^2 + g^4 + \cdots + g^{2(m-1)})$ if $n = 2m$ is even and positive and $\epsilon_g = 1$,
- $C_g(1 + g + \cdots + g^{n-1}) + (vu)^{g(1)-1}(1 + g^2 + g^4 + \cdots + g^{2(m-1)}) + (vu)^{g^{n-1}(1)-1}$ if $n = 2m+1$ is odd and positive and $\epsilon_g = 1$,
- $C_g(g^{-1} + g^{-2} + \cdots + g^n)$ if $n < 0$ and $\epsilon_g = 0$,
- $C_g(g^{-1} + g^{-2} + \cdots + g^n)(vu)^{g^{-1}(1)-1}(1 + g^{-2} + g^{-4} + \cdots + g^{-m+1})$ if $n = -2m$ is even and negative and $\epsilon_g = 1$,
- $C_g(g^{-1} + g^{-2} + \cdots + g^n)(vu)^{g^{-1}(1)-1}(1 + g^{-2} + g^{-4} + \cdots + g^{-m+1}) + (vu)^{g^{-n}(1)-1}$ if $n = -2m-1$ is odd and negative and $\epsilon_g = 1$.

These formulae are obtained by evaluating the Fox derivatives, using the fact that $(\underline{i}) = (\underline{1})^{v^{i-1}}$, and $(\underline{1},\underline{i}) = (U^2)^{(vu)^{i-1}}$.

The last cell in the straight line program corresponds to the relation r, and the fact that the relator evaluates to a matrix with determinant $= 1$ implies that $\epsilon_r = 0$. Moreover the relator s corresponding to r is $r(u^2)^{C_r}$.

As written the relator s may have length $O(d)$ assuming, as is the case for the presentations of S_d that we use, that the exponents that arise in the relations for

S_d are of modulus $O(d)$. We can easily reduce the length of the relator s to having length bounded by a small fixed multiple of the length of the relator r in one of two (more or less equivalent) ways. We can write the relators for S_d without exponents at the cost of multiplying the length of the relators by a factor of at most 4; or we can recursively write $1 + g + \cdots + g^n = (1 + g)(1 + g^2 + g^4 + \cdots + g^{2(n/2)})$ if n is even, and $1 + g + \cdots + g^n = (1 + g + \cdots + g^{n-1}) + g^n$ if n is odd.

Let $\alpha = \omega^{\gcd(d,q-1)}$. Then the centre of $\mathrm{SL}(d, q)$ is generated by the scalar matrix αI_d, and to get a presentation for $\mathrm{PSL}(d, q)$ from a presentation for $\mathrm{SL}(d, q)$ one needs to construct αI_d as a word on the chosen generators for $\mathrm{SL}(d, q)$.

6 A short presentation for $\mathrm{SU}(2n, q)$

Our presentation is similar to our presentation for $\mathrm{Sp}(2n, q)$. We take $G = \mathrm{SU}(2n, q)$ to act on the vector space V of dimension $2n$ over $\mathrm{GF}(q^2)$, preserving a non-degenerate unitary form defined with respect to the hyperbolic basis $(e_1, e_2, \ldots, e_{2n})$, so that $e_{2u-1}.e_{2u} = e_{2u}.e_{2u-1} = 1$, for $1 \leq u \leq n$, and $e_i.e_j = 0$ for other pairs i, j with $1 \leq i, j \leq 2n$, and choose a primitive element ω of $\mathrm{GF}(q^2)$, and set $\alpha = \omega^{(q+1)/2}$, so that α^2 is a primitive element of $\mathrm{GF}(q)$. Let $\omega_0 = \omega^{q+1} = \alpha^2$, a primitive element of $\mathrm{GF}(q)$.

Assume first that q is odd.

Let $1 \leq u \neq v \leq n$, and let $\{i, j\} = \{2u - 1, 2u\}$ and $\{k, \ell\} = \{2v - 1, 2v\}$. For $\theta \in \mathrm{GF}(q)$ and $\phi \in \mathrm{GF}(q^2)$, define $\tau_{ij}(\alpha\theta)$ by $e_i \mapsto e_i + \alpha\theta e_j$, fixing the other basis elements, and $\sigma_{ijk\ell}(\phi)$ by $e_i \mapsto e_i + \phi e_\ell$, $e_k \mapsto e_k - \phi^q e_j$, fixing the other basis elements. Clearly these elements lie in G.

We take our generating set for G to be

$$\{\tau = \tau_{12}(\alpha), \sigma = \sigma_{2143}(1), \Delta, Z, U, V\}.$$

The generator Δ acts as the matrix

$$\begin{pmatrix} \omega^q & 0 & 0 & 0 \\ 0 & \omega^{-1} & 0 & 0 \\ 0 & 0 & \omega & 0 \\ 0 & 0 & 0 & \omega^{-q} \end{pmatrix}$$

on the first 4 basis vectors, fixing the others.

The generator Z acts on the first two basis elements as the matrix

$$\begin{pmatrix} 0 & \alpha \\ -\alpha^{-1} & 0 \end{pmatrix},$$

centralising the other basis elements. Note that $\sigma_{2143}(\phi)^Z = \sigma_{1243}(-\alpha\phi)$, $\sigma_{1243}(\phi)^Z = \sigma_{2143}(\alpha^{-1}\phi)$, $\tau_{12}(\alpha\theta)^Z = \tau_{21}(-\alpha^{-1}\theta)$, $\tau_{21}(\alpha\theta)^Z = \tau_{12}(-\alpha^3\theta)$.

The generators U and V are the same as in the case of $\mathrm{Sp}(2n, q)$.

The Steinberg generators are then the set of symbols of the form $\sigma_{ijk\ell}(\theta)$ for suitable suffices (i,j,k,ℓ) and $\theta \in \mathrm{GF}(q^2)$, or of the form $\tau_{ij}(\theta)$ for suitable suffices (i,j) and $\theta \in \mathrm{GF}(q^2)$ satisfying $\theta+\theta^q = 0$. Here 'suitable suffices' means that $\{i,j\} = \{2u-1, 2u\}$ and $\{k,\ell\} = \{2v-1, 2v\}$, where $1 \leq u \neq v \leq n$.

Note that $\theta + \theta^q = 0$ is equivalent to the condition $\theta = \alpha\theta_0$ for some $\theta_0 \in \mathrm{GF}(q)$.

The Steinberg relations are the following.

S(i) $\tau_{ij}(\theta)\tau_{ij}(\phi) = \tau_{ij}(\theta + \phi)$.

S(ii) $\sigma_{ijk\ell}(\theta)\sigma_{ijk\ell}(\phi) = \sigma_{ijk\ell}(\theta + \phi)$.

S(iii) $[\sigma_{ijk\ell}(\theta), \sigma_{jik\ell}(\phi)] = \tau_{kl}(\theta\phi^q - \phi\theta^q)$.

S(iv) $[\sigma_{ijk\ell}(\theta), \tau_{ij}(\phi)] = 1$.

S(v) $[\sigma_{ijk\ell}(\theta), \tau_{ji}(\phi)] = \sigma_{jik\ell}(\theta\phi^q)\tau_{k\ell}(-\phi\theta^{1+q})$.

S(vi) $[\tau_{ij}(\theta), \tau_{k\ell}(\phi)] = 1$.

S(vii) $\sigma_{ijk\ell}(\theta) = \sigma_{k\ell ij}(-\theta^q)$.

S(viii) $[\sigma_{ijk\ell}(\theta), \sigma_{k\ell xy}(\phi)] = 1$.

S(ix) $[\sigma_{ijk\ell}(\theta), \sigma_{\ell kxy}(\phi)] = \sigma_{ijxy}(\theta\phi)$.

S(x) $[\sigma_{ijk\ell}(\theta), \sigma_{xyuv}(\phi)] = 1$.

S(xi) $[\sigma_{ijk\ell}(\theta), \tau_{xy}(\phi)] = 1$.

In all suffices, distinct letters are assumed to represent distinct integers.

We can now describe our presentation for $\mathrm{SU}(2n,q)$.

When U, V and Z are referred to as permutations we will take these to be $(1,3)(2,4)$, $(1,3,\ldots,2n-1)(2,4,\ldots,2n)$ and $(1,2)$ respectively.

Theorem 6.1 *Let $G = \mathrm{SU}(2n,q)$ where $q = p^e$ for some odd prime p, and $n \geq 2$. Let H be the group defined by $\{X \mid \mathcal{R}\}$ where $X = \{\tau, \sigma, \Delta, Z, U, V\}$, with V omitted if $n = 2$. $\mathcal{R}$ is the union of the following sets of relations.*

R1. A set of defining relations for S_n on $\{U, V\}$, (or on U if $n = 2$) where U maps to the transposition $(1,2)$, and V to the cycle $(1,2,\ldots,n)$.

R2. A set of relations that, with R1, express the fact that $\langle U, V, Z\rangle$ (omit V if $n = 2$) is isomorphic to $C_4 \wr S_n$. These are the following:
$[Z, Z^U] = 1$.
$Z^4 = 1$.
If $n > 2$ then $[Z, U^V] = 1$.
If $n > 3$ then $[Z, VU] = 1$.

R3. *A set of relations that express the fact that the subgroup of* $\langle U, V, Z\rangle$ *that centralises* e_1 *and* e_2*, and that is isomorphic to* $C_4 \wr S_{n-1}$*, centralises* τ *and* $\delta := [\Delta, Z]$*. These are the following:*
$[\tau, Z^U] = [\delta, Z^U] = 1$.
If $n > 2$ *then* $[\tau, U^V] = [\delta, U^V] = 1$.
If $n > 3$ *then* $[\tau, VU] = 1$. ($[\delta, VU] = 1$ *is superfluous in light of R[5]).*

R4. *?? A set of relations that express the fact that conjugating by* Z^2 *centralises* τ *and inverts* σ*, and that conjugating by* Z *inverts* δ*.*

R5. *A set of relations that express the fact that the subgroup of* $\langle U, V, Z\rangle$ *that centralises* e_i *for* $1 \le i \le 4$*, and that is isomorphic to* $C_4 \wr S_{n-2}$*, centralises* σ *and* Δ*. These are the following:*
If $n > 2$ *then* $[\sigma, Z^{V^2}] = [\Delta, Z^{V^2}] = 1$.
If $n > 3$ *then* $[\sigma, U^{V^2}] = [\Delta, U^{V^2}] = 1$.
If $n > 4$ *then* $[\sigma, VUU^V] = [\Delta, VUU^V] = 1$.

R6. *A set of relations expressing the fact that* $\langle \delta, \tau, Z\rangle$ *is isomorphic to* $\mathrm{SL}(2, q)$*, where* δ *and* τ *and* Z *stand for* $\begin{pmatrix} \omega_0^{-1} & 0 \\ 0 & \omega_0 \end{pmatrix}$ *and* $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ *and* $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ *respectively. see Theorem 8.4*
If $e = 1$ *then* ω_0 *may be regarded as an integer* k *(defined modulo* p*).*

R7. *A set of relations expressing the fact that* $\langle \sigma, \Delta\rangle$ *is isomorphic to the group of upper triangular matrices in* $\mathrm{SL}(2, q^2)$*, where* Δ *maps to* $\begin{pmatrix} \omega^{-1} & 0 \\ 0 & \omega \end{pmatrix}$*, and* σ *maps to* $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$*, see Theorem 8.2.*

R8. *A set of relations that, with R6 and R7, give presentations of* $\langle \tau, \delta, \Delta, U\rangle$ *and* $\langle \sigma, \Delta, U\rangle$*. These are:* $[\delta, \Delta] = 1$; $\Delta^U = \Delta^q$; $\tau^\Delta = \prod_i \tau^{a_i \delta^i}$ *where* $\omega_0^{-1} = \sum_i a_i \omega_0^{2i}$; $[\tau, \tau^U] = [\delta, \delta^U] = [\tau, \delta^U] = 1$; $\sigma^U = \sigma^{-1}$.

R9. *Relations that express instances of the Steinberg relations. The instances are labelled to correspond to the above listing. We state the relation as a relation in the presentation, and in parentheses the relation as an instance of a Steinberg relation. The numbering of these relations corresponds to the Steinberg relations as above.*

S3. $[\sigma, \sigma'] = 1$, *where* $\sigma' = (\prod_i \sigma^{a_i \Delta^i})^Z$ *and* $-\alpha^{-1} = \sum_i a_i \omega^{2i}$
($[\sigma_{2143}(1), \sigma_{1243}(1)] = 1$);
$[\sigma^\Delta, \sigma'] = \left(\prod_i \tau^{c_i \delta^i}\right)^{\Delta Z U}$, *where* $\omega^{(7q+3)/2} - \omega^{(3q+7)/2} = \sum_i c_i \omega_0^{2i}$
($[\sigma_{2143}(\omega^2), \sigma_{1243}(1)] = \tau_{43}(\omega^2 - \omega^{2q})$.)

S4. $[\sigma, \tau^Z] = 1$;($[\sigma_{2143}(1), \tau_{21}(-\omega_0^{-1}\alpha)] = 1$).

S5. $[\sigma,\tau]=\sigma^Z\tau^{\Delta^{-1}ZU}$*;(*$[\sigma_{2143},\tau_{12}(\alpha)]=\sigma_{1243}(-\alpha)\tau_{43}(-\alpha)$*).*

S6. $[\tau,\tau^U]=1$*; (*$[\tau_{12}(1),\tau_{34}(1)]=1$*); but see R[8].*

S8. If $n>2$ *then* $[\sigma,\sigma^V]=1$*;(*$[\sigma_{2143}(1),\sigma_{4365}(1)]=1$*);*

S9. $[\sigma,\sigma^{ZV}] = \left(\prod_i\sigma^{a_i\Delta^i}\right)^{(U^V)}$*; where* $-\alpha=\sum_i a_i\omega^{2i}$*;(*$[\sigma_{2143}(1),\sigma_{3465}(-\alpha)] = \sigma_{2165}(-\alpha)$*).*

S10. If $n>3$ *then* $[\sigma,\sigma^{V^2}]=1$*; (*$[\sigma_{2143}(1),\sigma_{6587}(1)=1.$*)*

S11. If $n>2$ *then* $[\sigma,\tau^{V^2}]=1$*; (*$[\sigma_{2143}(1),\tau_{56}(\alpha)]=1.$*)*

R10. Relations to express Δ *and* U *as words in conjugates of* σ *and* τ*, namely* $\Delta = \sigma\sigma_1\sigma_2\sigma_3$*, where* $\sigma_1=\left(\sigma^{\Delta^U}\sigma^{-\Delta^Z\Delta^{-1}}\right)^{ZZ^U}$*, and* $\sigma_2=\prod_i\sigma^{a_i\Delta^i}$ *where* $-\omega=\sum_i a_i\omega^{2i}$*, and* $\sigma_3 = \left(\sigma^{\Delta^{-U}\Delta^Z\Delta^{-1}}\sigma^{\Delta^{-U}\Delta^{-1}}\right)^{ZZ^U}$*; and* $U = \sigma_4^{-Z^U}\sigma_4^Z\sigma_4^{-Z^U}Z^2$*, where* $\sigma_4 = \prod_i\sigma^{b_i\Delta^i}$*, and* $\sum_i b_i\omega^{2i} = -\omega^{-(q+1)/2}$*.*

(Here $\sigma_1=\sigma_{1234}(1-\omega^{-q})$*,* $\sigma_2=\sigma_{2143}(-\omega)$*,* $\sigma_3=\sigma_{1234}(\omega^{-2q}-\omega^{-q})$*,* $\sigma_4^Z=\sigma_{1243}(1)$*, and* $\sigma_4^{-Z^U}=\sigma_{2134}(1)$*.)*

Then $H=\langle X\mid\mathcal{R}\rangle$ *is isomorphic to* $\mathrm{SU}(2n,q)$*.*

Proof: Again we define elements $\tau_{ij}(\theta)$ and $\sigma_{ijk\ell}(\phi)$ of H for suitable suffices, where $\theta,\phi\in\mathrm{GF}(q^2)$, and $\theta\alpha^{-1}\in\mathrm{GF}(q)$.

Now $\tau_{12}(\theta)$ is an allowed element for $\theta=\theta_0\alpha$, where $\theta_0\in\mathrm{GF}(q)$. If $e=1$, so that θ_0 may be equated with an integer k defined modulo p, define $\tau_{12}(\theta_0\alpha)$ to be τ^k. If $e>1$ define $\tau_{12}(\theta_0\alpha)$ to be $\prod_i\tau^{c_i\delta^i}$ where $\theta_0=\sum_i c_i\omega_0^{-2i}$. These formulae are well defined by R6.

Define $\tau_{21}(\theta)$ to be $\tau_{12}(-\omega_0\theta)^Z$.

Now define $\tau_{2u-1,2u}(\theta)$ for suitable θ to be $\tau_{12}(\theta)^g$, where g is an element of $\langle U,V\rangle$ (as permutation group) that maps $(1,2)$ to $(2u-1,2u)$, and define $\tau_{2u,2u-1}(\theta)$ to be τ_{21}^g. This is well defined by R3.

Define $\sigma_{2143}(\theta)$ to be $\sigma^{\sum_i a_i\Delta^i}$ where $\theta=\sum_i a_i\omega^{2i}$. This is well defined by R7; define $\sigma_{1243}(\theta)=\sigma_{2143}(\alpha^{-1}\theta)^Z$; define $\sigma_{2134}(\theta)=\sigma_{2143}(\alpha^{-1}\theta)^{Z^U}$; and define $\sigma_{1234}(\theta)=\sigma_{2143}(-\omega_0^{-1}\theta)^{ZZ^U}$. If $\{i,j\}=\{3,4\}$ and $\{k,\ell\}=\{1,2\}$ define $\sigma_{ijk\ell}(\theta)=\sigma_{k\ell ij}(-\theta^q)$.

Define $\sigma_{2u,2u-1,2v,2v-1}(\theta)$ to be $\sigma_{2143}(\theta)^g$, where $g\in\langle U,V\rangle$, as a permutation, takes $(1,2,3,4)$ to $(2u-1,2u,2v-1,2v)$ if $u<v$, and to $(2v-1,2v,2u-1,2u)$ if $u>v$. Then define $\sigma_{2u-1,2u,2v,2v-1}(\theta)$ to be $\sigma_{1243}(\theta)^g$ if $u<v$ and to be $\sigma_{2134}(\theta)^g$ if $u>v$; and define $\sigma_{2u,2u-1,2v-1,2v}(\theta)$ and $\sigma_{2u-1,2u,2v-1,2v}(\theta)$ similarly.

This is well defined by R5.

We now check the Steinberg relations.

S(i) and S(ii) follow from R6 and R7.

S(iv) states that $[\sigma_{ijk\ell}(\theta),\tau_{ij}(\alpha\phi)]=1$, for all $\theta\in\mathrm{GF}(q^2)^\times$ and all $\phi\in\mathrm{GF}(q)^\times$, and we may take (i,j,k,ℓ) to be $(2,1,4,3)$. The case $\theta=1$, $\phi=\omega_0$ is given by S[4]. If the result holds for one value of (θ, ϕ), then, by conjugation with Δ, it follows that it holds for $(\theta\omega^2, \phi\omega_0)$, by R[?]. $(\sigma_{2143}(\theta)^{\Delta} = \sigma_{2143}(\omega^2\theta)$; $\tau_{21}(\phi)^{\Delta} = \tau_{21}(\omega_0\phi))$. So conjugating by $\Delta^{(q-1)/2}$ it follows that S(iv) then holds for $(\theta\omega^{q-1}, -\phi)$ and hence for $(\theta\omega^{q-1}, \phi)$. But if it holds for (θ_1, ϕ) and for (θ_2, ϕ) then it holds for $(\theta_1 + \theta_2, \phi)$; and ω^{q-1} generates GF(q^2) as a field over GF(p): so the result then holds for (θ, ϕ) where θ takes any value in GF(q^2), and this fixed value of ϕ. Hence, by conjugation by Δ, the result holds for all (θ, ϕ), and S(iv) is proved.

S(iii) states that $[\sigma_{ijk\ell}(\theta), \sigma_{jik\ell}(\phi)] = \tau_{k\ell}(\theta\phi^q - \phi\theta^q)$, and we take $(i, j, k, \ell) = (2, 1, 4, 3)$. By conjugation by Δ and by Δ^Z, if it holds for (θ, ϕ) it also holds for $(\theta\omega^2, \phi\omega^{1-q})$ and $(\theta\omega^{1-q}, \phi\omega^2)$. $(\sigma_{1243}(\theta)^{\Delta} = \sigma_{1243}(\omega^{1-q}\theta)$; $\sigma_{2143}(\theta)^{\Delta^Z} = \sigma_{2143}(\omega^{1-q}\theta)$; $\sigma_{1243}(\theta)^{\Delta^Z} = \sigma_{1243}(\omega^2\theta))$. Thus it holds for $(\theta\omega^{1-q^2}, \phi\omega^{2(1+q)}) = (\theta, \phi\omega_0^2)$. But since σ_{2143} and σ_{1243} centralise $\tau_{43}(\alpha\psi)$ for any ψ in GF(q) it follows that if S(iii) holds for (θ, ϕ_1) and for (θ, ϕ_2) it holds for $(\theta, \phi_1 + \phi_2)$. Since ${\omega_0}^2$ generates GF(q) as an extension field of GF(p) it follows then that if S(iii) holds for (θ, ϕ) it also holds for $(\theta, \phi\psi)$ for any $\psi \in$ GF(q), and similarly for $(\theta\psi, \phi)$. Now it holds for $(\theta, \phi) = (1, 1)$, and for $(\theta, \phi) = (1, \omega)$ by S[3], and hence for all $(1, \phi)$. Hence, by conjugation by Δ it holds for all (θ, ϕ), and S(iii) is proved.

S(vi) states that $[\tau_{ij}(\theta), \tau_{k\ell}(\phi)] = 1$. Again it suffices to take $(i, j, k, \ell) = (2, 1, 4, 3)$. Conjugating by δ and δ^U reduces to the case $\theta = \phi = \alpha$, which is S[6], or R[8].

S(v) states that $[\sigma_{ijk\ell}(\theta), \tau_{ji}(\phi)] = \sigma_{jik\ell}(\theta\phi^q)\tau_{k\ell}(-\phi\theta^{1+q})$. It follows from S(i) to S(iv) that if this holds for (θ_1, ϕ) and (θ_2, ϕ) it holds for $(\theta_1 + \theta_2, \phi)$, and if it holds for (θ, ϕ_1) and (θ, ϕ_2) then it holds for $(\theta, \phi_1 + \phi_2)$.

Conjugating by Δ proves that if S(v) holds for (θ, ϕ) then it holds for $(\omega^2\theta, \omega_0\phi)$, and hence for $(\omega^{2(q-1)}, \phi)$. Since $\omega^{2(q-1)}$ generates GF(q^2) over GF(p) it follows that if S(v) holds for one value of (θ, ϕ) then it holds for that value of ϕ and any value of θ. But conjugating by Δ shows that if it holds for one value of ϕ and some value of θ then it holds for all allowable values of ϕ and all values of θ; so if it holds for $(\theta, \phi) = (1, \alpha)$, as in S[v], then it holds for all (θ, ϕ).

S(vii) states that $\sigma_{ijk\ell}(\theta) = \sigma_{k\ell ij}(-\theta^q)$. This is built into our definitions.

Clearly S(viii) to S(xi) follow similarly from one instance, and this completes the proof of the Steinberg relations.

The proof of the theorem is now completed as in the symplectic case.

□

Now suppose that $p = 2$. In this case $Z^2 = 1$, and $\alpha = \omega_0$. The following changes in the presentation for odd q then follow.

In S3, replace the expression $\omega^{(7q+3)/2} - \omega^{(3q+7)/2}$ by $\omega^{3q+1} + \omega^{q+3}$, and replace α by ω_0.

In S9 replace α by ω_0.

In R10, in the expression $\sum_i b_i\omega^{2i} = -\omega^{-(q+1)/2}$, replace $\omega^{-(q+1)/2}$ by $\omega^{-(q+1)}$.

α should also be replaced by ω_0 in defining the element for which τ stands, and in the parenthetical expressions giving the relations in terms of the Steinberg generators.

7 A short presentation for $\mathrm{SU}(2n+1,q)$

Suppose first that q is odd.

The extra complication for the unitary groups in odd dimension is that the short root groups are now non-abelian of order q^3.

To deal with this phenomenon define

$$v_{ij}(\alpha) = \begin{pmatrix} 1 & \alpha & \alpha\overline{\alpha}\omega^{(q+1)/2} \\ 0 & 1 & -\overline{\alpha} \\ 0 & 0 & 1 \end{pmatrix}.$$

Here $(i,j) = (2k-1, 2k)$ for some k, and this matrix is with respect to the ordered basis (e_k, v, f_k). Other basis elements are centralised, and ω is a primitive element of $\mathrm{GF}(q^2)$.

Now introduce a new generator $v = v_{12}(1)$.

Also define

$$\Delta_{ij}(\theta) = \begin{pmatrix} \theta & 0 & 0 \\ 0 & \theta^{q-1} & 0 \\ 0 & 0 & \theta^{-q} \end{pmatrix},$$

and $\Delta = \Delta_{12}(\omega)$ with the same conventions, this Δ replacing the Δ for $\mathrm{SU}(2n, q^2)$.

For suitable suffices (i,j) a short root group K_{ij} is the normal closure in $H_{ij} = \langle v_{ij}(1), \Delta_{ij}(\omega)\rangle$ of $\langle v_{ij}(1)\rangle$.

7.1 A short bounded presentation for $H = H_{12}$

We first construct a short bounded presentation of $H = H_{12}$. The size of the presentation is $O(\log(q))$, and the number of relators is 9. The presentation is on $\{\tau, v, \Delta\}$, where, with respect the ordered basis (e_1, v, f_1), the matrix of τ is

$$\begin{pmatrix} 1 & 0 & \omega^{(q+1)/2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let $\theta = \omega^{q-2}$. Now $q \neq 2$, so θ generates $\mathrm{GF}(q^2)$ as a field. Let $m_1(x)$ denote the minimal polynomial of θ over $\mathrm{GF}(p)$. Since $q \neq 2$ it follows that $\theta + 1 \neq 0$. At this point there is a problem, solved below. The powers of θ are the cubes in $\mathrm{GF}(q^2)$. Suppose that $\theta+1$ is a cube in $\mathrm{GF}(q^2)$. Then define k_1 by $\theta^{k_1} = \theta + 1$. If $\theta+1$ is not a cube, we might replace $\theta+1$ by some suitable binomial in θ that is a cube. In the meanwhile, assume that $\theta+1$ is a cube; for example $p = 3$. Now define w_i for $1 \leq i \leq 4$ by the equations:

$$w_1 = v^{\Delta^{k_1} - \Delta - 1} \quad w_2 = v^{m_1(\Delta)} \quad w_3 = v^p \quad w_4 = [v^\Delta, v].$$

Note that w_i is, for all i, an element of $\langle\tau\rangle^{\langle\Delta\rangle}$. Thus w_i can be written as a word

$$\tau^{\Delta(c_e + \Delta(c_{e-1} + \Delta(+\cdots)))\ldots)},$$

in the style of Horner's algorithm, as an expression of size $O(\log q)$, with the c_i in $\mathrm{GF}(p)$.

Let $m_2(x)$ denote the minimum polynomial of $\omega_0^{-1} = \omega^{-(q+1)}$, and let $\omega_0^{-k_2} = \omega_0^{-1}+1$.

Lemma 7.1 *If $\theta+1$ is a cube in $\mathrm{GF}(q^2)$ then H has the following presentation.*

$$\begin{aligned}\{\tau, \upsilon, \Delta \quad | \quad & \upsilon^{\Delta^{k_1}-\Delta-1} = w_1, \upsilon^{m_1(\Delta)} = w_2, \upsilon^p = w_3, [\upsilon^\Delta, \upsilon] = w_4,\\ & \Delta^{q^2-1} = \tau^{\Delta^{k_2}-\Delta-1} = \tau^{m_2(\Delta)} = \tau^p = [\tau^\Delta, \tau] = 1\}\end{aligned}$$

Proof: Let G be the group defined by this presentation, and by abuse of notation for the duration of the proof, let τ, υ, Δ denote the images of these elements in G, or in specified images of G. Let $N = \langle\tau\rangle^G$, and let $K = G/N$. Note that these relations do hold in the root group in question, so G has order at least $q^3(q^2-1)$, and K has order at least $q^2(q^2-1)$.

Step 1. Prove that $K \cong \mathrm{GF}(q^2)^+ : \mathrm{GF}(q^2)^\times$. Here a generator of the copy of $\mathrm{GF}(q^2)^\times$ acts as the cube of a primitive element of $\mathrm{GF}(q^2)$, and hence does not act faithfully if $q \equiv 2 \bmod 3$. This is because Δ acts as ω^{q-2}, and the greatest common divisor of q^2-1 and $q-2$ is 3 if $q \equiv 2 \bmod 3$, and is otherwise 1.

In Step 1 we take τ, υ, Δ to denote the images of these elements in K. Let $A = \langle\upsilon\rangle^K$, so $K/A \cong \mathrm{GF}(q^2)^\times$. It suffices to prove that $[\upsilon, \upsilon^{\Delta^i}] = 1$ for all i, since then the relation $\upsilon^{m_1(\Delta)} = 1$ in K proves that A has order at most q^2. Now let J be the set of integers j such that $[\upsilon, \upsilon^{\Delta^j}] = 1$. Observe the following properties of J.

(i) $j \in J \iff -j \in J$.

(ii) $0, 1 \in J$.

(iii) If, for some integer i, two of the integers $i, i+1, i+k_1$ lie in J then so does the third.

Of these, (i) is trivial, (ii) is a relation from the presentation, and (iii) arises from the identity $\upsilon^{\Delta^{k_1+i}} = \upsilon^{\Delta^{i+1}}\upsilon^{\Delta^i}$, which in turn arises from the given case $i=0$ by conjugation by Δ^i.

We now prove by induction that these three properties imply that $\{\pm c, \pm(k_1+c)\} \subset J$ for every $c \geq 0$. The case $c=0$ is trivial; so assume that the result holds for some value of c. Now $\{c, k_1+c\} \subset J$; so by (iii) $c+1 \in J$. Then $\{-k_1-c, -c-1\} \subset J$, so by (iii) $-k_1-c-1 \in J$. Thus $\{\pm(c+1), \pm(k_1+c+1)\} \subset J$, and the induction is complete.

Step 2. Prove that N is elementary abelian of order q. The proof is effectively the same as the proof of Step 1, with k_2 playing the role of k_1. □

Now suppose that $\theta+1$ is not a cube, and that no suitable binomial term can be constructed for which it is a cube. There is a primitive element λ of $\mathrm{GF}(q^2)$ such that $\theta = \lambda^3$. Now G/N is a subgroup of index 3 in $\mathrm{GF}(q^2)^+ : C_{3(q^2-1)}$. Here the cyclic group acts as λ, so the q^2-1-st power of the cyclic group centralises $\mathrm{GF}(q^2)^+$. Now define k_1

by $\lambda^{k_1} = \lambda + 1$, and take $m_1(x)$ to be the minimum polynomial of λ over $\mathrm{GF}(p)$. Now the method of the Lemma will apply, and since G/N has index 3 in this group, a short bounded presentation of the larger group gives rise to a short bounded presentation of this subgroup.

It remains to deal with the Steinberg relations of type (ii). One of these is approximately as follows.

$$[\upsilon_{ij}(\alpha), \upsilon_{k\ell}(\beta)] = \sigma_{ijk\ell}(\alpha\overline{\beta})$$

when $i = j + 1$ and $k = \ell + 1$, or when $i = j - 1$ and $k = \ell - 1$.

8 Presentations for $\mathrm{SL}(2, p^e)$

Our treatment of these presentations follows [4]. We assume throughout this section that $e > 1$.

Theorem 8.1 *Let ω be a primitive element of $\mathrm{GF}(p^e)$, where p is any prime, and define m so that $\omega^m = 1 + \omega$. Let $\sum_{i=0}^{e} c_i x^i$ be the characteristic polynomial of ω. Then*

$$H = \{a, b : a^{p^e - 1} = b^p = [b, b^a] = 1, b^{a^m} = b.b^a, \prod_i b^{c_i a^i} = 1\}$$

is a presentation of the Frobenius group $\mathrm{GF}(p^e) : \mathrm{GF}(p^e)^\times$, where a stands for ω in $\mathrm{GF}(p^e)^\times$, and b stands for 1 in $\mathrm{GF}(p^e)$. Here $\mathrm{GF}(p^e)^\times$ acts on $\mathrm{GF}(p^e)$ by multiplication.

PROOF: Clearly b lies in H'; so H/H' is cyclic of order $p^e - 1$, generated by the image of a. Now suppose that, for some $i > 0$, $[b, b^{a^j}] = 1$ whenever $1 \leq j < i$; so in particular, $[b, b^{a^{i-1}}] = 1$. Conjugating by a^m gives $[b.b^a, b^{a^{i-1}} b^{a^i}] = 1$, and hence $[b, b^{a^i}] = 1$. It follows by induction that b commutes with b^{a^i} for all i, and hence that H' is an elementary abelian p-group. The last relation imposes the condition that a acts on H' by multiplication by ω. □

Theorem 8.2 *Let $F = \mathrm{GF}(p^e)$ where p is an odd prime and F has size at least 5. Let ω be a primitive element of F, and let $\ell = (p^e - 1)/2$. Let b_i and u_i for $0 \leq i \leq e$ be defined as above. Let H be the image of the subgroup of upper triangular matrices of $\mathrm{SL}(2, F)$ in $\mathrm{PSL}(2, F)$. Suppose first that $1 + \omega$ is a square, and let m satisfy $\omega^{2m} = 1 + \omega$. Then H has the following presentation:*

$$\begin{aligned}\{a, b_0, b_1 \;\mid\; & b_2 = a^{-1} b_0 a,\ a^\ell = b_0^p = [b_0, b_1] = [b_1, b_2] = \prod_{i=0}^{e} b_i^{u_i} = 1, \\ & a^{-m} b_0 a^m = b_0 b_1, \\ & a^{-m} b_1 a^m = b_1 b_2\}.\end{aligned}$$

If $1+\omega$ is not a square, let m satisfy $\omega^{2m+1} = 1+\omega$. Then H has the following presentation:

$$\begin{aligned}\{a, b_0, b_1 \quad | \quad & b_2 = a^{-1}b_0a,\ a^\ell = b_0^p = [b_0, b_1] = [b_1, b_2] = \prod_{i=0}^{e} b_i^{u_i} = 1,\\ & a^{-m}b_1a^m = b_0b_1,\\ & a^{-m-1}b_0a^{m+1} = b_1b_2\}.\end{aligned}$$

Proof: We list elements of $\mathrm{PSL}(2, p^e)$ whose images satisfy these presentations.

$$\delta := \begin{pmatrix}\omega^{-1} & 0\\ 0 & \omega\end{pmatrix} \longmapsto a, \quad \tau := \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} \longmapsto b_0, \quad \tau_1 := \begin{pmatrix}1 & \omega\\ 0 & 1\end{pmatrix} \longmapsto b_1.$$

Note that $1+\omega \neq 0$ since $p^e > 3$.

The proof that H/H' is cyclic of order ℓ, and that H' is elementary abelian, are similar to the proofs of the corresponding facts in the previous theorem. Now H' is generated, as a normal subgroup of H, by b_0; so the last relation enforces the condition that H' is isomorphic to $\mathrm{GF}(p^e)$, and that a acts by multiplication by ω^2. □

We expand this presentation to give a presentation of $\mathrm{PSL}(2, p^e)$ on the generating set $\{\tau, \delta, Z\}$, where Z is defined by $e_1 \mapsto e_2 \mapsto -e_1$. Our starting point is the presentation $G = \langle w, x, y, r \mid R_1 \cup R_2\rangle$ of $\mathrm{PSL}(2, p^e)$, where

$$w = \begin{bmatrix}0 & -1\\ 1 & -1\end{bmatrix}, \quad x = \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}, \quad y = \begin{bmatrix}1 & \omega\\ 0 & 1\end{bmatrix}, \quad r = \begin{bmatrix}\omega & \omega^{-1}\\ 0 & \omega^{-1}\end{bmatrix},$$

and R_1 is $\{w^3 = (wr)^2 = (wx)^2 = (wyr)^3 = 1\}$, and R_2 is a set of $2n+3$ relations on $\{x, y, r\}$. This presentation is due to Sinkov [11], and is given explicitly on p 335 of [4]. We replace this generating set by $\{\tau, \delta, Z\}$. Note that $w = (\tau Z)^{-1}$, $x = \tau$, $y = \tau^{\delta^m}\tau$, $r = \tau\delta$. Now R_2 can be replaced by any presentation of the group of upper triangular matrices on $\{\tau, \tau_1, \delta\}$, and we take the presentation of Theorem 8.2. Thus we arrive at the following.

Theorem 8.3 *Let $\omega = \sum_{i=0}^{e-1} a_i\omega^{2i}$. With the notation of the previous theorem,* $\mathrm{PSL}(2, p^e)$ *has the presentation $\{\tau, \delta, Z \mid R_1 \cup R_2\}$, where*

$$R_1 = \{\tau_1 = \prod_i \tau^{a_i\delta^i},\ (\tau Z)^3 = (Z\delta)^2 = Z^2 = (\tau_1 Z\delta)^3 = 1\},$$

and R_2 is obtained from the relations of Theorem 8.2 by replacing a, b_0 and b_1 by δ, τ, and τ_1 respectively.

This is may be compared with Theorem 2.2 of [4], which gives a thirteen relator presentation on the generators w, x, y, r. Ours is a three generator eleven relator presentation (disregarding the relation that defines τ_1). Our presentation is essentially the same as theirs, but with the redundant relations $y^p = s^{tf(t)} = 1$ in their notation omitted.

Finally we need to construct from this a presentation for $\mathrm{SL}(2, p^e)$. Let S_1 be the set of relations $\{\tau_1 = \prod_i \tau^{a_i \delta^i}, Z^4 = (\tau_1 Z\delta)^3 = 1, (\tau Z^{-1})^3 = (Z\delta)^2 = Z^2\}$. Let S_2 be the set of relations in R_2 with $\delta^{(p^e-1)/2} = 1$ replaced with the relation $\delta^{(p^e-1)/2} = Z^2$.

Theorem 8.4 *With the above notation* $\{\tau, \delta, Z \mid S_1 \cup S_2\}$ *is a presentation for* $\mathrm{SL}(2, p^e)$.

PROOF: This follows at once from the previous theorem. □

Thus we have a thirteen relator presentation of $\mathrm{SL}(2, p^e)$ on our chosen set of three generators, ignoring the relation used to define τ_1. In considering the length of the presentation we must include this relation. But in considering the length of the relations we may assume that δ^m is only counted once. Thus the length of the presentation is the sum of the lengths associated to the lengths of the words δ^ℓ, δ^m and τ^p, plus the length of the relation required to define τ_1, plus the length of the relation that encodes the characteristic polynomial of ω, plus a small absolute constant. This comes to approximately $8 \log_2 p^e + 2 \log_2 p$ if one is calculating the number of group multiplications and inversions required to evaluate the relators. If one bears in mind the fact that δ^ℓ and δ^m may be replaced by expressions in which successive squaring is used to eliminate exponents, then since the squaring operations need only be carried out once the length of the presentation becomes approximately $7 \log_2 p^e + 2 \log_2 p$; but these estimates are not of much practical significance as matrices may be raised to high powers by more efficient means than successive squaring.

Note that the time required to find the presentation includes the time required to compute m, which is a discrete log problem in the field F, and this is asymptotically slower than evaluating the words in the presentation.

If $p^e \equiv 3 \bmod 4$ then [4] Theorem 2.4 gives a simpler presentation for $\mathrm{PSL}(2, p^e)$. This arises as follows. The presentation in Theorem 8.1 is simpler than the presentation in Theorem 8.2. The reason for this is that ω^2 does not generate $\mathrm{GF}(p^e)^\times$ as a cyclic group, and so in Theorem 8.2 τ and τ_1 are not conjugate in H. But if $p^e \equiv 3 \bmod 4$ then τ and $-\tau_1$ are conjugate in H, and the simpler presentation arises from using this fact.

Translating their presentation into a presentation on our generators, and expanding the presentation to a presentation of $\mathrm{SL}(2, p^e)$, we get the following presentation.

Theorem 8.5 *If* $p^e \equiv 3 \bmod 4$ *define* $b_0 = \tau$; *and* $b_1 = \tau^{-\delta^n}$, *where* $n = (p^e+1)/4$; *and* $b_{i+2} = b_i^\delta$ *for* $i > 1$; *and let* $\sum_i u_i x^i$ *be the minimum polynomial of* ω. *Let* $\ell = (p^e-1)/2$; *and let* $\gamma = \delta^{\lfloor k/2 \rfloor}$ *where* $1 + \omega = \omega^k$. *Then* $\mathrm{SL}(2, p^e)$ *has the presentation* $\{\tau, \delta, Z \mid Z^4 = \delta Z \delta Z^{-1} = \prod_i b_i^{u_i} = [\tau, \tau^{\delta^n}] = 1, (\tau Z^{-1})^3 = Z^2, (\tau\delta)^\ell = \tau^p Z^2, \tau^\gamma = [\tau^{-1}, \delta^{(-1)^k n}]\}$.

Similar results hold for characteristic 2.

Theorem 8.6 *Let $F = \mathrm{GF}(2^e)$. Let ω be a primitive element of F, let m satisfy $\omega^{2m} = 1 + \omega^2$, and let $\ell = 2^e - 1$. Let the minimal polynomial of ω^2 be $\sum_i u_i x^i$. Define $b_0 = b$, and $b_i = b_{i-1}^a$ for $i > 0$. Let H be the group of upper triangular matrices in $\mathrm{SL}(2, F)$. Then H has the following presentation:*

$$H = \langle a, b \mid a^\ell = b^2 = 1 \ , \ a^{-m} b a^m = [b, a] \ , \ \prod_i b_i^{u_i} = 1 \rangle.$$

This is Theorem 3.1 of [4], with $j = 1$. Here a and b stand for δ and τ respectively, as in the case of odd characteristic. The reason that the presentation is simpler than the corresponding presentation given in Theorem 8.2 is the fact that $\mathrm{GF}(p^e)^\times$ is generated by ω^2, as a cyclic group, if and only if $p = 2$.

Now Theorem 3.2 of [4], translated into our generating system, gives the following:

Theorem 8.7 *With the above notation, for $e > 2$, with $b_0 = \tau$ and $b_i = b_{i-1}^\delta$ for $i > 0$, $\mathrm{SL}(2, 2^e)$ has the presentation*

$$\mathrm{SL}(2, 2^e) = \langle \tau, \delta, Z \mid (Z\tau)^3 = Z^2 \ , \ (Z\delta)^2 = (\tau\delta)^\ell = \tau^2 \ , \ \tau^{\delta^m} = [\tau, \delta] \ , \ \prod_i b_i^{u_i} = 1 \rangle.$$

In fact Theorem 3.2 of [4] is somewhat more general. They allow a relation $\tau^{\delta^m} \tau^{\delta^j} \tau$ for any m and j, with j odd, such that $\omega^{2m} + \omega^{2j} + 1 = 0$. This is our penultimate relation with $j = 1$. The purpose of this extra degree of generality is that, provided that the polynomial $t^m + t^j + 1 = 0$ has precisely e zeros in $\mathrm{GF}(2^e)$, the relation $\prod_i b_i^{u_i} = 1$ is redundant, both here and in the previous theorem. They conjecture that m and j can always be chosen to satisfy this condition, which will be satisfied if $m = e$ and the polynomial is irreducible. Unfortunately this comes at the cost of specifying the primitive element ω (up to conjugacy). Suppose then that we find m and j to satisfy the above conditions, thus defining a primitive element α in $\mathrm{GF}(2^e)$. If $\alpha = \omega^s$ it is easy to convert the presentation defined in terms of α into a presentation defined in terms of ω. Suitable trinomials have been constructed for e up to 1000 (reference 19 of CRW paper). For small values of e MAGMA will have a preferred primitive element, and for these values we can pre-compute s. For large values of e any primitive element will do, and we can allow the chosen values of m and j to determine the primitive element.

Finally, we note the following presentation in [3].

Theorem 8.8 *Let p be an odd prime. Then $\mathrm{SL}(2, p)$ has the following presentation*

$$\{a, b \mid b^2 = (ba)^3, (ba^4 b a^{(p+1)/2})^2 a^p b^{2k}\}$$

where $k = \lfloor p/3 \rfloor$.

If $p \equiv 1 \bmod 3$ then the matrices in $\mathrm{SL}(2, p)$ corresponding to a and b are $-\tau$ and Z respectively, otherwise they are τ and $-Z$. To turn a presentation on $\{-\tau, Z\}$ to a presentation on $\{\tau, Z\}$ one can simply replace a by ab^2 in the relators, and to replace a presentation on $\{\tau, -Z\}$ into a presentation on $\{\tau, Z\}$ similarly replace b by b^{-1}.

9 Outstanding issues

SOME THOUGHT NEEDS TO GO INTO KILLING THE CENTRE IN ODD CHARACTERISTIC.

References

[1] László Babai, William M. Kantor, Péter P. Pálfy and Ákos Seress, Black box recognition of finite simple groups of Lie type by statistics of element orders, *J. Group Theory* **5** (2002), 383–401.

[2] L. Babai, A.J. Goodman, W.M. Kantor, E.M. Luks, and P.P. Pálfy. Short presentations for finite groups. *J. Algebra*, **194**, 79–112, 1997.

[3] C.M. Campbell and E.F. Robertson. A deficiency zero presentation for $\mathrm{SL}(2, p)$. *Bull. London Math. Soc.* **12**:17–20, 1980.

[4] C.M. Campbell, E.F. Robertson, and P.D. Williams. On Presentations of $\mathrm{PSL}(2, p^n)$. *J. Austral. Math. Soc* (Series A) **48**:333–346, 1990.

[5] Charles W. Curtis. Central extensions of groups of Lie type. *J. Reine Angew. Math.*, 220:174–185, 1965.

[6] M.D.E. Conder, C.R. Leedham-Green, and E.A. O'Brien. Short presentations for the alternating and symmetric groups. Preprint, 2005.

[7] Alexander Hulpke and Ákos Seress. Short presentations for three-dimensional unitary groups. *J. Algebra*, 245(2):719–729, 2001.

[8] W.M. Kantor and Á. Seress, Black box classical groups. *Mem. Amer. Math. Soc.* **149**, (2001), viii+168.

[9] C.R. Leedham-Green and E.A. O'Brien, Constructive recognition for classical groups. in odd characteristic. Preprint.

[10] E.A. O'Brien, Towards effective algorithms for linear groups. To appear in *Finite Geometries, Groups and Computation*, (Colorado), September 2004.

[11] A.Sinkov. A note on a paper by J.A. Todd. *Bull. Amer. Math. Soc.***32**, 762–765, 1939.

[12] Robert Steinberg. Générateurs, relations et revêtements de groupes algébriques. In *Colloq. Théorie des Groupes Algébriques (Bruxelles, 1962)*, pages 113–127. Librairie Universitaire, Louvain, 1962.

[13] Robert Steinberg. Generators, relations and coverings of algebraic groups. II. *J. Algebra*, 71(2):527–543, 1981.

[14] Charles C. Sims, *Computation with finitely presented groups.* Cambridge University Press, 1994.

[15] Michio Suzuki. On a class of doubly transitive groups. *Ann. of Math. (2)*, 75:105–145, 1962.

[16] H.J. Zassenhaus. A presentation of the groups PSL$(2, p)$ with three defining relations. *Canad. J. Math.*, 21:310–311, 1969.

Marston Conder ⟨conder@math.auckland.ac.nz⟩
E.A. O'Brien ⟨obrien@math.auckland.ac.nz⟩:
Department of Mathematics, University of Auckland, Private Bag 92019, Auckland, New Zealand

C.R. Leedham-Green ⟨C.R.Leedham-Green@qmul.ac.uk⟩: School of Mathematical Sciences, Queen Mary, University of London, London E1 4NS, United Kingdom