Short presentations for alternating and symmetric groups

J.N. Bray, M.D.E. Conder, C.R. Leedham-Green, and E.A. O'Brien

Abstract

We derive new families of presentations (by generators and relations) for the alternating and symmetric groups of finite degree n. These include presentations of length that are linear in $\log n$, and 2-generator presentations with a bounded number of relations independent of n.

1 Introduction

A long-standing question of interest is the existence of short-length presentations for the alternating and symmetric groups of degree n; see [5, 7] for example.

Babai *et al.* [1] define the *length* of a presentation as the number of symbols required to write down the presentation. In this context, each generator is a single symbol, and relations can be transformed into relators (namely words in the generators that present the trivial element), each of which can be written as a string of symbols. Accordingly, the length of a presentation is equal to the number of generators plus the sum of the lengths of the relators. For powers of elements, the exponent can be written in binary form, and so (for example) the length of $(xy)^m$ for large m is $2 + \log_2(m)$.

Perhaps the best known family of presentations for the finite symmetric groups are their presentations as Coxeter groups (see [7] for example), attributable to E.H. Moore [13]. In these, the symmetric group S_n of degree n is presented in terms of the transpositions $t_k = (k, k+1)$ for $1 \leqslant k < n$, which generate S_n and satisfy the defining relations ${t_k}^2 = 1$ for $1 \leqslant k < n$, and $(t_{k-1}t_k)^3 = 1$ for $1 < k < n$, and $(t_j t_k)^2 = 1$ for $1 \leqslant j < k-1 < n-1$. The number of relations is $n-1+(n-1)(n-2)/2 = n(n-1)/2$, and the presentation length is $n-1+2(n-1)+6(n-2)+4(n-2)(n-3)/2 = 2n^2-n-3$, both of which are quadratic in the degree n.

In this paper we construct new families of presentations for the symmetric groups that reduce the presentation length, first to linear in n (in Section 2), then to quadratic

2000 *Mathematics Subject Classification*: 20F05, 20B30 (primary). This work was supported in part by the Marsden Fund of New Zealand via grant UOA 0412. We thank George Havas, Bill Kantor and Igor Pak for comments on the paper.

in $\log n$ (in Section 3), then to $O(\log^\alpha n \log\log n)$ where $\alpha = \log_2(1+\sqrt{2}) \approx 1{\cdot}27155$ (in Section 4). All these presentations are obtained recursively from symmetric subgroups of symmetric groups. We then turn to a different technique, using presentations of $\mathrm{PSL}(2,p)$ and $\mathrm{PGL}(2,p)$ for suitable primes p. We construct a family of presentations for S_n in which the number of relations is bounded, answering a long-standing question that was raised again recently by Alex Lubotzky and Bill Kantor (in a personal communication), and finally we give families of presentations for S_n of length $O(\log n)$ having a uniformly bounded number of generators and relations.

In Section 3, we present the key idea of obtaining a presentation for S_{m+n} or S_{m+n-1} by 'gluing' together presentations for S_m and S_n. If $m = n$ we can avoid repeating the presentation for S_n, and this enables us to construct efficiently a presentation for S_{2m}, or for S_{2m-1}, from a presentation for S_m. This provides an inductive process of obtaining a presentation for S_n in $O(\log n)$ steps. In Section 4 a similar technique is used to combine presentations for S_m and S_n to produce a presentation for S_{mn}. Our gluing technique can then be used to obtain short-length presentations for symmetric groups of other degrees.

In Section 5, we take quite a different approach: we use short presentations for the 3-transitive permutation groups $\mathrm{PGL}(2,p)$ of degree $p+1$, for prime p, to build presentations for symmetric groups of degree n where $n-1$ or $n-2$ is expressible as a sum of two primes. An application of the Goldbach conjecture (or more precisely its three-prime version proved by Vinogradov [19] for sufficiently large odd n, or its six-prime version proved by Ramaré [15] for all even n), together with our gluing techniques then gives presentations for S_n with two generators and a bounded number of relations, for all n. We improve these in Section 6, giving presentations for S_n not just with a bounded number of generators and relations, but also having length $O(\log n)$ — which is the best possible since merely specifying the integer n needs $O(\log n)$ bits.

In Section 7, we explain briefly how short-length presentations for the alternating groups A_n can be obtained from those for S_n. In particular, this gives presentations for A_n with two generators and a bounded number of relations, for all large n, thereby answering a question of Campbell *et al.* from [5, Section 7].

In Section 8, we record the shortest presentations we have found (by other means) for the symmetric groups of degree up to 9. From a computational perspective, these presentations are important base cases, which can improve results obtained using the constructions given in Sections 2 to 5.

Simultaneously and independent of our work, Guralnick, Kantor, Kassabov and Lubotzky [8] obtained similar results for A_n and S_n. In a major extension, they use the short presentations for the corresponding Weyl groups to prove that every nonabelian finite simple group of rank r over $\mathrm{GF}(q)$ with the possible exception of the Ree groups ${}^2G_2(q)$ has a presentation with a bounded number of generators and relations and total length $O(\log r + \log q)$. In [6], we use our short presentations for the Weyl groups to write down explicit presentations of this length for the classical groups. The previously shortest known presentations have length at most $O(r^4(\log q)^2)$; see [1].

The various presentations for the alternating and symmetric groups satisfy differ-

ent demands. In particular, for computational purposes, we often need a presentation on the "standard generators", $\{(1,2),(1,2,\ldots,n)\}$, of S_n. For example, such presentations can be used in conjunction with the algorithms of [2] to verify constructive recognition of black-box representations of these groups. For further applications of short presentations, see [14], where a preliminary version of our results was announced.

We conclude by commenting on the concept of presentation length. Our definition (due to Babai *et al.* [1]) accords well with the requirement of computational complexity. We could, however, assume instead that the generators in a presentation are indexed symbols, and count the length of subscripts as part of the length; if the generating set has size k then this would increase the length of the presentation by a factor of at most $O(\log k)$. Such a simplification does not change the asymptotic complexity of the lengths of the presentations in this paper.

2 A presentation for S_n of length $O(n)$

In this section, we show how to construct a presentation for S_n having length that is linear in n.

First, consider the presentation for S_n as a Coxeter group given in the Introduction. A shorter presentation can be obtained by introducing the n-cycle $y = (1,2,3,\ldots,n)$ as a new generator and then using the fact that $y^{-(i-1)}t_1y^{i-1} = (i,i{+}1) = t_i$ to eliminate the generator t_i for $2 \leqslant i < n$. Letting $x = t_1$ and eliminating further redundancy of relations under conjugation by the long cycle y, the presentation may be taken as

$$\{x, y \,|\, x^2 = y^n = (xy)^{n-1} = 1,\ (xy^{-1}xy)^3 = 1,\ (xy^{-j}xy^j)^2 = 1 \text{ for } 2 \leqslant j \leqslant \lfloor n/2 \rfloor\}.$$

Note that the relation $(xy)^{n-1} = 1$ can be used to express y as a product of conjugates of x and hence as a product of the t_i. The number of relations is now $3 + \lfloor n/2 \rfloor$, which is linear in n, and the presentation length is $O(n \log n)$.

Now, however, we may define $y_i = y^i$ for $1 \leqslant i \leqslant \lfloor n/2 \rfloor$, and add these to the presentation as additional (redundant) generators, by adding also the relations $y_1 = y$ and $y_i = y_{i-1}y$ for $2 \leqslant i \leqslant \lfloor n/2 \rfloor$. Once we do that, we can replace the relation $(xy^{-j}xy^j)^2 = 1$ by $(xy_j^{-1}xy_j)^2 = 1$ for $2 \leqslant j \leqslant \lfloor n/2 \rfloor$, which reduces the presentation length from quadratic in n to linear in n. Indeed we have the following:

Theorem 2.1 *For all $n \geqslant 3$, the symmetric group S_n of degree n has presentation*

$$\begin{aligned}\{x, y_1, y_2, \ldots, y_{\lfloor n/2 \rfloor} \,|\, & x^2 = y_1^n = (xy_1)^{n-1} = 1,\ (xy_1^{-1}xy_1)^3 = 1,\\ & y_{j-1}y_1y_j^{-1} = 1 \text{ for } 2 \leqslant j \leqslant \lfloor n/2 \rfloor,\ (xy_j^{-1}xy_j)^2 = 1 \text{ for } 2 \leqslant j \leqslant \lfloor n/2 \rfloor\},\end{aligned}$$

with $\lfloor n/2 \rfloor + 1$ generators, $2\lfloor n/2 \rfloor + 2$ relations, and presentation length bounded above by $(\lfloor n/2 \rfloor + 1) + (3n + 12 + 11(\lfloor n/2 \rfloor - 1)) = 3n + 12\lfloor n/2 \rfloor + 2 \leqslant 9n + 2$.

Note that linearity of the presentation length has been achieved by introducing additional (redundant) generators.

3 Building presentations for S_{m+n} from S_m and S_n

In this section, we describe a 'doubling' construction that produces presentations for S_{2n-1} and S_{2n} from a given presentation for S_n involving the generators $x = (1, 2)$ and $y_n = (1, 2, \ldots, n)$. Inductive use of this construction then gives a presentation for S_n with length that is quadratic in $\log n$.

We first give a simpler construction for building a presentation for S_{m+n} from presentations for S_m and S_n, which will be used also in Sections 4 and 5. The key idea is to 'glue' the presentations together using a single transposition:

Theorem 3.1 *Let $P = \{A \,|\, \mathcal{R}\}$ and $Q = \{B \,|\, \mathcal{S}\}$ be presentations for the symmetric groups S_m and S_n of degrees $m, n \geqslant 3$, such that the generating set A for S_m contains elements a and v standing for the transposition $(1, 2)$ and the m-cycle $(1, 2, \ldots, m)$ respectively, and the generating set B for S_n contains elements b and w standing for the transposition $(1, 2)$ and the n-cycle $(1, 2, \ldots, n)$ respectively. Then*

$$\begin{aligned}\{\, A, B, t, y \,|\, \mathcal{R}, \mathcal{S}, t^2, (at)^3, (tb)^3, [a, b], [a, w], [v, b], [v, w],\\ [av, t], [vav^{-1}, t], [t, wb], [t, w^{-1}bw], y^{-1}wtv \,\}\end{aligned}$$

is a presentation for S_{m+n} on a generating set that includes the elements y standing for the $(m+n)$-cycle $(1, 2, \ldots, m+n)$ and t standing for a transposition of the form $(i, i+1)$. In particular, this presentation has $|A| + |B| + 2$ generators and $|\mathcal{R}| + |\mathcal{S}| + 12$ relations, and presentation length at most $\ell(P) + \ell(Q) + 64$ where $\ell(P)$ and $\ell(Q)$ are the lengths of the presentations P and Q. Moreover, the elements of $A \cup B$ are redundant generators, with $a = yty^{-1}$, $v = (yt)^{m-1}y^{-(m-1)}$, $b = y^{-1}ty$ and $w = y^{-(n-1)}(ty)^{n-1}$. Hence the above presentation can be simplified to one on just 2 generators (t and y) subject to at most $|\mathcal{R}| + |\mathcal{S}| + 12$ relations.

Proof: First, let G be the group defined by the given presentation, and observe that there exists a homomorphism from G to S_{m+n} under which

$$\begin{array}{lllllll} a & \mapsto & (m-1, m), & & v & \mapsto & (1, 2, \ldots, m), \\ b & \mapsto & (m+1, m+2), & & w & \mapsto & (m+1, m+2, \ldots, m+n), \\ t & \mapsto & (m, m+1), & & y & \mapsto & (1, 2, \ldots, m, m+1, m+2, \ldots, m+n), \end{array}$$

since these permutations satisfy the given relations.

Now in G, define $t_{m-1} = a$ and $t_{m-i-1} = v^i a v^{-i}$ for $1 \leqslant i \leqslant m-2$, and then also $t_m = t$, and $t_{m+1} = b$ and $t_{m+i+1} = w^{-i}bw^i$ for $1 \leqslant i \leqslant n-2$. We will show that these $m+n-1$ elements t_j satisfy the usual Coxeter relations for S_{m+n}, and that they generate G.

By the hypotheses on P, the elements t_i for $1 \leqslant i < m$ generate a subgroup isomorphic to S_m and therefore satisfy the usual Coxeter relations for S_m, and similarly, the elements t_{m+i} for $1 \leqslant i < n$ satisfy the the usual Coxeter relations for S_n. In particular, the elements t_i (for $1 \leqslant i < m+n$) satisfy the Coxeter relations $t_i^2 = 1$. Similarly, because we included the relations $(at)^3 = (tb)^3 = 1$, we have $(t_i t_{i+1})^3 = 1$ for $1 \leqslant i \leqslant m+n-2$.

Next, we will show that $[t_j, t_k] = (t_j t_k)^2 = 1$ whenever $1 \leqslant j < k-1 < m+n-1$. If $j < k < m$, then this follows from the presentation P; if $m < j < k$, then it follows from the presentation Q; and if $j < m < k$, then it follows from the four relations $[a, b] = [a, w] = [v, b] = [v, w] = 1$, because these imply that $\langle a, v\rangle$ commutes with $\langle b, w\rangle$. Similarly, because av and vav^{-1} generate a subgroup of index m in $\langle a, v\rangle \cong S_m$ containing the involutions $t_1, t_2, \ldots, t_{m-2}$, and wb and $w^{-1}bw$ generate a subgroup of index n in $\langle b, w\rangle \cong S_n$ containing $t_{m+2}, t_{m+3}, \ldots, t_{m+n-1}$, the four relations $[av, t] = [vav^{-1}, t] = [t, wb] = [t, w^{-1}bw] = 1$ imply that t centralises $\langle t_1, t_2, \ldots, t_{m-2}\rangle$ and $\langle t_{m+2}, t_{m+3}, \ldots, t_{m+n-1}\rangle$, and so we obtain also $[t_j, t_m] = 1$ for $1 \leqslant j \leqslant m-2$ and $[t_m, t_k] = 1$ for $m+2 \leqslant k \leqslant m+n-1$.

Hence the $m+n-1$ involutions t_i generate a subgroup satisfying the usual Coxeter relations for S_{m+n}.

The relations in P and Q imply that each of the elements of A or B is expressible as a word in these t_i, and the relation $y^{-1}wtv = 1$ then implies the same for y, and therefore the involutions t_i generate G.

Thus G is isomorphic to S_{m+n}, and the rest follows easily. □

In the above proof, t stands for the transposition $(m, m+1)$. Since conjugation by the long cycle $(1, 2, \ldots, n)$ is equivalent to the obvious cyclic relabelling of the points, the generator t can stand for any transposition of the form $(i, i+1)$, including the transposition $(1, 2)$. This observation will apply also to later theorems.

By ignoring the elements b and w, we have also the following in the case $n = 1$:

Theorem 3.2 *Let $P = \{A \,|\, \mathcal{R}\}$ be a presentation for S_m as in Theorem 3.1. Then*

$$\{\, A, t, y \,|\, \mathcal{R}, t^2, (at)^3, [av, t], [vav^{-1}, t], y^{-1}tv \,\}$$

is a presentation for S_{m+1} on a generating set that includes the elements y standing for the $(m+1)$-cycle $(1, 2, \ldots, m+1)$ and t standing for a transposition of the form $(i, i+1)$. In particular, this presentation has $|A| + 2$ generators and $|\mathcal{R}| + 5$ relations, and presentation length at most $\ell(P) + 27$, with the generators in A redundant.

A very similar construction without the gluing transposition can be used to prove the following:

Theorem 3.3 *Let $P = \{A \,|\, \mathcal{R}\}$ and $Q = \{B \,|\, \mathcal{S}\}$ be presentations for S_m and S_n, as in Theorem 3.1. Then*

$$\{\, A, B, y \,|\, \mathcal{R}, \mathcal{S}, (ab)^3, [a, wb], [a, w^{-1}bw], [v, wb], [v, w^{-1}bw], [av, b], [vav^{-1}, b], y^{-1}wv \,\}$$

is a presentation for S_{m+n-1} on a generating set that includes the elements y standing for the $(m+n-1)$-cycle $(1, 2, \ldots, m+n-1)$ and b standing for a transposition of the form $(i, i+1)$. In particular, this presentation has $|A| + |B| + 1$ generators and $|\mathcal{R}| + |\mathcal{S}| + 8$ relations, and presentation length at most $\ell(P) + \ell(Q) + 52$. Moreover, the elements of $A \cup B \setminus \{b\}$ are redundant generators.

Here the commutation relations ensure that $\langle a, v\rangle \cong S_m$ commutes with the subgroup $\langle wb, w^{-1}bw\rangle \cong S_{n-1}$ of index n in $\langle b, w\rangle$, and that $\langle b, w\rangle \cong S_n$ commutes with the subgroup $\langle av, vav^{-1}\rangle \cong S_{m-1}$ of index m in $\langle a, v\rangle$.

As a basis for induction, the constructions used in Theorems 3.1 and 3.3 are not the best possible, since doubling the degree roughly doubles both the number of relations and the presentation length. By taking $m = n$, however, we can express a copy of S_n as a conjugate of another, and thus do much better.

Theorem 3.4 *Let $P = \{A \,|\, \mathcal{R}\}$ be a presentation for the symmetric group S_n of degree $n \geqslant 3$, such that the generating set A contains elements x and w standing for the transposition $(1, 2)$ and the n-cycle $(1, 2, \ldots, n)$ respectively. Then*

$$\{\, A, y \,|\, \mathcal{R}, y^{2n}, (xy)^{2n-1}, [x, wy^{-1}], [w^2xw^{-1}, wy^{-1}], [x, y^n]^2, [x, y^{n-1}]^2 \,\}$$

is a presentation for S_{2n} on a generating set that includes the elements y standing for the $2n$-cycle $(1, 2, \ldots, 2n)$ and x standing for a transposition of the form $(i, i+1)$. In particular, this presentation has $|A| + 1$ generators and $|\mathcal{R}| + 6$ relations. Moreover, the elements of $A \setminus \{x\}$ are redundant generators.

Proof: Let G be the group defined by the given presentation, and observe that there exists a homomorphism from G to S_{2n} under which $x \mapsto (1, 2)$, $w \mapsto (1, 2, \ldots, n)$, and $y \mapsto (1, 2, \ldots, n, n+1, n+2, \ldots, 2n)$. In G, define $t_1 = x$ and $t_{i+1} = w^{-i}xw^i$ for $1 \leqslant i \leqslant n-2$, and then $t_n = y^{-(n-1)}xy^{n-1}$, and $t_{n+i} = y^{-n}t_iy^n$ for $1 \leqslant i \leqslant n$.

The elements t_i for $1 \leqslant i < n$ generate the same subgroup as x and w, so satisfy the usual Coxeter relations for transpositions generating S_n. Also the elements t_i for $1 \leqslant i \leqslant n-2$ generate the same subgroup as x and w^2xw^{-1} (standing for the stabiliser in S_n of the point n). Now the relations $[x, wy^{-1}] = [w^2xw^{-1}, wy^{-1}] = 1$ imply that wy^{-1} centralises $\langle x, w^2xw^{-1}\rangle = \langle t_1, t_2, \ldots, t_{n-2}\rangle$, and it follows that conjugation by y has the same effect as conjugation by w on the $n-2$ elements $t_1, t_2, \ldots, t_{n-2}$. In particular, $y^{-1}t_iy = w^{-1}t_iw = t_{i+1}$ for $1 \leqslant i \leqslant n-2$. But then also $y^{-1}t_{n-1}y = y^{-(n-1)}t_1y^{n-1} = t_n$ and $y^{-1}t_{n+i}y = y^{-(n+1)}t_iy^{n+1} = t_{n+i+1}$ for $1 \leqslant i \leqslant n$. From these and the relation $y^{2n} = 1$ it follows that conjugation by y cyclically permutes the $2n$ elements $t_1, t_2, \ldots, t_{2n}$.

The Coxeter relations $t_i^2 = 1$ and $(t_it_{i+1})^3 = 1$ now follow from conjugating the analogous relations in the subgroup $\langle\, t_1, t_2, \ldots, t_{n-1}\,\rangle = \langle x, w\rangle \cong S_n$ by powers of y. Similarly, so do the relations $[t_j, t_k] = (t_jt_k)^2 = 1$ whenever $2 \leqslant |j-k| \leqslant n-2$ (where $|j-k|$ is considered modulo $2n$). Also the relations $[x, y^{n-1}]^2 = [x, y^n]^2 = 1$ can be rewritten as $1 = [x, y^{-(n-1)}xy^{n-1}] = [t_1, t_n]$ and $1 = [x, y^{-n}xy^n] = [t_1, t_{n+1}]$, and conjugating these by powers of y gives all the remaining Coxeter relations of the form $[t_j, t_k] = (t_jt_k)^2 = 1$.

Thus the $2n-1$ involutions $t_1, t_2, \ldots, t_{2n-1}$ generate a subgroup satisfying the usual Coxeter relations for S_{2n}. The relations in $P = \{A \,|\, \mathcal{R}\}$ ensure that each of the elements of A is expressible as a word in the t_i, and in particular, $w = t_{n-1}t_{n-2}\ldots t_2t_1$. But similarly, from the relations $y^{2n} = (xy)^{2n-1} = 1$ it follows that $t_{2n-1}t_{2n-2}\ldots t_2t_1 =$

$(y^{-(2n-2)}t_1y^{2n-2})(y^{-(2n-3)}t_1y^{2n-3})\dots(y^{-1}t_1y)t_1 = y(yt_1)^{2n-1} = y(yx)^{2n-1} = y,$ and so the involutions t_i generate G. Thus G is isomorphic to S_{2n}, and the rest follows easily. □

The above construction can be viewed as taking a given presentation for S_n on generators $x \mapsto (1,2)$ and $y_n \mapsto (1,2,\dots,n)$, then duplicating it as a presentation for an isomorphic copy of S_n acting on $\{n+1, n+2, \dots, 2n\}$ with generators $t_{n+1} \mapsto (n+1, n+2)$ and $v_n \mapsto (n+1, n+2, \dots, 2n)$, 'gluing' the two presentations together by introducing the transposition $t_n \mapsto (n, n+1)$, and then replacing this element t_n by the $2n$-cycle $y = y_{2n} = v_n t_n y_n \mapsto (1,2,3,\dots,2n)$.

Again we can drop the gluing transposition, to obtain the following:

Theorem 3.5 *Let $P = \{A \,|\, \mathcal{R}\}$ be a presentation for S_n, as in Theorem 3.4. Then*

$$\{\, A, y \,|\, \mathcal{R}, y^{2n-1}, (xy)^{2n-2}, [x, wy^{-1}], [w^2xw^{-1}, wy^{-1}], [x, y^n]^2 \,\}$$

is a presentation for S_{2n-1} on a generating set that includes the elements y standing for the $(2n-1)$-cycle $(1,2,\dots,2n-1)$ and x standing for a transposition of the form $(i, i+1)$. In particular, this presentation has $|A|+1$ generators and $|\mathcal{R}|+5$ relations. Moreover, the elements of $A \setminus \{x\}$ are redundant generators.

Here the second (isomorphic) copy of S_n acts on $\{n, n+1, \dots, 2n-1\}$ with generators $t_n \mapsto (n, n+1)$ and $z_n \mapsto (n, n+1, \dots, 2n-1)$, and we introduce the $(2n-1)$-cycle $y = y_{2n-1} = z_n y_n \mapsto (1,2,3,\dots,2n-1)$. The relations for the second copy of S_n may be taken as conjugates under y^{n-1} of the relations in the starting presentation for S_n, and so can be eliminated (once y is introduced).

Proof: As for Theorem 3.4, but with the Coxeter generators defined by $t_1 = x$ and $t_{i+1} = w^{-i}xw^i$ for $1 \leqslant i \leqslant n-2$, and $t_{n+i-1} = y^{-(n-1)}t_i y^{n-1}$ for $1 \leqslant i \leqslant n$. Here conjugation by y cyclically permutes the $2n-1$ elements $t_1, t_2, \dots, t_{2n-1}$, and we have $t_{2n-2}t_{2n-3}\dots t_2t_1 = y(yt_1)^{2n-2} = y(yx)^{2n-2} = y$. Only one additional relation not already provided by the presentation $P = \{A \,|\, \mathcal{R}\}$, such as $[x, y^n]^2 = 1$ (or equivalently, $[x, y^{n-1}]^2 = 1$), is required to obtain all the Coxeter relations in this case. □

Together, Theorems 3.4 and 3.5 can be used to produce short presentations for symmetric groups of arbitrarily large degree, whether odd or even. For instance, a presentation for S_{439} can be obtained by inductively finding presentations for S_m for all m in the sequence 4, 7, 14, 28, 55, 110, 220, 439. As building blocks for this general process, we need short presentations for S_n involving $(1,2)$ and an n-cycle for just two cases: S_3 and S_4. Examples are provided in Section 8.

Theorem 3.6 *For every positive integer n, the symmetric group S_n has a presentation in which the number of relations is $O(\log n)$ and the presentation length is $O((\log n)^2)$.*

Proof: For any n, a suitable presentation may be obtained by induction using Theorems 3.4 and 3.5 as appropriate. If the number of generators in the starting presentation for S_n is $g(n)$, then we have the recurrence $g(2n-1) = g(2n) = g(n) + 1$,

which implies that g is a logarithmic function. Similarly, if the number of relators in the starting presentation for S_n is $f(n)$, then we have the recurrence property $f(n) \leqslant f(2n-1) \leqslant f(2n) \leqslant f(n)+6$, which implies that f is logarithmic. Each relator has fixed syllable length, however, and each syllable (such as y^n or y^{n-1} or $(xy)^{2n-1}$ or y^{2n-2}) can be rewritten as a number of relators of length $O(\log n)$, so the total presentation length is quadratic in $\log n$. □

4 Building presentations for S_{mn} from S_m and S_n

The construction involved in Theorem 3.4 started with a presentation for S_n in terms of generators x and w, and duplicated this using an involution y^n, in order to obtain a presentation for S_{2n}. In the resulting group, the elements x, w and y^n generate a subgroup isomorphic to the wreath product $S_n \wr S_2$.

We can generalise this to a construction that uses the wreath product $S_m \wr S_n$ to help produce a short presentation for S_{mn}, whenever $m, n \geqslant 3$. In particular, this gives a short presentation for S_{m^2} for all $m \geqslant 3$. We can then obtain a short presentation for S_n for any n.

Theorem 4.1 *Let $P = \{A \,|\, \mathcal{R}\}$ and $Q = \{B \,|\, \mathcal{S}\}$ be presentations for the symmetric groups S_m and S_n of degrees $m, n \geqslant 3$, such that the generating set A for S_m contains elements a and v standing for the transposition $(1,2)$ and the m-cycle $(1,2,\ldots,m)$ respectively, and the generating set B for S_n contains elements b and w standing for the transposition $(1,2)$ and the n-cycle $(1,2,\ldots,n)$ respectively. Then*

$$\begin{aligned}\{\, &A, B, t, y \,|\, \mathcal{R}, \mathcal{S}, t^2, b^{-1}(v^{-1}tw^{-1}v^{-1}w)^m, w^{-1}y^m, y^{-1}wv(wtv)^{n-1}, y^{-1}vyav^{-1}t,\\ &(v^2av^{-2}t)^3, (tw^{-1}aw)^3, [a,t], [v^2av^{-1},t], [a,vy^{-1}], yty^{-1}v^2av^{-2}, y^{-1}tyw^{-1}aw,\\ &[a,w^{-1}aw], [a,w^{-1}vw], [v,w^{-1}aw], [v,w^{-1}vw], [a,wb], [a,w^{-1}bw], [v,wb], [v,w^{-1}bw]\,\}\end{aligned}$$

is a presentation for S_{mn} on a generating set that includes the elements y standing for the mn-cycle $(1,2,\ldots,mn)$ and t standing for a transposition of the form $(i,i+1)$. In particular, this presentation has $|A|+|B|+2$ generators and $|\mathcal{R}|+|\mathcal{S}|+20$ relations. Moreover, the generators in $A \cup B$ are redundant.

Proof: First, let G be the group defined by the given presentation, and observe that there exists a homomorphism from G to S_{mn} under which

$$\begin{aligned}
a &\mapsto (1,2),\\
v &\mapsto (1,2,\ldots,m),\\
b &\mapsto (1,m{+}1)(2,m{+}2)\ldots(m,2m),\\
w &\mapsto (1,m{+}1,\ldots,(n{-}1)m{+}1)(2,m{+}2,\ldots,(n{-}1)m{+}2)\ldots(m,2m,\ldots,mn),\\
t &\mapsto (m,m{+}1),\\
y &\mapsto (1,2,\ldots,m,m{+}1,m{+}2,\ldots,2m,2m{+}1,2m{+}2,\ldots,mn),
\end{aligned}$$

since these permutations satisfy the given relations. Note that the images of a, v, b and w generate a subgroup isomorphic to the wreath product $S_m \wr S_n$; in particular,

the images of a and v generate a subgroup isomorphic to S_m, and conjugates of this subgroup under powers of w generate a direct product $(S_m)^n$ of n copies of S_m.

Now in G, define $t_1 = a$ and $t_{i+1} = v^{-i}av^i$ for $1 \leqslant i \leqslant m-2$, and then $t_m = t$, and also $t_{im+j} = w^{-i}t_jw^i$ for $1 \leqslant i < n$ and $1 \leqslant j \leqslant m$. As in previous proofs, we will show that the $mn-1$ elements $t_1, t_2, \ldots, t_{mn-1}$ satisfy the usual Coxeter relations for S_{mn}, and that they generate G. Note that the element t_{mn} has been defined here also, for later use.

By the hypotheses on P, the elements $t_1, t_2, \ldots, t_{m-1}$ generate a subgroup A_1 isomorphic to S_m and satisfy the usual Coxeter relations for S_m. Conjugating by w^i, we see the same is true for the subgroup A_{i+1} generated by $t_{im+1}, t_{im+2}, \ldots, t_{im+m-1}$, for $1 \leqslant i < n$. Moreover, the four commutation relations $[a, w^{-1}aw] = [a, w^{-1}vw] = [v, w^{-1}aw] = [v, w^{-1}vw] = 1$ ensure that $A_1 = \langle t_1, t_2, \ldots, t_{m-1}\rangle = \langle a, v\rangle$ commutes with its conjugate under w, namely $\langle t_{m+1}, t_{m+2}, \ldots, t_{2m-1}\rangle = A_2$. The next four commutation relations $[a, wb] = [a, w^{-1}bw] = [v, wb] = [v, w^{-1}bw] = 1$ imply that A_1 is centralised by $B_1 = \langle wb, w^{-1}bw\rangle$, which has index n in $B = \langle b, w\rangle \cong S_n$, and it follows that B permutes the subgroups $A_1, A_2, \ldots, A_n$ by conjugation according to the natural action of S_n on the index set $\{1, 2, \ldots, n\}$. In particular, as this action is doubly-transitive, we deduce that A_i commutes with its conjugate A_j whenever $1 \leqslant i < j \leqslant n$, and that the subgroup of G generated by a, v, b and w is isomorphic to the wreath product $S_m \wr S_n$.

It is easy to see that the elements t_i satisfy $t_i^2 = 1$ for $1 \leqslant i < mn$. Similarly, $(t_it_{i+1})^3 = 1$ for $1 \leqslant i \leqslant m-2$, and the relations $(v^2av^{-2}t)^3 = (tw^{-1}aw)^3 = 1$ give $(t_{m-1}t_m)^3 = (t_mt_{m+1})^3 = 1$, and then conjugation of these by powers of w gives all remaining relations of the form $(t_jt_{j+1})^3 = 1$, namely for $m+1 \leqslant j \leqslant mn-2$.

Next, we will show that $[t_j, t_k] = (t_jt_k)^2 = 1$ whenever $1 < j+1 < k < mn$. If $j < k < m$, then this follows from the presentation P; and if $im < j < k < (i{+}1)m$ for some $i > 0$, then it follows from P after conjugation by w^i. If t_j and t_k lie in different B-conjugates of the subgroup A_1, then $[t_j, t_k] = 1$ because those conjugates commute with each other. Also the two commutation relations $[a, t] = [v^2av^{-1}, t] = 1$ ensure that $t_m = t$ commutes with all elements of $\langle a, v^2av^{-1}\rangle = \langle t_1, t_2, \ldots t_{m-2}\rangle$.

To obtain the rest, we show that the relations $[a, vy^{-1}] = y^{-1}vyav^{-1}t = yty^{-1}v^2av^{-2} = y^{-1}tyw^{-1}aw = 1$ imply that y conjugates each element in the sequence $t_1, t_2, \ldots, t_{mn}$ to its successor. First, $y^{-1}t_1y = y^{-1}ay = v^{-1}av = t_2$ since vy^{-1} centralises a, and then we find by induction that

$$\begin{aligned} y^{-1}t_{i+1}y &= y^{-1}v^{-1}t_ivy = (y^{-1}vy)^{-1}(y^{-1}t_iy)(y^{-1}vy) \\ &= (av^{-1}t)t_{i+1}(tva) = av^{-1}t_{i+1}va = at_{i+2}a = t_{i+2} \end{aligned}$$

for $1 \leqslant i \leqslant m-3$, since t commutes with t_{i+1}, and $a = t_1$ commutes with t_{i+2}. Next $y^{-1}t_{m-1}y = y^{-1}v^2av^{-2}y = t = t_m$, and $y^{-1}t_my = y^{-1}ty = w^{-1}aw = w^{-1}t_1w = t_{m+1}$, and finally because $w = y^m$ is centralised by y (which has order mn), we find

$$y^{-1}t_{im+j}y = y^{-1}w^{-i}t_jw^iy = w^{-i}y^{-1}t_jyw^i = w^{-i}t_{j+1}w^i = t_{im+j+1}$$

for $1 \leqslant i < n$ and $1 \leqslant j \leqslant m$.

Now since the known relations of the form $[t_j, t_k] = 1$ cover all possible index differences $k - j$ modulo mn (other than ± 1 mod mn), conjugation by powers of y gives $[t_j, t_k] = 1$ whenever $1 < j + 1 < k < mn$.

Hence the $mn - 1$ involutions $t_1, t_2, \ldots, t_{mn-1}$ generate a subgroup satisfying the usual Coxeter relations for S_{mn}.

Finally, we show that each of the given generators for G can be expressed as a word in these t_i. The relations in P give this for each element of the set A. In particular, $a = t_1$ and $v = (v^{-(m-2)}av^{m-2})(v^{-(m-3)}av^{m-3})\ldots(v^{-1}av)a = t_{m-1}t_{m-2}\ldots t_2t_1$, and it follows that $w^{-i}vw^i = t_{im+m-1}t_{im+m-2}\ldots t_{im+2}t_{im+1}$ for $1 \leqslant i < n$. Similarly $t = t_m$, and of course $w^{-i}tw^i = t_{(i+1)m}$ for $1 \leqslant i \leqslant n-2$. From the relation $y = wv(wtv)^{n-1}$ we can now deduce that

$$\begin{aligned} y &= (w^{-(n-1)}vw^{n-1})(w^{-(n-2)}tw^{n-2})(w^{-(n-2)}vw^{n-2})\ldots(w^{-1}tw)(w^{-1}vw)tv \\ &= (t_{mn-1}\ldots t_{(n-1)m+1})t_{(n-1)m}\ldots(t_{2m-1}\ldots t_{m+2}t_{m+1})t_m(t_{m-1}\ldots t_2t_1) \\ &= t_{mn-1}t_{mn-2}\ldots t_2t_1. \end{aligned}$$

Next, from the relations $w = y^m$ and $b = (v^{-1}tw^{-1}v^{-1}w)^m$ it follows that both b and w are expressible as words in the t_i as well, and the same then holds for all the redundant elements of the set B (which are expressible in terms of b and w).

Hence the involutions t_i generate G, which is therefore isomorphic to S_{mn}, and the rest follows easily. □

Corollary 4.2 *For all $m \geqslant 3$, if the symmetric group S_m has a presentation involving the long cycle $(1, 2, 3, \ldots, m)$ and a transposition of the form $(i, i+1)$, and of length at most $K \log m \log\log m$ where K is any constant $\geqslant 100$, then S_{m^2} also has such a presentation, of length at most $K \log m^2 \log\log m^2$.*

Proof: Taking $n = m$ in Theorem 4.1, we find that if the presentation for S_m has length $\ell(P)$, then a presentation can be constructed for S_{m^2} with length bounded above by $2\ell(P) + 3\log m + c$ for some constant $c \leqslant 148$. Now $(K-3)\log m \geqslant 97 \log 3 > c$, so it follows that if $\ell(P) \leqslant K \log m \log\log m$, then

$$\begin{aligned} 2\ell(P) + 3\log m + c &\leqslant 2\ell(P) + K\log m \\ &\leqslant 2K \log m \log\log m + K \log m \\ &\leqslant 2K \log m(\log\log m + \log 2) \\ &= K \log m^2 \log(2\log m) \\ &= K \log m^2 \log\log m^2 \end{aligned}$$

as required. □

It remains to construct a presentation for S_n from these building blocks. To that end we require the following lemma.

Lemma 4.3 *Let $f : \mathbb{N} \to \mathbb{N}$ satisfy the following, for some positive constants u and v:*

(i) $f(a+b) \leqslant f(a) + f(b) + u\log(ab)$ *for all* $a, b \in \mathbb{N}$, *and*

(ii) $f(ab) \leqslant f(a) + f(b) + v\log(ab)$ *for all* $a, b \in \mathbb{N}$.

Then $f(n) < k\log^\alpha n \log\log n$ *for* n *sufficiently large, and some constant* k, *where* $\alpha = \log_2(1+\sqrt{2}) \approx 1 \cdot 27155$.

PROOF: For $n \in \mathbb{N}$, let $n_1 \in \mathbb{N}$ be minimal subject to $n_1^2 \geqslant n$, and let n_2 be minimal subject to $n_1^2 - n_2^2 \leqslant n$. Then clearly $n_1 < \sqrt{n}+1$ and $n_2 < 1+\sqrt{2(n_1-1)} < 1+\sqrt{2}\sqrt[4]{n}$. Further $n = (n_1+n_2)(n_1-n_2)+n_3$ where $0 \leqslant n_3 < 2n_2$.

Now consider the following inductive hypothesis: for some $k > 0$ and for some sufficiently large n_0 and for some $n > n_0$, suppose that $f(a) < k\log^\alpha a \log\log a$ for every integer a in the range $n_0 < a < n$.

It is easy to see that

$$\log^\alpha(n_1+n_2)\log\log(n_1+n_2) = \log^\alpha\sqrt{n}\log\log\sqrt{n} + o(1)$$

and similarly when n_1+n_2 is replaced by n_1-n_2. Hence it follows that

$$\begin{aligned} f(n) &\leqslant f(n_1+n_2) + f(n_1-n_2) + f(n_3) + (5u/4+v)\log n + O(1) \\ &\leqslant k(2\log^\alpha\sqrt{n}\log\log\sqrt{n} + \log^\alpha(2^{3/2}n^{1/4})\log\log(2^{3/2}n^{1/4})) \\ &\quad + (5u/4+v)\log n + c \quad \text{for some absolute constant } c \\ &\leqslant k(2\log^\alpha\sqrt{n}\log\log\sqrt{n} + \log^\alpha n^{1/4}\log\log\sqrt{n}) \end{aligned}$$

for all $n > N$, where N is determined by u, v, k. By definition of α, the last expression is $k\log^\alpha n\log\log n$. The least value of N that satisfies this condition decreases monotonically with k, and hence may be regarded as being determined by u, v alone. Given N, we may take k to be large enough to ensure that $f(a) < k\log^\alpha a\log\log a$ for all a such that $a \leqslant n$ and $\log\log a > 0$. The above argument then proves the lemma inductively. □

Note that the proof can also be organised to demonstrate that for any $k > 0$, $f(a) < k\log^\alpha a\log\log a$ for all sufficiently large values of a.

This leads to the following result.

Theorem 4.4 *For all* $n \geqslant 3$, *the symmetric group* S_n *has a presentation of length* $O(\log^\alpha n\log\log n)$ *on a generating set that includes elements* x *and* y *standing for the transposition* $(1,2)$ *and the* n*-cycle* $(1,2,\ldots,n)$, *and* α *is as in the above lemma.*

PROOF: Let $f(n)$ be the minimum length of a presentation of S_n on a generating set as in the theorem. Then by Theorem 3.1 it follows that $f(a+b) \leqslant f(a)+f(b)+u$ for some constant u, and by Theorem 4.1 it follows that $f(ab) \leqslant f(a)+f(b)+v\log(ab)$ for some constant v. The result then follows from the previous lemma. □

5 Presentations for S_n with a bounded number of relations

In this section we combine certain presentations for 2-dimensional projective general linear groups over prime fields, exploiting the 3-transitive permutation representations of these groups, in order to build presentations for symmetric groups with a bounded number of relations.

For every odd prime p, the group $\mathrm{PGL}(2,p)$ has order $p(p^2-1)$, and is 3-transitive in its natural action on the $p+1$ points of the projective line over the field $\mathrm{GF}(p)$. The stabiliser of any point is a Frobenius group of order $p(p-1)$, complemented in $\mathrm{PGL}(2,p)$ by a cyclic subgroup of order $p+1$ (any generator of which induces a single cycle of length $p+1$, called a *Singer cycle*), and the stabiliser of any unordered pair of points is a dihedral subgroup of order $2(p-1)$. Further, $\mathrm{PGL}(2,p)$ itself can be generated by two elements, such as a Singer cycle and any element of order p or $p-1$.

Lemma 5.1 *Let p be any odd prime, and α any primitive element of the field $\mathrm{GF}(p)$. Then for some integer j satisfying $1 \leqslant j \leqslant p-2$,*

$$\{\, Q, R, S, T \mid S^p, T^2, (ST)^3, (S^4TS^{(p+1)/2}T)^2, Q^{p-1}, Q^{-1}SQS^{-\alpha}, (QT)^2, \\ Q^{-1}TQTS^{-1/\alpha}TS^{-\alpha}TS^{-1/\alpha}T, R^{-1}QTS^j, SRS^{(1/\alpha)-j}RS^{1-j}T \,\},$$

is a 4-generator 10-relator presentation for the group $\mathrm{PGL}(2,p)$. The three generators S, T and Q may be taken as standing respectively for the linear fractional transformations $z \mapsto z+1$, $z \mapsto -1/z$ and $z \mapsto \alpha z$, in which case the first and third generate a Frobenius subgroup of order $p(p-1)$ fixing the point ∞, the second and third generate a dihedral subgroup of order $2(p-1)$ stabilising the pair $\{0,\infty\}$, and the fourth generator R stands for a Singer cycle that takes ∞ to 0. Moreover, in this presentation the generators T and Q are redundant, and their elimination gives rise to an 8-relator presentation for $\mathrm{PGL}(2,p)$ on the two generators S and R.

Proof: A presentation for $\mathrm{PGL}(2,p)$ is given in Section 7.5 of [7] in terms of the three generators S, T and Q, subject to the first eight of the relations given, and standing for elements with the given properties. Then for any j, the pre-image in $\mathrm{GL}(2,q)$ of the element QTS^j can be represented by a 2×2 matrix $\pm \begin{pmatrix} \alpha j & \alpha \\ -1 & 0 \end{pmatrix}$ with trace $\pm\alpha j$ and determinant α, and so the value of j can be chosen to ensure that $R = QTS^j$ lies in a conjugacy class of elements of order $p+1$. With R so chosen, we find that $SRS^{(1/\alpha)-j}RS^{1-j} = T$, and thus we obtain the above presentation. The last two relations can be used to eliminate T and Q, giving a 2-generator 8-relator presentation on the generating set $\{R,S\}$. □

We now exploit the existence of such short presentations and the 3-transitivity of $\mathrm{PGL}(2,p)$, to give the following:

Theorem 5.2 *For odd primes p and q, let $P = \{a, g, c, e \,|\, \mathcal{R}\}$ be a presentation for* $\mathrm{PGL}(2, p)$*, such that in the natural action on the projective line over* $\mathrm{GF}(p)$*, the generator g stands for a Singer cycle taking ∞ to 0, the generator a stands for an element of order p fixing ∞, the subgroup generated by a and c stands for the stabiliser of the point ∞, and the subgroup generated by c and e stands for the stabiliser of the pair $\{0, \infty\}$, and let $Q = \{b, h, d, f \,|\, \mathcal{S}\}$ be an analogous presentation for* $\mathrm{PGL}(2, q)$*. Then there exist words v and w in the letters t and y such that*

$$\begin{aligned}\{\, a, b, c, d, e, f, g, h, t, y \,|\, \mathcal{R}, \mathcal{S}, t^2, y^{-1}htg^{-1}, a^{-1}v, g(yt)^p y^{-p}, b^{-1}w, h^{-1}y^{-q}(ty)^q,\\ [g,h], [a,t], [c,t], [t,b], [t,d], [c, yty^{-1}], [e, yty^{-1}], [y^{-1}ty, d], [y^{-1}ty, f],\\ [t,y]^3, [t, y^{-2}ty^2], [t, y^{-1}gy^2], [t, yhy^{-2}], [g, y^{-1}ty], [yty^{-1}, h]\,\}\end{aligned}$$

is a presentation for S_{p+q+2} on a generating set that includes the elements y standing for the $(p{+}q{+}2)$-cycle $(1, 2, \ldots, p{+}q{+}2)$ and t standing for a transposition of the form $(i, i{+}1)$. In particular, this presentation has 10 generators and $|\mathcal{R}| + |\mathcal{S}| + 21$ relations. Also the generators a, b, c, d, e, f, g and h are redundant.

Proof: First, let G be the group defined by the given presentation, and label the points of the projective line over $\mathrm{GF}(p)$ as $1, 2, \ldots, p, p{+}1$ such that p labels 0, and $p{+}1$ labels ∞, and similarly, label the points of the projective line over $\mathrm{GF}(q)$ as $p{+}2, p{+}3, \ldots, p{+}q{+}1, p{+}q{+}2$ such that $p{+}2$ labels ∞, and $p{+}3$ labels 0. This labelling and the words v and w can be chosen in such a way that it gives a homomorphism from G to S_{p+q+2} under which

- $$\begin{aligned} g &\mapsto (p{+}1, p, \ldots, 2, 1),\\ h &\mapsto (p{+}2, p{+}3, \ldots, p{+}q{+}1, p{+}q{+}2),\\ t &\mapsto (p{+}1, p{+}2),\\ y &\mapsto (1, 2, \ldots, p, p{+}1, p{+}2, p{+}3, \ldots, p{+}q{+}1, p{+}q{+}2),\end{aligned}$$
- the subgroup generated by a and g is transitive on $\{1, 2, \ldots, p, p{+}1\}$ and fixes the other $q + 1$ points $p{+}2, p{+}3, \ldots, p{+}q{+}1, p{+}q{+}2$,
- the subgroup generated by b and h is transitive on $\{p{+}2, p{+}3, \ldots, p{+}q{+}1, p{+}q{+}2\}$ and fixes the other $p + 1$ points $1, 2, \ldots, p, p{+}1$,
- the subgroups $\langle a, c\rangle$ and $\langle b, d\rangle$ fix $p{+}1$ and $p{+}2$ respectively, and
- the subgroups $\langle c, e\rangle$ and $\langle d, f\rangle$ preserve the pairs $\{p, p+1\}$ and $\{p+2, p+3\}$ respectively.

The words v and w exist because the images of t and y generate S_{p+q+2}, and in fact they can be found by expressing the images of a and c as words in the transpositions $(i, i{+}1)$ for $1 \leqslant i \leqslant p{+}q{+}1$, which are images of conjugates of t by powers of y. Observe also that the generators c and e in the presentation P for $\mathrm{PGL}(2, p)$ are redundant, as are the generators d and f in the presentation Q for $\mathrm{PGL}(2, q)$.

Now in G, define $t_{p+1} = t$, and $t_p = yty^{-1}$ and $t_{p+2} = y^{-1}ty$, and then $t_{p-i} = g^{-i}t_pg^i$ for $1 \leqslant i \leqslant p-1$, and $t_{p+2+j} = h^{-j}t_{p+2}h^j$ for $1 \leqslant j \leqslant q-1$. As earlier, we will show that the elements $t_1, t_2, \ldots, t_{p+q+1}$ satisfy the usual Coxeter relations for S_{p+q+2}, and that they generate G.

To do that, we make some key observations about conjugates of the element t_p by elements of $A = \langle a, g\rangle$. One is that $t_p = yty^{-1}$ is centralised by c and e, which generate a dihedral subgroup of order $2(p-1)$ standing for the stabiliser of an unordered pair $\{p, p{+}1\}$ of points in a known 3-transitive action of A of degree $p+1$. In particular, the number of conjugates of t_p by elements of A is exactly $p(p+1)/2$, and there is a one-to-one correspondence between these conjugates and the unordered pairs of distinct points in the 3-transitive action of A. Also a and c generate a subgroup of order $p(p-1)$ that stands for the stabiliser in A of the point $p+1$, and by 3-transitivity, this subgroup has two orbits on pairs, of lengths p and $p(p-1)/2$. Moreover, this subgroup is a complement to the subgroup generated by g (standing for a Singer cycle that moves $p{+}1$ to p), and $\{p, p{+}1\}^{g^i}$ lies in the long orbit of $\langle a, c\rangle$ on pairs for $1 \leqslant i \leqslant p$ (while $\{p, p{+}1\}$ lies in the short orbit). It follows that $t_1, t_2, \ldots, t_{p-1}$ all lie in the same equivalence class under conjugation by $\langle a, c\rangle$, of size $p(p-1)/2$, but not containing t_p.

By the analogous argument, $t_{p+3}, t_{p+4}, \ldots, t_{p+q+1}$ all lie in the same equivalence class under conjugation by $\langle b, d\rangle$, because $t_{p+2} = y^{-1}ty$ is centralised by the dihedral subgroup of order $2(q-1)$ generated by d and f in $B = \langle b, h\rangle$.

Next, we consider conjugation by y and t. The relation $[t, y^{-1}gy^2] = 1$ gives $y^2t_{p+1}y^{-2} = y^2ty^{-2} = (y^{-1}g)^{-1}t(y^{-1}g) = g^{-1}yty^{-1}g = t_{p-1}$; similarly, $[t, yhy^{-2}] = 1$ gives $y^{-2}t_{p+1}y^2 = y^{-2}ty^2 = (yh)^{-1}t(yh) = h^{-1}y^{-1}tyh = t_{p+3}$. The relation $[t, y^{-2}ty^2] = 1$ now gives $[t_{p-1}, t_{p+1}] = [y^2ty^{-2}, t] = 1$, and by our observations above (and since a and c both centralise t), conjugating this by elements of $\langle a, c\rangle$ gives $[t_i, t_{p+1}] = 1$ for $1 \leqslant i \leqslant p-1$. Similarly, we have $[t_{p+1}, t_{p+3}] = [t, y^{-2}ty^2] = 1$, and conjugating this by elements of $\langle b, d\rangle$ gives $[t_{p+1}, t_j] = 1$ for $p+3 \leqslant i \leqslant p+q+1$.

More easily, the relation $[yty^{-1}, h] = 1$ gives $[t_p, h] = 1$, and conjugation by powers of g (which all centralise h since $[g, h] = 1$) then gives $[t_i, h] = 1$ for $1 \leqslant i \leqslant p$; and similarly, conjugation of the relation $[g, y^{-1}ty] = 1$ by powers of h gives $[g, t_j] = 1$ for $p+2 \leqslant j \leqslant p+q+1$. Because $y = htg^{-1}$, it now follows that

$$y^{-1}t_iy = gth^{-1}t_ihtg^{-1} = gtt_itg^{-1} = gt_ig^{-1} = t_{i+1} \quad \text{for } 1 \leqslant i \leqslant p-1, \quad \text{while}$$
$$y^{-1}t_jy = gth^{-1}t_jhtg^{-1} = gtt_{j+1}tg^{-1} = gt_{j+1}g^{-1} = t_{j+1} \quad \text{for } p+2 \leqslant j \leqslant p+q.$$

As also $y^{-1}t_py = t_{p+1}$ and $y^{-1}t_{p+1}y = t_{p+2}$, we have $y^{-1}t_ky = t_{k+1}$ for $1 \leqslant k \leqslant p{+}q{+}1$.

Now conjugation of the relation $t^2 = 1$ gives $t_i^2 = 1$ for all i. Similarly, the relation $[t, y]^3 = 1$ gives $(t_{p+1}t_{p+2})^3 = (ty^{-1}ty)^3 = 1$, and then conjugation of this by powers of y gives $(t_it_{i+1})^3 = 1$ for $1 \leqslant i \leqslant p+q+1$. As we saw above, the relation $[t, y^{-2}ty^2] = 1$ gives $[t_{p+1}, t_{p+3}] = [t, t_{p+3}] = 1$, and conjugation of this by powers of y gives $[t_i, t_{i+2}] = 1$ for $1 \leqslant i \leqslant p+q$. In particular, we have $[yty^{-1}, y^{-1}ty] = [t_p, t_{p+2}] = 1$. Conjugation of this by powers of g gives $[t_i, t_{p+2}] = 1$ for $1 \leqslant i \leqslant p$, and further conjugation by powers of h gives $[t_i, t_j] = 1$ for $1 \leqslant i < p+1 < j \leqslant p+q+1$. Because the known relations of the form $[t_j, t_k] = 1$ now cover all possible index differences $k-j$ (other than ±1) mod $p{+}q{+}2$, conjugation by powers of y gives $[t_j, t_k] = 1$ whenever $1 < j{+}1 < k < p{+}q{+}2$.

Hence the involutions $t_1, t_2, \ldots, t_{p+q+1}$ generate a subgroup satisfying the usual Coxeter relations for S_{p+q+2}.

The relations $g(yt)^p y^{-p} = 1$ and $h^{-1}y^{-q}(ty)^q = 1$ give

$$g^{-1} = (yty^{-1})(y^2ty^{-2})\ldots(y^{p-1}ty^{-(p-1)})(y^pty^{-p}) = t_pt_{p-1}\ldots t_2t_1, \quad \text{and}$$

$$h = (y^{-q}ty^q)(y^{-(q-1)}ty^{q-1})\ldots(y^{-2}ty^2)(y^{-1}ty) = t_{p+q+1}t_{p+q}\ldots t_{p+3}t_{p+2}.$$

These, together with the two relations expressing a and c as words in the conjugates of t by powers of y, and the relation $y = htg^{-1}$, show that each of the given generators for G can be expressed as a word in the elements t_i.

Hence the involutions t_i generate G, which is therefore isomorphic to S_{p+q+2}, and the rest follows easily. □

Again, a very similar construction without the gluing transposition $(p+1, p+2)$ can be used to obtain the following:

Theorem 5.3 *Let $P = \{a, g, c, e \,|\, \mathcal{R}\}$ and $Q = \{b, h, d, f \,|\, \mathcal{S}\}$ be presentations for* $\mathrm{PGL}(2,p)$ *and* $\mathrm{PGL}(2,q)$ *as in Theorem 5.2. Then there exist words v and w in the letters t and y such that*

$$\{\, a, b, c, d, e, f, g, h, t, y \,|\, \mathcal{R}, \mathcal{S}, t^2, y^{-1}hg^{-1}, a^{-1}v, g(ty)^py^{-p}, b^{-1}w, h^{-1}y^{-q}(ty)^q,$$
$$[a,h], [c,h], [g,b], [g,d], [c,t], [e,t], [y^{-1}ty, d], [y^{-1}ty, f],$$
$$[t,y]^3, [t, y^{-2}ty^2], [t, gy], [t, yhy^{-2}], [g, y^{-2}ty^2], [yty^{-1}, h] \,\}$$

is a presentation for S_{p+q+1} on a generating set that includes the elements y standing for the $(p+q+1)$-cycle $(1, 2, \ldots, p+q+1)$ and t standing for a transposition of the form $(i, i+1)$. In particular, this presentation has 10 *generators and $|\mathcal{R}| + |\mathcal{S}| + 20$ relations. Also the generators a, b, c, d, e, f, g and h are redundant.*

Proof: Here the Coxeter generators are $t_p = t$ and $t_{p+1} = y^{-1}ty$, plus $t_{p-i} = g^{-i}t_pg^i$ for $1 \leqslant i \leqslant p-1$, and $t_{p+1+j} = h^{-j}t_{p+1}h^j$ for $1 \leqslant j \leqslant q-1$. The relations $[c,t] = [e,t] = 1$ ensure that t has exactly $p(p+1)/2$ conjugates under elements of $A = \langle a, g\rangle$, and hence that $t_1, t_2, \ldots, t_{p-1}$ all lie in the same equivalence class under conjugation by $\langle a, c\rangle$, and similarly, the relations $[y^{-1}ty, d] = [y^{-1}ty, f] = 1$ ensure that $t_{p+2}, t_{p+3}, \ldots, t_{p+q}$ are all equivalent under conjugation by $\langle b, d\rangle$. The relation $[t, gy] = 1$ gives $yt_py^{-1} = g^{-1}t_pg = t_{p-1}$, and similarly, $[t, yhy^{-2}] = 1$ gives $y^{-2}t_py^2 = t_{p+2}$. The relation $[yty^{-1}, h] = 1$ gives $[t_{p-1}, h] = 1$, and conjugation by elements of $\langle a, c\rangle$ (which all centralise h since $[a,h] = [c,h] = 1$) then gives $[t_i, h] = 1$ for $1 \leqslant i \leqslant p$; and similarly, conjugation of the relation $[g, y^{-2}ty^2] = 1$ by elements of $\langle b, d\rangle$ gives $[g, t_j] = 1$ for $p+2 \leqslant j \leqslant p+q$. These observations imply that $y^{-1}t_ky = t_{k+1}$ for $1 \leqslant k \leqslant p+q$, and the rest follows as previously. □

The presentations from Lemma 5.1 may be used as input into Theorem 5.2 or 5.3, and then eliminating the redundant generators a, b, c, d, e, f, g and h from the result (using one relation each), we find the following:

Corollary 5.4 *Let n be any positive integer such that either $n-1$ or $n-2$ is the sum of odd primes p and q. Then the symmetric group S_n has a defining presentation with generators $x = (1,2)$ and $y = (1,2,3,\ldots,n)$ subject to at most 33 relations.*

If the Goldbach conjecture is true, then every integer $n \geqslant 7$ has this property. But independent of the veracity of the Goldbach conjecture, we have the following:

Corollary 5.5 *For every positive integer n, the symmetric group S_n has a defining presentation with generators $x = (1,2)$ and $y = (1,2,3,\ldots,n)$ subject to at most 123 relations, or at most 78 relations if n is sufficiently large.*

PROOF: By a theorem of Vinogradov [19], every sufficiently large odd integer is expressible as a sum of three primes, and it follows that every sufficiently large even integer is expressible as a sum of four odd primes. Now suppose $n-2$ or $n-1$ is so expressible, say as $p+q+r+s$. Then by Corollary 5.4 there exist 2-generator 33-relator presentations for S_{p+q+1} and S_{r+s+1}, which can be taken as input to Theorem 3.1 or 3.3, to produce 2-generator presentations for S_n ($= S_{p+q+r+s+2}$ or $S_{p+q+r+s+1}$), with at most $33+33+12 = 78$ relations. Alternatively, we use the theorem of Ramaré [15] that gives every even $n > 1$ as a sum of at most six primes, and take the analogous approach, obtaining a presentation with at most $78+33+12 = 123$ relations. □

Unfortunately the lengths of the presentations in the last two results are difficult to measure, because of the words v and w used in Theorems 5.2 and 5.3. This problem is tackled using a somewhat different approach in the next section.

6 Short presentations for S_n with a bounded number of relations

We now construct short presentations for S_n with a bounded number of relations. As a first step, we introduce a more general version of the 'gluing' construction described in Section 3.

Theorem 6.1 *Let $P = \{A, t \,|\, \mathcal{R}\}$ and $Q = \{B, t \,|\, \mathcal{S}\}$ be presentations for the symmetric groups S_{m+1} and S_{n+1} of degrees $m+1$ and $n+1$, with $m, n \geqslant 1$, such that the generator t denotes a transposition in both cases. Let M and N be sets of words on the sets $A \cup \{t\}$ and $B \cup \{t\}$ standing for generators of the natural subgroups S_m and S_n fixing a point moved by t, respectively, and let $[M,N]$ denote the set of all commutators $[u,v]$ with $u \in M$ and $v \in N$. Then*

$$\{\, A, B, t \,|\, \mathcal{R}, \mathcal{S}, [M,N] \,\}$$

is a presentation for S_{m+n}, again with the generator t standing for a transposition. In particular, this presentation has $|A|+|B|+1$ generators and at most $|\mathcal{R}|+|\mathcal{S}|+|M||N|$ relations. If y and z stand for an $(m+1)$-cycle in S_{m+1} and an $(n+1)$-cycle in S_{n+1} respectively, then zyt stands for an $(m+n)$-cycle in S_{m+n}.

PROOF: First, observe that there is a natural homomorphism from the group G with the given presentation onto S_{m+n} under which the subgroups generated by M and N map to $\mathrm{Sym}(\{1,\ldots,m\})$ and $\mathrm{Sym}(\{m{+}1,\ldots,m{+}n\})$, respectively, and t maps to the transposition $(m, m{+}1)$.

Now choose elements $a_1, a_2, \ldots, a_m \in \langle M, t\rangle$ that generate $\langle M, t\rangle = \langle A, t\rangle \cong S_{m+1}$ and satisfy the standard Coxeter presentation for S_{m+1}, with also $\langle a_1, \ldots, a_{m-1}\rangle = \langle M\rangle$ and $a_m = t$. Similarly, choose n elements $b_1, b_2, \ldots, b_n \in \langle N, t\rangle$ that generate $\langle N, t\rangle = \langle B, t\rangle \cong S_{n+1}$ and satisfy the standard Coxeter presentation for S_{n+1}, with also $b_1 = t$ and $\langle b_2, \ldots, b_n\rangle = \langle N\rangle$.

Next define $a_{m+j-1} = b_j$ for $1 < j \leqslant n$, and let $S = \{a_1, \ldots, a_{m+n-1}\}$. This set of involutions generates G, since it contains $t = a_m = b_1$ as well as generating sets for both $\langle M, t\rangle = \langle A, t\rangle$ and $\langle N, t\rangle = \langle B, t\rangle$. Consider the required Coxeter relations $(a_i a_j)^{m_{ij}} = 1$ for $1 \leqslant i \leqslant j \leqslant m + n - 1$. Observe that either $1 \leqslant i \leqslant j \leqslant m$ and this relation is implied since it is satisfied by $\langle M, t\rangle$; or $m \leqslant i \leqslant j \leqslant m + n - 1$ and this relation is implied since it is satisfied by $\langle N, t\rangle$; or $1 \leqslant i < m < j$ and this relation is implied since $a_i \in M$ commutes with $a_j \in N$. Thus $G = \langle S\rangle \cong S_{m+n}$.

Finally if y and z are elements of G standing for the $(m{+}1)$-cycle $(1, 2, \ldots, m + 1)$ and the $(n{+}1)$-cycle $(m, m{+}1, \ldots, m{+}n)$, respectively, then zty stands for the $(m{+}n)$-cycle $(1, 2, \ldots, m + n)$. □

We have the following special case.

Corollary 6.2 *Let $P = \{A, t \,|\, \mathcal{R}\}$ be a presentation for the symmetric group S_{m+1}, such that the generator t denotes a transposition, and let $\{a, c\}$ be a set of two words on the set $A \cup \{t\}$ standing for generators of the natural subgroup S_m fixing a point moved by t. Then $\{\, A, b, t \mid \mathcal{R}, b^2, (bt)^3, [a, b], [c, b]\,\}$ is a presentation for S_{m+2}, with the generators b and t standing for transpositions. In particular, this presentation has $|A|{+}2$ generators and at most $|\mathcal{R}|{+}4$ relations. Also an element standing for an $(m{+}2)$-cycle in S_{m+2} can be expressed in the form yb, where y stands for an $(m{+}1)$-cycle in S_{m+1}.*

PROOF: Take $n = 2$, $B = N = \{b\}$ and $M = \{a, c\}$ in Theorem 6.1. □

Also we need a variant of the Coxeter presentations:

Theorem 6.3 *For all $n \geqslant 3$, the finitely-presented group G_n having presentation*

$$\{\, a_1, a_2, \ldots, a_n \mid a_i^2, (a_i a_j)^3, (a_i a_j a_i a_k)^2 \text{ for all distinct } i, j, k\,\}$$

is isomorphic to the symmetric group S_{n+1}. In fact, the relations $a_i^2 = (a_i a_j)^3 = 1$ together with the relations $(a_1 a_j a_1 a_k)^2 = 1$ for distinct $j, k > 1$ are sufficient to give a presentation of S_{n+1}. Similarly, for each unordered triple $\{i, j, k\}$, it suffices to use only one relation of the form $(a_i a_j a_i a_k)^2$.

PROOF: The displayed presentation is due to Burnside [4] and Miller [12], but we give a complete proof of the theorem. First, clearly G_n has an image S_{n+1} (acting on $\{0, 1, 2, \ldots, n\}$), where a_i stands for the transposition $(0, i)$ for $1 \leqslant i \leqslant n$. Also $G_3 \cong S_4$, which we leave as an exercise for the reader. Now suppose by induction that G_n presents S_{n+1}. Then G_{n+1} is generated by a_{n+1}, a_1, and $a_j^{a_1}$ for $2 \leqslant j \leqslant n$. Let $t = a_1$, $M = \langle a_{n+1} \rangle$ and $N = \langle\, a_1 a_j a_1 : 2 \leqslant j \leqslant n \,\rangle$. Clearly $M \cong S_2$ and $\langle M, t \rangle \cong S_3$, with t being a transposition therein. Also $\langle N, t \rangle = \langle a_1, \ldots, a_n \rangle$ is a homomorphic image of G_n, and so $\langle N, t \rangle \cong S_{n+1}$, with t a transposition, and as N is a conjugate of $\langle a_2, \ldots, a_n \rangle \leqslant \langle N, t \rangle$, also $N \cong \mathrm{S}_n$. Next, the relations $(a_1 a_j a_1 a_{n+1})^2 = 1$ imply that the generator a_{n+1} of M commutes with the generators $a_1 a_j a_1$ of N, and hence $[M, N] = 1$. Since G_{n+1} is generated by M, N and t, Theorem 6.1 gives $G_{n+1} \cong S_{n+2}$. Finally, note that for each unordered triple $\{i, j, k\}$, the relator $(a_i a_j a_i a_k)^2$ is conjugate to $(a_i a_k a_i a_j)^2$, and equivalent to $(a_j a_i a_j a_k)^2$ because of the relation $(a_i a_j)^3 = 1$. □

We now proceed as in Section 5, but using the presentations from Theorem 6.3 in place of the Coxeter presentations, and without assuming that elements standing for a Singer cycle (or a longer cycle) are included in the generating set.

Lemma 6.4 *Let p be an odd prime, and α a primitive element of the field* $\mathrm{GF}(p)$. *Then* $\{\, a, b, c \,|\, a^p, b^{p-1}, c^2, a^b a^{-\lambda}, (bc)^2, (ac)^3 \,\}$ *is a presentation for* $\mathrm{PGL}(2, p)$, *in which the three generators a, b and c may be taken as standing respectively for the linear fractional transformations $z \mapsto z + 1$, $z \mapsto \lambda z$ and $z \mapsto -1/z$.*

PROOF: See Todd [18]. □

Theorem 6.5 *Let p be an odd prime, and let λ be a primitive element of* $\mathrm{GF}(p)$. *Then*

$$\{\, a, b, c, t \,|\, a^p, b^{p-1}, c^2, a^b a^{-\lambda}, (bc)^2, (ac)^3, t^2, [t, a], [t, b], [t, c]^3, (tt^c tt^{ca})^2, (abt^c)^p \,\}$$

is a 4-generator 12-relator presentation for the symmetric group S_{p+2}, in which att^c stands for a $(p+2)$-cycle $(1, 2, \ldots, p+2)$ and t stands for a transposition of the form $(i, i+1)$.

PROOF: First, observe that there is a natural homomorphism from the group G with the given presentation onto the symmetric group S_{p+2} acting on $\{\star, \infty\} \cup \mathrm{GF}(p)$, given by mapping t to the transposition $(\star, \infty)$ and mapping a, b, c to elements that fix $\star$ and act on the projective line $\{\infty\} \cup \mathrm{GF}(p)$ as the linear fractional transformations given in Lemma 6.4. In particular, att^c maps to the $(p+2)$-cycle $(0, 1, 2, \ldots, p-1, \star, \infty)$. Moreover, the elements a, b and c generate a subgroup H isomorphic to $\mathrm{PGL}(2, p)$.

The images of the elements a and b generate the stabiliser of the point ∞ in this permutation representation, and as a and b centralise t, it follows that t has exactly $p + 1$ conjugates under the action of $H = \langle a, b, c \rangle$, and these elements map to the transpositions $t_i = (*, i)$ for $i \in \{\infty\} \cup \mathrm{GF}(p)$.

The relations $t^2 = [t,c]^3 = (tt^ctt^{ca})^2 = 1$ now give all that is required to invoke Theorem 6.3, and so the subgroup K generated by these $p+1$ conjugates of t is isomorphic to S_{p+2}. Hence the given group is a quotient of a semi-direct product of S_{p+2} by $\mathrm{PGL}(2,p)$. But S_{p+2} has trivial outer automorphism group, so this group must be a quotient of the direct product $S_{p+2} \times \mathrm{PGL}(2,p)$.

On the other hand, the relation $(abt^c)^p = 1$ may be written as

$$(ab)^p((ab)^{-(p-1)}t^c(ab)^{p-1})((ab)^{-(p-2)}t^c(ab)^{p-2})\ldots((ab)^{-2}t^c(ab)^2)((ab)^{-1}t^c(ab))t^c = 1,$$

and as the element ab is known to have order $p-1$ (because it represents the linear fractional transformation $z \mapsto \lambda(z+1)$), it follows that $ab = (ab)^p$ is expressible as a product of conjugates of t^c (and hence of t). In particular, ab lies in K, and therefore all conjugates of ab lie in K, and so the subgroup $H \cong \mathrm{PGL}(2,p)$ generated by these conjugates lies in K. Thus $G = HK = K \cong S_{p+2}$. □

An alternative (and shorter) presentation of S_{p+2} on the generators a, c and t can be obtained from the following:

Lemma 6.6 *Let p be any odd prime. Then $\{\, a, c \,|\, a^p, acacac^{-1}, (a^{(p+1)/2}ca^4c)^2 \,\}$ is a presentation for $\mathrm{PSL}(2,p)$, in which the two generators a and c may be taken as standing respectively for the linear fractional transformations $z \mapsto z+1$ and $z \mapsto -1/z$.*

Proof: See Sunday [17]. □

Theorem 6.7 *Let p be an odd prime, and let λ be a primitive element of $\mathrm{GF}(p)$, with inverse μ. Then*

$$\{\, a, c, t \,|\, a^p, acacac^{-1}, (a^{(p+1)/2}ca^4c)^2, \\ t^2, [t,a], [t, ca^\lambda ca^\mu c], [t,c]^3, (tt^ctt^{ca})^2, (tt^ctt^{ca^\lambda})^2, (at^c)^{p+1} \,\}$$

is a 3-generator 10-relator presentation for the symmetric group S_{p+2}, in which att^c stands for a $(p+2)$-cycle $(1,2,\ldots,p+2)$ and t stands for a transposition of the form $(i,i+1)$. Moreover, if $p \equiv 3$ mod 4 then the relator $(tt^ctt^{ca^\lambda})^2$ is redundant and just 9 relators are needed.

Proof: This is similar to the proof of Theorem 6.5, but now $H = \langle a, c\rangle$ is isomorphic to $\mathrm{PSL}(2,p)$, by Lemma 6.6, and the images of the elements a and $ca^\lambda ca^\mu c$ generate the stabiliser of the point ∞. In contrast to $\mathrm{PGL}(2,p)$, the group $\mathrm{PSL}(2,p)$ is not 3-transitive on the projective line $L = \{\infty\} \cup \mathrm{GF}(p)$, but it has only two orbits on ordered triples of distinct points of L, and so the required relations for the subgroup $K(\cong S_{p+2})$ generated by conjugates of t can be obtained from just the six relations $t^2 = [t,a] = [t, ca^\lambda ca^\mu ca^\mu] = [t,c]^3 = (tt^ctt^{ca})^2 = (tt^ctt^{ca^\lambda})^2 = 1$. If $p \equiv 3$ mod 4, then the relator $(tt^ctt^{ca^\lambda})^2$ can be dropped because $\mathrm{PSL}(2,p)$ has a single orbit on unordered triples of points of L in that case — see the last sentence of Theorem 6.3. □

Ideally, we would like to convert this presentation to one on the standard generators of S_{p+2}. Taking these to be $x = (\star, \infty)$ and $y = (\star, \infty, 0, 1, \ldots, p-1)$, we have $x = t$ and $y = att^c$, and conversely $a = yxy^{-1}xy$. However, we cannot express c as a short word in x and y, and suspect that this cannot be done.

Instead we apply Theorem 6.1 to Theorem 6.5 and obtain the following:

Theorem 6.8 *Let p and q be odd primes, and let λ and μ be primitive elements of* GF(p) *and* GF(q)*, respectively. Then*

$$\begin{aligned}\{\, a,b,c,d,e,f,t \mid\ & a^p, b^{p-1}, c^2, a^b a^{-\lambda}, (bc)^2, (ac)^3, d^q, e^{q-1}, f^2, d^e d^{-\mu}, (ef)^2, (df)^3,\\ & t^2, [t,a], [t,b], [t,d], [t,e], [t,c]^3, [t,f]^3, (tt^c tt^{ca})^2, (tt^f tt^{fd})^2, (abt^c)^p, (det^f)^q,\\ & [a,d], [a,f], [c,d], [cf,t]^2 \,\}\end{aligned}$$

is a 7-generator 27-relator presentation for the symmetric group S_{p+q+2}, in which $at^f tt^c d$ stands for a $(p+q+2)$-cycle $(1, 2, \ldots, p+q+2)$ and t stands for a transposition of the form $(i, i+1)$.

Proof: Take $m = p+1$ and $n = q+1$, and also $A = \{a, b, c\}$ and $B = \{d, e, f\}$, and $M = \{a, t^{ct}\}$ and $N = \{d, t^{ft}\}$. Then using Theorem 6.1 we obtain a presentation for $S_{m+n} = S_{p+q+2}$ the same as above but with the relators $[a, t^{ft}]$, $[d, t^{ct}]$ and $[t^{ct}, t^{ft}]$ in place of $[a, f]$, $[c, d]$ and $[cf, t]^2$. Since t commutes with both a and d in this group, the relations $[a, f] = 1$ and $[c, d] = 1$ imply that $[a, t^{ft}] = 1$ and $[d, t^{ct}] = 1$; further the relation $[t^{ct}, t^{ft}] = 1$ is equivalent to $1 = [t^c, t^f] = (ctcftf)^2$ and hence to $[cf, t]^2 = (fctcft)^2 = 1$. □

A similar application of Theorem 6.1 to Theorem 6.7 gives:

Theorem 6.9 *Let p and q be odd primes, and let λ and μ be primitive elements of* GF(p) *and* GF(q)*, with inverses ρ and σ respectively. Then*

$$\begin{aligned}\{\, a,c,d,f,t \mid\ & a^p, acacac^{-1}, (a^{(p+1)/2} ca^4 c)^2, d^q, dfdfdf^{-1}, (d^{(q+1)/2} f d^4 f)^2,\\ & t^2, [t,a], [t, ca^\lambda ca^\rho c], [t,d], [t, fd^\mu fd^\sigma f],\\ & [t,c]^3, [t,f]^3, (tt^c tt^{ca})^2, (tt^f tt^{fd})^2, (tt^c tt^{ca^\lambda})^2, (tt^f tt^{fd^\mu})^2, (at^c)^{p+1}, (dt^f)^{q+1},\\ & [a,d], [a,f], [c,d], [cf,t]^2 \,\}\end{aligned}$$

is a 5-generator 23-relator presentation for the symmetric group S_{p+q+2}, in which $at^f tt^c d$ stands for a $(p+q+2)$-cycle $(1, 2, \ldots, p+q+2)$ and t stands for a transposition of the form $(i, i+1)$. Moreover, if $p \equiv 3 \bmod 4$ then the relator $(tt^c tt^{ca^\lambda})^2$ is redundant, and if $q \equiv 3 \bmod 4$ then the relator $(tt^f tt^{fd^\mu})^2$ is redundant.

We now adopt the approach taken at the end of Section 5. Namely we use the results of Vinogradov or Ramaré and a gluing construction of the type given in Theorem 6.1 to produce the following.

Corollary 6.10 *Let n be any positive even integer such that $n-2$ is the sum of odd primes p and q. Then the symmetric group S_n has a defining presentation of length*

$O(\log n)$ having five generators and at most 23 relations. One of the generators can be taken as the transposition $(1,2)$, and the n-cycle $(1,2,\ldots,n)$ can be expressed as a word of length at most 9 in the five generators.

Corollary 6.11 *For every even integer $n > 1$, the symmetric group S_n has a defining presentation of length $O(\log n)$ having 13 generators and at most 77 relations, or 9 generators and at most 50 relations if n is sufficiently large. One of the generators can be taken as the transposition $(1,2)$, and the n-cycle $(1,2,\ldots,n)$ can be expressed as a word of length at most 29 in the 13 generators, or at most 19 in the 9 generators if n is sufficiently large.*

Similar presentations for S_n with n odd can be obtained from the above, together with Corollary 6.2.

It remains to be decided whether or not there is a presentation of S_n on the standard generating set (consisting of the n-cycle $(1,2,\ldots,n)$ and a transposition of the form $(i,i+1)$), with a uniformly bounded number of relators of total length $O(\log n)$.

7 Alternating groups

Presentations for the alternating groups A_n can be obtained easily from those for S_n, using the Reidemeister–Schreier process [16, Chapter 6].

If P_n is any presentation for S_n, let E_n and O_n be the subsets of generators in P_n that stand for even and odd permutations in S_n, respectively. A Schreier transversal for A_n in S_n is $\{1, x\}$, where x is any one of the generators in O_n (such as one standing for a single transposition). The Schreier generators for A_n are then all elements of the form y and xyx^{-1} where $y \in E_n$, together with all elements of the form $xz^{\pm 1}$ where $z \in O_n$. The Reidemeister–Schreier relators are just those in the presentation for S_n, rewritten in terms of these Schreier generators, plus conjugates of these by x^{-1}.

In particular, the number of generators in the resulting presentation for A_n is at most twice the number of generators in the presentation P_n, and the number of relations is at most twice the number of relations in P_n. Accordingly, the length of the resulting presentation is at most twice that of P_n.

In fact, many of the relators are equivalent to their conjugates under x, so the number of relations in the presentation for A_n is very similar to that for S_n. For example, conjugation of $(xy^{-1}xy)^k$ by the involution x gives $(y^{-1}xyx)^k$, which is simply the inverse of $(xy^{-1}xy)^k$.

Also if P_n has just two generators x and y, standing for a transposition and the n-cycle $(1,2,\ldots,n)$, the Schreier generators for A_n are either y and xyx (when n is odd) or xy and xy^{-1} (when n is even).

As a result, the constructions in Sections 2, 3 and 4 give rise to presentations for the alternating group A_n that are (respectively) linear in n, quadratic in $\log n$ and linear in $\log n \log\log n$, and the ones in Sections 5 and 6 give presentations for A_n with a bounded number of generators and relations, for all n. In particular, those in Section 5

give rise to presentations for A_n on two generators with a bounded number of relations, giving a positive answer to the second part of Question (3) in Section 7 of [5].

8 Short presentations for small degree cases

We have investigated, using computational techniques, short presentations for the symmetric groups of degree at most 9. One motivation for this work is to use these in the base cases of the general constructions described here.

Theorem 8.1 reports the shortest-length presentations we have found for S_n for small values of n, on generating sets which include a 2-cycle u and an n-cycle v. (If an additional generator appears, then that generator is redundant.)

Theorem 8.1

$$
\begin{aligned}
S_3 &\cong \langle\, u, v \mid u^2, v^3, (uv)^2 \,\rangle;\\
S_4 &\cong \langle\, u, v \mid u^2, v^4, (uv)^3 \,\rangle;\\
S_5 &\cong \langle\, u, v, w \mid u^2, v^5, w^3, (vw)^2, v^{-1}uv^2uw \,\rangle;\\
S_6 &\cong \langle\, u, v, w, x \mid u^2, vw^2x^{-1}, w^5, vx^{-1}v^{-1}wx, uxv^{-1}w^{-1}x \,\rangle;\\
S_7 &\cong \langle\, u, v, w, x \mid u^2, v^{-2}wx^2, v^{-1}wx^{-1}wx, uv^{-1}w^{-1}xwv^{-1}, (uv^{-1}x^{-1})^2 \,\rangle;\\
S_8 &\cong \langle\, u, v, w, x \mid u^2, w^2, (uw)^2, x^5, x^{-1}v^{-1}uvxu, x^{-1}wuv^{-1}x^{-1}w, v^4x^{-1}wx \,\rangle;\\
S_9 &\cong \langle\, u, v, w \mid u^2, w^2, uvwuwv^{-1}, v^9, uvuwuv^2uv^{-3}, (v^2w)^4 \,\rangle.
\end{aligned}
$$

These claims can be readily established by coset enumeration over a suitable subgroup (see [16]). We used the coset enumeration machinery of Havas and Ramsay [10], available within MAGMA [3], extensively in these investigations.

References

[1] 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.

[2] Robert Beals, Charles R. Leedham-Green, Alice C. Niemeyer, Cheryl E. Praeger, and Ákos Seress, A black-box group algorithm for recognizing finite symmetric and alternating groups. I. *Trans. Amer. Math. Soc.* **355** (2003), 2097–2113.

[3] W. Bosma, J. Cannon, and C. Playoust, The MAGMA algebra system I: The user language, *J. Symbolic Comput.* **24** (1997), 235–265.

[4] W. Burnside, *Theory of Groups of Finite Order*, 2nd ed. Dover Publications, Inc., New York, 1955, xxiv+512 pp.

[5] C.M. Campbell, G. Havas, C. Ramsay, and E.F. Robertson, Nice efficient presentations for all small simple groups and their covers. *LMS J. Comput. Math.* **7** (2004), 266–283.

[6] M.D.E. Conder, C.R. Leedham-Green, and E.A. O'Brien, Short presentations for classical groups. Preprint, 2006.

[7] H.S.M. Coxeter and W.O.J. Moser, *Generators and Relations for Discrete Groups*, 4th ed. Springer-Verlag (Berlin), 1980, ix+169 pp.

[8] R.M. Guralnick, W.M. Kantor, M. Kassabov and A. Lubotzky, "Presentations of finite simple groups: a quantitative approach". Preprint, 2006.

[9] G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers*, 5th ed. The Clarendon Press, Oxford University Press, New York, 1979, xvi+426 pp.

[10] George Havas and Colin Ramsay, Proving a group trivial made easy: a case study in coset enumeration. *Bull. Austral. Math. Soc.* **62** (2000), 105–118.

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

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

[13] E.H. Moore, Concerning the abstract groups of order $k!$ and $\frac{1}{2}k!$, *Proc. London Math. Soc.* **28** (1897), 357–366.

[14] E.A. O'Brien, Towards effective algorithms for linear groups. *Finite Geometries, Groups and Computation*, (Colorado), September 2004. De Gruyter, Berlin, 163-190, 2006.

[15] O. Ramaré, On Šnirelman's constant. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **22** (1995), 645–706.

[16] Charles C. Sims, *Computation with finitely presented groups*. Cambridge University Press, 1994, xiii+604 pp.

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

[18] J.A. Todd, A note on the linear fractional group. *J. London Math. Soc.* **7** (1932), 195–200.

[19] I.M. Vinogradov, *The method of trigonometrical sums in the theory of numbers* (Russian). *Trav. Inst. Math. Stekloff* **23** (1947). 109 pp.

J.N. Bray ⟨J.N.Bray@qmul.ac.uk⟩:
School of Mathematical Sciences, Queen Mary, University of London, London E1 4NS, United Kingdom

M.D.E. Conder ⟨m.conder@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

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