Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Algebra 292 (2005) 4–46
JOURNAL OF Algebra
www.elsevier.com/locate/jalgebra

Constructive recognition of finite alternating and symmetric groups acting as matrix groups on their natural permutation modules

Robert Beals [a], Charles R. Leedham-Green [b], Alice C. Niemeyer [c,1], Cheryl E. Praeger [c,*,1], Ákos Seress [d,2]

[a] *IDA Center for Communications Research, Princeton, NJ 08540, USA*
[b] *School of Mathematical Sciences, Queen Mary and Westfield College, London E1 4NS, United Kingdom*
[c] *School of Mathematics and Statistics, University of Western Australia, Crawley, WA 6009, Australia*
[d] *Department of Mathematics, The Ohio State University, Columbus, OH 43210, USA*

Received 1 March 2004

Available online 14 March 2005

Communicated by Eamonn O'Brien

Abstract

We present a Las Vegas algorithm which, for a given matrix group known to be isomorphic modulo scalars to a finite alternating or symmetric group acting on the fully deleted permutation module, produces an explicit isomorphism with the standard permutation representation of the group. This algorithm exploits information available from the matrix representation and thereby is faster than existing 'black-box' recognition algorithms applied to these groups. In particular, it uses the fact that certain types of elements in these groups can be identified and constructed from the structure of their characteristic polynomials. The algorithm forms part of a large-scale program for computing with groups of matrices over finite fields. When combined with existing 'black-box' recognition algorithms, the results of this paper prove that any d-dimensional absolutely irreducible matrix rep-

* Corresponding author.
E-mail addresses: beals@idaccr.org (R. Beals), crlg@maths.qmw.ac.uk (C.R. Leedham-Green), alice@maths.uwa.edu.au (A.C. Niemeyer), praeger@maths.uwa.edu.au (C.E. Praeger), akos@math.ohio-state.edu (Á. Seress).

[1] Partially supported by Australian Research Council Discovery Project Grant DP0209760.
[2] Partially supported by the National Science Foundation.

doi:10.1016/j.jalgebra.2005.01.035

resentation of a finite alternating or symmetric group, over a finite field, can be recognised with $O(d^{1/2})$ random group elements and $O(d^{1/2})$ matrix multiplications, up to some logarithmic factors.

Keywords: Constructive recognition; Matrix groups; Alternating and symmetric groups; Probabilistic method

1. Introduction

In this paper we present an algorithm designed to recognise finite alternating and symmetric groups acting naturally as matrix groups in their smallest dimensional, faithful, absolutely irreducible representations over a finite field of characteristic p. The reason for focusing on the special case of these representations of A_n and S_n is that they arise in a special way as maximal subgroups (modulo scalars) of finite classical groups. The algorithm given in this paper requires $O(n^\alpha)$ random selections and $O(n^\alpha \log^2 n)$ matrix multiplications, where $\alpha = 1/3$ if $p \neq 3$ and $\alpha = 1/2$ for $p = 3$, and is asymptotically faster than an implementation for these groups of the fastest known 'black-box' algorithm to recognise finite alternating and symmetric groups. Moreover, the algorithm given in this paper, combined with the 'black-box' algorithm in [5], provides a uniform complexity of $O(d^{1/2})$ random selections and $O(d^{1/2})$ matrix multiplications (up to some logarithmic factors) to recognise any d-dimensional absolutely irreducible representation of a finite alternating or symmetric group over a finite field, see Section 2.1 for details.

Aschbacher [1] described eight families of maximal subgroups of the finite classical groups of dimension d over a field $\mathbb{F}$ of order q (where $q = p^a$ for some prime p). He proved that any maximal subgroup G not lying in one of these eight families must be nearly simple, that is $G/(G \cap Z)$ has a simple socle S where Z denotes the subgroup of non-zero scalar matrices. Moreover, for these nearly simple groups, the pre-image of S in G is absolutely irreducible on the underlying vector space V, is not realisable over a proper subfield, and is not a classical group in its natural representation. Every abstract finite simple group can occur in this way as the simple group S. In Section 2 we briefly describe how Aschbacher's result has been used as the underpinning framework for a matrix recognition project for matrix group computation, and how the algorithm of this paper fits into this framework.

Moreover, it was shown by Liebeck [22] that, for sufficiently high dimensions, the largest among the nearly simple maximal subgroups mentioned above are the groups $Z \times S_n$ acting on the fully deleted permutation module over $\mathbb{F}$ corresponding to the natural transitive permutation action of S_n of degree n. This module will be described in detail in Section 3.1. Its dimension is $n-1$ if the characteristic p does not divide n, and is $n-2$ if p does divide n.

Our main result is Theorem 1.1. It involves several parameters, namely ω, ρ_F and ξ. The parameter ξ is an upper bound on the cost of producing one random element of G; ρ_F is an upper bound on the cost of performing one operation (addition, multiplication or finding an inverse) in the finite field $\mathbb{F}$ of order q; and ω is a real number for which

there exists an algorithm for multiplying two $n \times n$ matrices over a field with $O(n^\omega)$ field operations. There are algorithms known for which $\omega < 2.376$, see [11].

Theorem 1.1. *There is a Las Vegas algorithm with the following specifications. It takes as input a positive real number ε such that $0 < \varepsilon < 1$ and a subset X of $\mathrm{GL}(n-\delta, q)$, where $n \geqslant 5$, q is a power of a prime p, $\delta = 1$ or 2 according as p does not or does divide n, and if $G = \langle X \rangle$ then $G' \cong A_n$. The output is a monomorphism λ from G to $Z_{q-1} \times S_n$. The algorithm succeeds with probability at least $1-\varepsilon$, and the cost is*

$$O\big(\log\big(\varepsilon^{-1}\big)n^\alpha\big(\xi + \rho_F \log^2 n\big(n^\omega + n \log nq \log\log n\big)\big) + |X|\rho_F n^\omega\big),$$

where $\alpha = 1/3$ if $p \neq 3$ and $1/2$ if $p = 3$. The cost of evaluating λ on a given element of G is $O(n^\omega \rho_F)$, and similarly the cost of evaluating λ^{-1} on a given element of $\lambda(G)$ is $O(n^\omega \rho_F)$.

Thus, for small fields, and small generating sets, the cost of constructing the monomorphism is

$$O\big(\log\big(\varepsilon^{-1}\big)\big(n^\alpha \xi + \rho_F n^{\omega+1/2} \log^2 n\big)\big)$$

with α as above. The assumption that $n \geqslant 5$ covers all parameter values for this family of nearly simple matrix groups. Clearly to prove the theorem we may ignore small values of n, and in fact for one part of the algorithm presented in the paper we assume that $n \geqslant 13$. Our approach is to find a new basis for the underlying vector space, and if the group G were replaced by a conjugate under the corresponding change of basis matrix, then our procedures given in Section 10 evaluate λ and λ^{-1} on given group elements at a cost of $O(n^2 \rho_F)$ per element.

After the commentary in Section 2 on the matrix group project we describe, in Section 3, the context of the algorithm and in particular we define the fully deleted permutation module. We explain there the various components of the algorithm, and the proof of Theorem 1.1.

A complete implementation of the algorithm has been made by Stephen Howe, assisted by Maska Law, in the computer language GAP4 [12]. The authors wish to thank Stephen for his care in reading and implementing the various procedures in the paper, and in particular for locating several mistakes and misprints. The authors also acknowledge the advice from an anonymous referee that led to an improved exposition and layout of the paper.

2. Commentary on matrix group recognition

From the practical point of view the algorithm presented in this paper forms part of one of the matrix recognition projects. The objective of such projects (see [2,16,19,20]) is to produce a computer software system that accepts as input a subset X of $\mathrm{GL}(d, q)$ for some $d > 0$ and prime power q, and determines, among other things, a composition series (or composition tree) for the group $G = \langle X \rangle$. The project described in [19,20] is

heavily dependent on the theorem of Aschbacher mentioned in the introduction, which can be interpreted as stating that if G does not contain the special linear group, and is not almost simple modulo scalars, then G preserves a geometrical structure on the underlying vector space V. Making this theorem constructive reduces matrix recognition to dealing with groups that are almost simple groups modulo scalars.

If G is found to preserve a geometrical structure on V, then the geometric structure that is preserved by G is determined explicitly. This usually involves finding a basis for the underlying vector space V that exhibits the structure. For example, if a G-invariant direct sum decomposition of V is discovered, a basis is found that is the union of bases for the direct summands. If a G-invariant tensor decomposition is discovered, a change of basis is performed so that the elements of G are exhibited as Kronecker products of smaller matrices, etc. This change of basis has various useful consequences. Firstly, in terms of the new basis, the elements of G can be written in a more compact form: in the first of the above examples, as an element of a wreath product of a general linear group of smaller dimension by a symmetric group, and in the second example as a Kronecker product. This produces a saving in the time taken to multiply two group elements, which may be very dramatic, as well as a useful saving in space. Secondly, given any element of $\mathrm{GL}(d, q)$, with respect to this basis one can see at once whether it preserves the given structure, and if so, write it in the appropriate form. Thus, recognising the fact that G preserves some geometric structure reduces further problems of processing G to easier ones.

On the other hand, if G is found not to preserve such a structure, then in general no such reduction is possible, and usually we have to deal with G as it stands, as an almost simple group modulo scalars, using black-box recognition techniques. In addition, if G is almost simple modulo scalars, and is realisable over a proper subfield, it is sometimes desirable to recognise it as given, rather than first re-writing the group over the smaller field. In the case of finite alternating and symmetric groups, these algorithms construct an isomorphism with the natural permutation representation of the group.

If G acts on the fully deleted permutation module V as A_n or S_n, the situation of interest here, then G is such a group, but in this case we can do better than implementing the black-box group algorithms. In this special case the structure of G is made explicit by a suitable change of basis for V, and so our approach is very similar to the approach above for the earlier Aschbacher categories.

2.1. The complexity of recognising A_n and S_n

The asymptotically most efficient black-box recognition algorithm known for A_n and S_n is in [5], and requires $O(n)$ random selections and $O(n \log n)$ group multiplications. Applying this algorithm in the matrix group setting: if A_n or S_n, or one of their covering groups, were given as an irreducible subgroup in $\mathrm{GL}(d, q)$, and if n were $O(d^{1/2})$, then the time complexity of this algorithm would be $O(d^{1/2}\xi + \rho_F d^{\omega+1/2})$ (up to logarithmic factors).

Now it follows from results of James [14, Theorem 7] and Wagner [29] that, for $n \geqslant 15$, any faithful irreducible representation of A_n or S_n or one of their covering groups, apart from the representation on their deleted permutation modules, must have dimension $d \geqslant n(n-5)/4$, and hence $n = O(d^{1/2})$. Hence the algorithm presented in this paper ensures

that all d-dimensional irreducible representations of A_n and S_n, or their covering groups, can be constructively recognised in $O(d^{\omega+1/2})$ time up to some logarithmic factors.

The principal tool at our disposal that makes use of the fact that we are working with matrices rather than with a black-box group is the computation of the characteristic polynomial of group elements. For example, a crucial step in all 'black-box' recognition algorithms for alternating and symmetric groups is to find an element that is a transposition, a 3-cycle, or a double transposition in the natural representation. The algorithm presented here includes a faster method of finding such an element than the method of finding a 3-cycle given in [4,5,7]. In addition, the present algorithm does not require the construction of an n-cycle or $(n-1)$-cycle, instead making use of certain elements with order divisible by a prime greater than $3n/5$.

There are several reasons why we are able to make use of these faster methods. For example, we are able to recognise from their characteristic polynomials certain matrices from which we can construct a 3-cycle or double transposition (see Section 6), and upon identifying such matrices we are then able to extract the associated elements efficiently because we have available a fast method for determining the orders of these matrices from their characteristic polynomials (see Section 5). In addition, having constructed the standard basis for the fully deleted permutation module we obtain a positive identification of A_n and S_n. This obviates the need to confirm the supposed isomorphism type of G, which would otherwise have to be done by finding for the group a new generating set that satisfies a standard presentation.

Finally, we point out that the isomorphism λ in Theorem 1.1 evaluates images of elements of $Z_V \times H$ as pairs $(b, g) \in Z_{q-1} \times S_n$, where b is a non-zero scalar and g is a permutation. Similarly λ^{-1} computes the pre-image of such a pair as a matrix.

For applications of this algorithm in the matrix group recognition project, we would need also to construct straight-line programs from $\{\lambda(x) \mid x \in X\}$ to (b, g), and [5] contains an algorithm that does this, producing straight-line programs of length $O(n \log n)$, in $O(n^2 \log n)$ time. However, the evaluation in $Z_V \times H$ of such a straight-line program would cost $O(\rho_F n^{\omega+1} \log n)$ which is more expensive than the running time of our recognition algorithm. In order to construct (and evaluate within $Z_V \times H$) straight-line programs at no greater cost than the rest of the algorithm, the underlying open problem that needs to be solved is to find an algorithm that, for the natural permutation representation of S_n, computes a straight-line program of length $O(n^{1/3} \log^2 n)$ from the standard generating set $\{(12), (12 \ldots n)\}$ to an arbitrary permutation in S_n.

2.2. Other complexity issues

Another delicate issue arises from the construction of random elements. The complexity analysis is given in terms that involve the time required to construct a random element, but the algorithm loses its advantage in practice if this has a cost significantly worse than the cost of making a bounded number of group multiplications. Provided that the size of the given generating set is bounded the product replacement algorithm [9] will run in practice within these cost constraints; but despite very interesting theoretical progress, the assertion that the product replacement algorithm performs this well remains a well supported con-

jecture, and is not a theorem. This is one reason for including the cost of producing random elements as a parameter in the cost estimate for our algorithm.

In practice it seems unlikely that useful implementations of the algorithm in this paper will match its $o(d^3)$ complexity estimate. For example, Strassen's algorithm for multiplying two $d \times d$ matrices, which is useful in practice, has complexity $O(d^k)$ where $k = \log_2 7 > 2.8$, and using this would produce an algorithm that is slower than $O(d^3)$ (but still faster than an application of the 'black-box' algorithm from [5]). In addition, keeping the theoretical complexity below $O(d^3)$ meant, for example, that we could not calculate the minimum polynomial of a matrix, as we know of no algorithm for this that has complexity better than Las Vegas $O(d^3)$ field operations. This, in turn, is the reason for introducing the new algorithm in Section 5 for computing the order of an element of $Z_V \times H_0$ using only the characteristic polynomial, rather than the minimal polynomial.

There would have been some advantage in our algorithm, especially in Section 6, to pass from G to its derived subgroup $G' = A_n$. This would, in particular, have simplified the procedures in Section 6 for determining the scalar associated with a given group element. There is an easy algorithm [3] to pass from a generating set of G to a generating set for G', but its time requirement is asymptotically greater than the time requirement of our algorithm. Also, had we used such an algorithm, we would have needed to make random selections from two different groups, namely the input group and its derived subgroup. However, as algorithms for making random selections require a certain amount of preprocessing, it is not unreasonable from a practical as well as a theoretical point of view to abstain from doing this.

3. Context of the algorithm

In this section we define the deleted permutation module and its standard basis, we specify the algorithmic set-up, and we outline the principal steps in the algorithm, describing where these are presented and analysed in the paper. We shall use the notation introduced in this section throughout the paper.

3.1. Permutation modules and standard bases

Consider the group $\mathrm{GL}(n, q)$ acting naturally on the vector space $U = \mathbb{F}^n$ of n-dimensional row vectors, where $\mathbb{F}$ is a field of order $q = p^a$ (p a prime), and let $\mathcal{E}_0 := (e_1, \ldots, e_n)$ denote the standard (ordered) basis, where e_i is the row vector which has ith entry 1 and all other entries 0.

Let H_0 denote the subgroup of $\mathrm{GL}(n, q)$ consisting of all the permutation matrices. Then $H_0 \cong S_n$ and H_0 permutes the standard basis vectors and leaves invariant the all-1 vector $e = (1, \ldots, 1) = \sum e_i$. Set $E := \langle e \rangle$. Also H_0 leaves invariant the co-dimension 1 subspace $W := \{(x_1, \ldots, x_n) \mid \sum x_i = 0\}$ of U. Following [17, pp. 185–186], the subspace $V := W/(W \cap E)$ is called the *fully deleted permutation module*. Now $e \in W$ if and only if p divides n, and hence

$$\dim V = n - \delta, \quad \text{where } \delta = \begin{cases} 1 & \text{if } p \text{ does not divide } n, \\ 2 & \text{if } p \text{ divides } n. \end{cases} \tag{1}$$

If $n \geqslant 5$, then $H_0' \cong A_n$ acts irreducibly on V, while if $n \geqslant 10$ then by [27–29] and [14, Theorem 6], or see [17, 5.3.7], every faithful irreducible $\mathbb{F}H_0'$-module has dimension at least $n-2$, and V is the only such module of dimension at most n.

We shall need to compute with the actions of H_0 on both U and V, and since H_0 acts faithfully on V, we shall often regard H_0 as a subgroup of $\mathrm{GL}(V)$ as well as working with it (as defined) as a subgroup of $\mathrm{GL}(U)$. The normaliser of H_0' in $\mathrm{GL}(U)$ is $Z_U \times H_0$, where Z_U is the subgroup of non-singular scalar matrices in $\mathrm{GL}(U)$. Similarly the normaliser of H_0' in $\mathrm{GL}(V)$ is $Z_V \times H_0$, where Z_V is the subgroup of non-singular scalar matrices in $\mathrm{GL}(V)$. We shall sometimes write $Z \times H_0$ without specifying whether the action is on U or on V when it helps the flow of the discussion, and the meaning is clear from the context.

We shall work with the characteristic polynomials of elements of $Z \times H_0$ (where $Z = Z_U$ or Z_V). For $g \in H_0$ we often identify g with the permutation of S_n corresponding to it, and we say that g has *type* $1^{c_1}2^{c_2}\ldots n^{c_n} = \prod_i i^{c_i}$, where $\sum i c_i = n$, if g has c_i cycles of length i for each $i = 1, \ldots, n$. Our notation for the characteristic polynomials on U and V of elements in $Z \times H_0$ is given in Notation 4.1.

The standard basis for V we shall use as a reference basis in the algorithm is $\mathcal{B}_0 := (v_1, \ldots, v_{n-\delta})$, where

$$v_i = e_i - e_{i+1} + (W \cap E) \quad \text{for } 1 \leqslant i \leqslant n-\delta \tag{2}$$

and the e_i form the standard basis $\mathcal{E}_0 := (e_1, \ldots, e_n)$ for U as defined above. The important property of $\mathcal{B}_0$ is that each vector has an expression involving exactly two of the e_i and every e_i (apart from e_1, e_n and, if $\delta = 2$ also e_{n-1}) occurs exactly twice, with different signs, and in consecutive vectors of $\mathcal{B}_0$.

3.2. The algorithmic set-up

In the practical algorithmic application we shall be given an absolutely irreducible subgroup G of $\mathrm{GL}(d,q) = \mathrm{GL}(V)$, where $d = n - \delta$ with δ as in (1), such that G is conjugate to a subgroup of $Z_V \times H_0$ containing H_0'. The problem is the following.

Algorithmic Problem. *Given a subgroup $G = \langle X \rangle$ of $\mathrm{GL}(V) = \mathrm{GL}(d,q)$ satisfying $H' \leqslant G \leqslant Z_V \times H$, where H is conjugate to H_0 in $\mathrm{GL}(d,q)$, construct a monomorphism $\lambda : G \to Z_{q-1} \times S_n$.*

The monomorphism λ is constructed via a matrix that conjugates X into $Z_V \times H_0$. Equivalently, the key outcome of the algorithm is a basis for V on which $\langle X \rangle$ acts in the same way that H_0 acts on the standard basis $\mathcal{B}_0$ defined in (2). Given this basis inverse isomorphisms between $\langle X \rangle$ and the corresponding subgroup of $Z_V \times H_0$ can be read off very quickly; much faster than the corresponding isomorphisms when G is recognised as a black-box group.

We shall call a sequence of vectors $(w_1, \ldots, w_r)$ from V a *linked sequence relative to* H_0 if there exist distinct positive integers $j_1, j_2, \ldots, j_{r+1}$ and a field element $b \in \mathbb{F}^{\#}$ such that

$$w_j = b(e_{j_i} - e_{j_{i+1}}) + (W \cap E) \quad \text{for } 1 \leqslant i \leqslant r.$$

A linked sequence relative to H_0 of length $r = n - \delta$ is a basis of V and we call it a *linked basis relative to* H_0. For example, the reference basis $\mathcal{B}_0$ defined in (2) is a linked basis relative to H_0. In our algorithm the given group G will involve a conjugate H of H_0. We will construct a linked basis relative to H (defined below) that will enable us to conjugate G to a subgroup of $Z_V \times H_0$.

Each linked basis relative to H_0 is an image of $\mathcal{B}_0$ under an element of $Z_V \times H_0$. Set $\mathcal{L}_0 := \{\mathcal{B}_0 A \mid A \in Z_V \times H_0\}$, the set of all linked bases for V relative to H_0. Let $S \in \mathrm{GL}(V)$ be such that $S^{-1} H_0 S = H$.

Lemma 3.1. *The set $\mathcal{L}$ of images under S of all the linked bases relative to H_0 is independent of the choice of S.*

Proof. The set of images under S of the linked bases relative to H_0 is the set of all sequences of the form $\mathcal{B}_0 AS$, for some $A \in Z_V \times H_0$. Let T be another conjugating element, that is, $T^{-1} H_0 T = H$. Then ST^{-1} normalises H_0 and hence lies in $Z_V \times H_0$. Therefore $(Z_V \times H_0)S = (Z_V \times H_0)T$, and so $\mathcal{L} := \{\mathcal{B}_0 AS \mid A \in Z_V \times H_0\} = \{\mathcal{B}_0 AT \mid A \in Z_V \times H_0\}$, proving the lemma. □

Thus the set $\mathcal{L} = \mathcal{L}_0 S$ of images under S of all the linked bases relative to H_0 forms a family of bases for V that is invariant under $Z_V \times H$, and we call bases in this family *linked bases relative to* H.

3.3. Outline of the algorithm

The heart of our solution of the Algorithmic Problem is the construction of a linked basis $\mathcal{B}$ for V relative to H so that, by Lemma 3.1, $\mathcal{B} = \mathcal{B}_0 S$ for some S such that $H = S^{-1} H_0 S$. Once such a basis $\mathcal{B}$ is found, we use it to construct an isomorphism $\lambda : Z_V \times H \to Z_{q-1} \times S_n$ such that, for each $b \in \mathbb{F}^{\#}$ and $A \in H$, $\lambda(bA) = b\lambda(A)$ and $\lambda(A)$ is the permutation corresponding to the action of A on $\mathcal{B}$, or equivalently of $SAS^{-1} \in H_0$ on $\mathcal{B}_0$. We now give a summary of the main steps of the algorithm, and explain where these are presented and analysed in the paper.

Step 1. *Constructing a* 3*-cycle or double transposition*. Since we will use one of these elements to construct the first of the basis vectors, the initial step is to construct a matrix g in H conjugate to a (matrix of H_0 representing a) 3-cycle or double-transposition. Such an element can be obtained as a power of a (matrix in H representing a) pre-3-cycle or pre-double-transposition respectively (see Section 6 for definitions) and, based on some results about polynomials in Section 4, we show that scalar multiples of pre-3-cycles and pre-double-transpositions can be recognised from their characteristic polynomials, provided the characteristic p is not 3 or 2, respectively. In Section 6 we give algorithms to construct a matrix in H conjugate to a pre-3-cycle if $p \neq 3$, or a pre-double-transposition if $p = 3$.

To extract a matrix g corresponding to a 3-cycle or double transposition, we need to determine the orders of these elements. A new algorithm for computing the order of a matrix in H, based on knowing its characteristic polynomial, is given in Section 5, and used to construct a suitable element g in Section 6.

Step 2. *Constructing the first basis vector*. In Section 7, we show first how to construct a conjugate g' of the element g such that the permutations corresponding to g and g' have exactly one moved point in common. We then construct, using g and g', a vector v that lies in some linked basis relative to H.

Step 3. *Constructing a linked basis relative to* H. Extending v to a linked basis relative to H is done in two-stages. For the first stage, see Section 8, the vector v and element g are used to construct an element x of H whose corresponding permutation involves a cycle of prime length $r > 3n/5$, and v, x are then used to construct a linked sequence of vectors of length $r - 1$. Then, in Section 9, this linked sequence is extended to a linked basis relative to H. The reason for employing this two-stage process is that overall it requires asymptotically fewer random selections and matrix operations than the seemingly simpler alternative of finding an n-cycle or $(n-1)$-cycle for this purpose.

Step 4. *Constructing and evaluating the isomorphism*. A procedure is given in Section 10 that constructs λ and evaluates λ on elements of $Z_V \times H$. Evaluating λ^{-1} on elements of $Z_{q-1} \times S_n$ is discussed in Section 10.1.

The various procedures are drawn together in Section 10.1 to complete the proof of Theorem 1.1.

4. Characteristic polynomials

In this section we give some information about the characteristic polynomials on U and V of elements of $Z \times H_0$. We use the following notation throughout the paper.

Notation 4.1. Let $g \in H_0$ be fixed, and suppose that the permutation corresponding to g has cycle lengths $m_1, \ldots, m_l$, where $l \geqslant 1$ and $\sum_i m_i = n$. For each i, write $m_i = p^{a_i} r_i$, where $a_i \geqslant 0$ and r_i is coprime to p, and set

$$m := \sum_i p^{a_i}, \qquad R := \operatorname{lcm}\{r_1, \ldots, r_l\}, \qquad a := \max\{a_1, \ldots, a_l\}.$$

Then $|g| = Rp^a$. Let $b \in \mathbb{F}^{\#}$. Let $c_U^{(b)}(t)$, $c_V^{(b)}(t)$ denote the characteristic polynomials for the actions of bg on U, V respectively, and set $c_U(t) = c_U^{(1)}(t)$ and $c_V(t) = c_V^{(1)}(t)$. For a monic irreducible polynomial $f(t)$ let $\operatorname{mult}^{(b)}(f)$ denote the multiplicity of f in $c_V^{(b)}(t)$.

If $f(t)$ is an irreducible polynomial over $\mathbb{F}$ (our field of order $q = p^a$), then $f(t)$ divides $t^e - 1$ for some positive integer e, and we let

$$e(f) \text{ denote the least } e \text{ such that } f(t) \text{ divides } t^e - 1.$$

Note that if $f(t)$ divides $t^{e_1}-1$ and $t^{e_2}-1$, then $f(t)$ also divides t^e-1 where $e=\gcd(e_1,e_2)$. We usually deal with *monic* polynomials, that is, polynomials $f(t)$ for which the coefficient of the highest power of t occurring is 1.

For $b\in\mathbb{F}^{\#}$ let $f^{(b)}(t)=b^d f(tb^{-1})$ where $d=\deg f$; then $f^{(b)}(t)$ is monic if and only if $f(t)$ is monic, and $f^{(b)}(t)$ is irreducible if and only if $f(t)$ is irreducible. Basic facts about polynomials over finite fields can be found in [21, Section 2.4] and we record some that we shall need in the next lemma.

Lemma 4.2. *Let r,s be positive integers with r coprime to p, let $i\geqslant 0$, and let $b\in\mathbb{F}^{\#}$.*

(a) *Then $t^{p^i r}-1=(t^r-1)^{p^i}$, and the polynomial t^r-1 is a product of distinct irreducible polynomials over $\mathbb{F}$. Moreover, there exists an irreducible $f(t)$ over $\mathbb{F}$ such that $e(f)=r$.*
(b) *If $f(t)$ is a monic irreducible polynomial over $\mathbb{F}$, then*

$$f(t)\mid t^s-1 \quad\Longleftrightarrow\quad f^{(b)}(t)\mid t^s-b^s \quad\Longleftrightarrow\quad e(f)\mid s.$$

We use this information to examine the characteristic polynomials of elements of $Z\times H_0$. Recall that the order of a group element g is denoted by $|g|$; we also denote the multiplicative order of a non-zero element $b\in\mathbb{F}$ by $|b|$.

Lemma 4.3. *Suppose that Notation* 4.1 *holds. Then*

(a) $c_U(t)=\prod_{i=1}^{l}(t^{m_i}-1)=\prod_{i=1}^{l}(t^{r_i}-1)^{p^{a_i}}$, *and* $c_V(t)=c_U(t)/(t-1)^{\delta}$.
(b) *The characteristic polynomials for bg on U and V are $c_U^{(b)}(t)=b^n c_U(t/b)$ and $c_V^{(b)}(t)=b^{n-\delta}c_V(t/b)$, respectively.*
(c) *Let $f(t)$ be a monic irreducible polynomial over $\mathbb{F}$. Then*

$$\operatorname{mult}^{(1)}(f)=\operatorname{mult}^{(b)}\bigl(f^{(b)}\bigr)=\begin{cases}\sum_{\{i:\,e(f)\mid r_i\}}p^{a_i} & \text{if } f(t)\neq t-1,\\ m-\delta & \text{if } f(t)=t-1.\end{cases}$$

In particular, if $f(t)\neq t-1$, then $\operatorname{mult}^{(1)}(f)=1$ if and only if there exists a unique integer i such that r_i is divisible by $e(f)$, and for this i we have $a_i=0$.
(d) *If $g'\in\mathrm{GL}(V)$ is any element with characteristic polynomial equal to $c_V(t)$, then $|g'|=Rp^{a'}$ for some $a'\geqslant 0$.*

Proof. The characteristic polynomial $c_U(t)$ is equal to $\prod_{i=1}^{l}(t^{m_i}-1)$, and the second expression for it given in (a) follows from Lemma 4.2. Since g acts trivially on both E and U/W, it follows that $c_V(t)$ is as asserted. It is straightforward to check that the characteristic polynomials for bg on U and V are as in (b).

Let $f(t)$ be an irreducible polynomial over $\mathbb{F}$ and let r be a positive integer coprime to p. By Lemma 4.2, $f(t)$ divides t^r-1 if and only if $f^{(b)}(t)$ divides t^r-b^r if and only if $e(f)$ divides r, and in this case its multiplicity in t^r-1 is 1. The values of the multiplicities follow from these observations.

Finally suppose that $g' \in \mathrm{GL}(V)$ has characteristic polynomial $c_V(t)$. Let $\mathbb{F}'$ be the splitting field of $c_V(t)$, so that $c_V(t) = \prod_{i=1}^{n-\delta}(t-\zeta_i)$ for some $\zeta_i \in \mathbb{F}'$. We may regard g' as an element of $\mathrm{GL}(n-\delta, \mathbb{F}')$, and in this group g' is conjugate to an upper triangular matrix g'' with diagonal entries $\zeta_1, \ldots, \zeta_{n-\delta}$. Hence $|g'| = |g''| = \mathrm{lcm}\{|\zeta_1|, \ldots, |\zeta_{n-\delta}|\}p^{a'}$, for some $a' \geqslant 0$. In particular, $R' := \mathrm{lcm}\{|\zeta_1|, \ldots, |\zeta_{n-\delta}|\}$ is determined by $c_V(t)$. For the special choice of $g' = g$ we see from Notation 4.1 that $R' = R$, and part (d) follows. □

The next lemma gives important information regarding the problem of finding the scalar b from the characteristic polynomial $c_U^{(b)}(t)$ of bg. Note that the polynomial t^2+t+1 is irreducible if and only if $q \equiv 2 \pmod 3$.

Lemma 4.4. *Suppose that Notation* 4.1 *holds, let r be an integer coprime to p and let $c \in \mathbb{F}^{\#}$.*

(a) *Then $t-c$ divides $t^r - b^r$ if and only if $|cb^{-1}|$ divides r, so*

$$\mathrm{mult}^{(b)}(t-c) = \begin{cases} \sum_{\{i:\ |cb^{-1}| \text{ divides } r_i\}} p^{a_i} & \text{if } c \neq b, \\ m-\delta & \text{if } c = b. \end{cases}$$

(b) *For $q \equiv 2 \pmod 3$, t^2+ct+c^2 (is irreducible and) divides $t^r - b^r$ if and only if $3|cb^{-1}|$ divides r, so*

$$\mathrm{mult}^{(b)}\left(t^2+ct+c^2\right) = \sum_{\{i:\ 3|cb^{-1}| \text{ divides } r_i\}} p^{a_i}.$$

Proof. Now $t-c$ divides $t^r - b^r$ if and only if $c^r = b^r$, that is to say, $|cb^{-1}|$ divides r. By Lemma 4.2 the multiplicity of $t-c$ in $t^r - b^r$ is at most 1. It follows from Lemma 4.3 that the value of $\mathrm{mult}^{(b)}(t-c)$ is as claimed. Thus part (a) is proved.

Now suppose that $q \equiv 2 \pmod 3$ and let $f(t) = t^2+t+1$. Then $f(t)$ is irreducible and therefore also $f^{(c)}(t) = t^2+ct+c^2$ is irreducible. Set $d := c^{-1}b$ and note that $|d| = |cb^{-1}|$. Dividing $t^r - d^r$ by t^3-1 gives a remainder $g(t) = t^2 - d^r, t - d^r$ or $1-d^r$ according as $r \equiv 2, 1, 0 \pmod 3$ respectively. We claim that $f(t)$ divides $t^r - d^r$ if and only if $3|d|$ divides r. Since $f(t)$ divides t^3-1, it follows that $f(t)$ divides $t^r - d^r$ if and only if $f(t)$ divides $g(t)$, and this holds if and only if $g(t) = 0$, which is true if and only if $r \equiv 0 \pmod 3$ and $d^r = 1$. Since $|d|$ divides $q-1 \not\equiv 0 \pmod 3$, the latter conditions are equivalent to $3|d|$ divides r, and the claim follows. Finally setting $t = c^{-1}s$ it follows that s^2+cs+c^2 divides $s^r - c^r d^r = s^r - b^r$ if and only if $3|d| = 3|cb^{-1}|$ divides r.

The value for $\mathrm{mult}^{(b)}(t^2+ct+c^2)$ now follows from the fact that $t^r - b^r$ is multiplicity free (see Lemma 4.2). □

We conclude this section by stating some results about the costs of finding the characteristic polynomial of a matrix over $\mathbb{F}$, and the cost of finding all the distinct irreducible factors of small degree of a polynomial over $\mathbb{F}$.

Lemma 4.5 [8, p. 349]. *There is a deterministic algorithm that computes the characteristic polynomial of an $n \times n$ matrix over $\mathbb{F}$ at a cost of $O(n^\omega \log n \rho_F)$.*

There are several methods for factorising polynomials over finite fields, and recent discussions are given in [15] and [26, Chapter 14]. The most efficient methods are non-deterministic, and we use one of these described in [26, Chapter 14] for our complexity estimations.

Lemma 4.6. *Let $f(t), g(t)$ be polynomials of degree at most n with coefficients in a field $\mathbb{F}$ of order q, and with* $\deg g \leqslant \deg f$.

(a) *Then the product $f(t)g(t)$, and also the remainder on dividing $f(t)$ by $g(t)$ can be found at a cost of $O(\rho_F n \log n \log\log n)$.*
(b) *There is a Las Vegas algorithm that will find, for a given $\varepsilon \leqslant 1/2$, all the distinct linear factors of $f(t)$, or all distinct irreducible degree 2 factors of $f(t)$, at a cost of $O(\log(\varepsilon^{-1})\rho_F n \log^2 n \log(nq) \log\log n)$, and with probability of failure at most ε.*

Proof. Part (a) is proved by the results [26, Theorems 8.23 and 9.6] that multiplication or division of polynomials of degree at most n can be performed in $O(n \log n \log\log n)$ field operations.

We use a careful application of the algorithm presented in [26, Algorithm 14.19] for finding the distinct linear factors of $f(t)$. The heart of this algorithm is [26, Algorithm 14.10] that factorises a square-free monic polynomial of degree at most n for which all irreducible factors have the same degree. As explained in the proof of this latter algorithm in [26, Theorem 14.11], the workings of [26, Algorithm 14.10] can be illustrated by a labelled tree, and the probability that it requires at least k levels before succeeding is at most $n^2 2^{-k}$. Thus if [26, Algorithm 14.10] is allowed to run for up to $k = 4 \log n \log(\varepsilon^{-1}) > \log(n^2 \varepsilon^{-1})$ levels of the labelled tree, then the probability that it fails is at most $n^2 2^{-k} < \varepsilon$. With this value of k, the cost of [26, Algorithm 14.19] is $O(\log(\varepsilon^{-1}) n \log^2 n \log(nq) \log\log n)$ field operations in $\mathbb{F}$. For completeness of our proof, we note that the component [26, Algorithm 14.10] of [26, Algorithm 14.19] is only valid for fields of odd order. If q is even then an alternative algorithm is sketched in [26, Exercise 14.16(iii) on p. 399]; this algorithm also can be run the appropriate number of times to give a probability of failure at most $c_1 n^{-c_2}$, and its running time is asymptotically the same as that given for the case of odd q.

Finally, to find the distinct irreducible degree 2 factors of f, we use the above algorithm to find the linear factors of $f(t)$ over $\mathbb{F}$, and use it again to find its linear factors over a quadratic extension field of $\mathbb{F}$. □

5. Finding orders

If Notation 4.1 holds for bg, then g has order Rp^a where $R = \operatorname{lcm}\{r_1, \ldots, r_l\}$, $a = \max\{a_1, \ldots, a_l\}$, and $|bg| = Rp^a |b^{Rp^a}|$. In this section we present an algorithm for finding $|g|$ from $c_V^{(b)}(t)$, under the assumption that b is known.

We write $M(n)$ for the cost of multiplying two matrices in $\mathrm{GL}(V)$, and note that $M(n) = O(n^\omega)$. Throughout we denote $\log_2 x$ by $\log x$ and $\log_e x$ by $\ln x$.

Proposition 5.1. *Let $g \in \mathrm{GL}(V)$ be such that g is conjugate to an element of H_0, and $|g| = Rp^a$ where p does not divide R and $a \geqslant 0$. Also let $b \in \mathbb{F}^\#$ and let $c(t)$ be the characteristic polynomial of bg on V. Then there exists a deterministic algorithm that, given b and $c(t)$,*

(a) *computes R with $4n^2$ field operations;*
(b) *determines whether $R < n^{18\log n}$, and if so computes a, using at most $36M(n)\log^2 n = O(n^\omega(\log n)^2)$ field operations.*

Proof. Without loss of generality we may assume that $g \in H_0$ and that Notation 4.1 holds. Thus $c(t) = c_V^{(b)}(t)$ and $|g| = Rp^a$ with $R = \mathrm{lcm}\{r_1, \ldots, r_l\}$, $a = \max\{a_1, \ldots, a_l\}$. Since b is known, we also know $c_U^{(b)}(t) = c_V^{(b)}(t)(t-b)^\delta$. Consider the following procedure:

1. For each prime $s \leqslant n$, $s \neq p$, find the largest non-negative integer u such that $t^{s^u} - b^{s^u}$ divides $c_U^{(b)}(t)$, and denote this integer by $u(s)$. Define $R' := \prod_s s^{u(s)}$.
2. If $R' \geqslant n^{18\log n}$, then return R' and the assertion that $R \geqslant n^{18\log n}$; else go to Step 3.
3. Compute $h := (bg)^{R'}$, and find the least non-negative integer a' such that $h^{p^{a'}}$ is a scalar matrix. Return R' and a'.

First we show that the p'-part of $|g|$ is equal to the value R' returned by this procedure, and that, if a' is returned, then $a' = a$, and hence $|g| = R'p^{a'}$. Let s be a prime dividing R and let s^u be the highest power of s dividing R. Then s^u divides some r_i, and so $t^{s^u} - b^{s^u}$ divides $t^{r_i} - b^{r_i}$, which divides $c_U^{(b)}(t)$. Hence $u \leqslant u(s)$ and it follows that R divides R'. Conversely by Lemma 4.2 there exists a monic irreducible polynomial $f(t)$ such that $e(f) = u(s)$, so $f^{(b)}(t)$ is irreducible and divides $t^{s^u} - b^{s^u}$. Since $t^{s^u} - b^{s^u}$ divides $c_U^{(b)}(t)$, it follows that $f^{(b)}(t)$ divides $c_U^{(b)}(t)$, so there exists i such that $f^{(b)}(t)$ divides $t^{r_i} - b^{r_i}$. Again by Lemma 4.2, $e(f) = s^{u(s)}$ divides r_i. It follows that R' divides R, and hence $R' = R$. Thus g^R is a p-element and $|g^R|$ is equal to the order of h modulo scalars, that is to $p^{a'}$, so $a' = a$ and $|g| = R'p^{a'}$.

Now we need to determine the number of field operations required by the various steps of the procedure. The cost of finding R may be computed as follows. By the Prime Number Theorem there are $O(n/\log n)$ primes s to be considered in Step 1. In fact, using Chebyshev's estimates (see [23, Corollary 8.6]) the number of primes s is strictly less than $1.171n/\ln n = (1.171\log e)n/\log n$. For each s, since $s^{u(s)} \leqslant n$, the number of integers u for which we must test whether $t^{s^u} - b^{s^u}$ divides $c_U^{(b)}(t)$ is at most $\log n/\log s$. Each of these divisions requires at most $2n$ field operations. Thus finding $u(s)$ requires at most $2n\log n/\log s$ field operations, and so determining R requires fewer than cn^2 field operations, where $c = 1.171 \times \log e \times 2 < 4$.

If h is computed, its computation requires at most $2\log R$ matrix multiplications and hence requires at most $2M(n)\log R$ field operations. For each i we need at most $2\log p$

matrix multiplications to compute h^{p^i} from $h^{p^{i-1}}$, and hence determining a requires at most $M(n)(2\log R + 2a\log p) = 2M(n)\log(Rp^a)$ field operations. Since $R \leqslant n^{18\log n}$, computing a requires at most $36M(n)\log^2 n$ field operations. □

Remark 5.2. For all $k > 0$, and for all sufficiently large n, most elements of S_n have order less than $n^{k\log n}$. Indeed by [4, Theorem 4.1], the probability that a random permutation in S_n has order greater than $n^{18\log n}$ is less than n^{-7}. Thus the algorithm of the proposition will compute the order of g in almost all cases. For any $g \in S_n$, the p'-part R of $|g|$ satisfies $\log R = (1 + o(1))(n\log n)^{1/2}$ by [18, p. 222], and hence performing Step 3 for any value of R would compute a in $2(1 + o(1))M(n)(n\log n)^{1/2} = O(n^{\omega+1/2}(\log n)^{1/2})$ field operations.

6. 3-cycles and double-transpositions

Here we discuss the problem of finding elements $bg \in \mathrm{GL}(V)$ such that $b \in \mathbb{F}^{\#}$ and g is conjugate to a 3-cycle (an element of type $1^{n-3}3^1$) or a double transposition (an element of type $1^{n-4}2^2$) in H_0. The proportions of 3-cycles and double-transpositions in S_n or A_n are so small that we cannot easily find such elements by random selection from S_n or A_n. Instead we search for elements bg with g conjugate to an element of a larger subset of S_n such that certain powers give us 3-cycles or double-transpositions. These elements are defined as follows.

Definition 6.1. A *pre*-3-*cycle* is an element $\sigma \in S_n$ of order $3f$, where f is not divisible by 3, such that σ^f is a 3-cycle. A *pre-double-transposition* in S_n is an element $\sigma \in S_n$ of order $2f$ with f odd such that σ^f is a double transposition.

It turns out that, for almost all values of n and q, whenever $bg \in \mathrm{GL}(V)$ with $b \in \mathbb{F}^{\#}$ and g conjugate to a pre-3-cycle or pre-double-transposition, we can prove that the elements are of this form by examining their characteristic polynomials. We verify this assertion and then give a Las Vegas algorithm for constructing such elements. Suppose that g is conjugate to an element of H_0, and let $b \in \mathbb{F}^{\#}$. First we show that knowledge of both b and the characteristic polynomial $c_V^{(b)}(t)$ of bg on V allows us to detect whether or not g is conjugate to a pre-3-cycle or pre-double-transposition. Since conjugate matrices have the same characteristic polynomials it is sufficient to prove this property for $g \in H_0$.

Proposition 6.2. *Let $g \in H_0$, $b \in \mathbb{F}^{\#}$, as in Notation* 4.1.

(1) *If $p \neq 3$, then g is a pre-*3*-cycle if and only if*
   (a) $\mathrm{mult}^{(b)}(f) = 1$ *for each irreducible divisor $f(t)$ of $t^2 + bt + b^2$; and*
   (b) *for all primes $r \leqslant n/3$, $r \neq p$, $c_V^{(b)}(t)$ is not divisible by $t^{2r} + b^r t^r + b^{2r}$.*

(2) *If $p = 3$, then g is a pre-double-transposition if and only if*
   (a) $\mathrm{mult}^{(b)}(t + b) = 2$; *and*
   (b) *for all primes r such that $r = 2$, or $5 \leqslant r < n/2$, $c_V^{(b)}(t)$ is not divisible by $t^r + b^r$.*

Proof. Suppose that $p \neq 3$. Then $t-b$ does not divide t^2+bt+b^2 or $t^{2r}+b^rt^r+b^{2r}$ for any prime $r \neq p$. Suppose first that g is a pre-3-cycle. We may assume that $m_1 = r_1 = 3$ and r_i is coprime to 3 for $i > 1$. Then by Lemma 4.3(c), $\mathrm{mult}^{(b)}(f) = 1$ for each irreducible divisor $f(t)$ of t^2+bt+b^2 (since $f(t) \neq t-b$). Suppose that, for some prime $r \leqslant n/3$, $r \neq p$, $c_V^{(b)}(t)$ is divisible by $t^{2r}+b^rt^r+b^{2r} = (t^{3r}-b^{3r})/(t^r-b^r)$. By Lemma 4.2, there exists a monic irreducible $f(t)$ such that $e(f) = 3r$. For such an $f(t)$, $f^{(b)}(t)$ divides $(t^{3r}-b^{3r})/(t^r-b^r)$, and hence divides $c_V^{(b)}(t)$. Therefore $f^{(b)}(t)$ divides $t^{r_i}-b^{r_i}$ for some $i > 1$, and hence $e(f) = 3r$ divides r_i, which is a contradiction. Thus (a) and (b) of part 1 hold.

Conversely suppose that conditions (a) and (b) of part (1) hold. If $f(t)$ is an irreducible factor of t^2+bt+b^2, then $f(t) = h^{(b)}(t)$ where $e(h) = 3$, and since $f(t)$ divides $c_V^{(b)}(t)$, $f(t) = h^{(b)}(t)$ divides $t^{r_i}-b^{r_i}$ for some i. By Lemma 4.2, $e(h) = 3$ divides r_i. Since $\mathrm{mult}^{(b)}(f) = 1$, it follows from Lemma 4.3 that $a_i = 0$ and for all $j \neq i$, r_j is coprime to 3. If $r_i > 3$ then $r_i/3$ is divisible by some prime $r \neq p$, and hence $c_V^{(b)}(t)$ is divisible by $t^{2r}+b^rt^r+b^{2r} = (t^{3r}-b^{3r})/(t^r-b^r)$, contradicting (b). Hence $r_i = 3$, and so g is a pre-3-cycle.

Now suppose that $p = 3$. If g is a pre-double-transposition, then an analogous argument to the first paragraph of the proof shows that conditions (a) and (b) in part (2) hold. Conversely suppose that conditions (a) and (b) of part (2) hold for $c_V^{(b)}(t)$. Since $f(t) := t+b$ has multiplicity 2 in $c_V^{(b)}(t)$ and since $e(f) = 2$, we may assume that f divides $t^{r_i}-b^{r_i}$, with $a_i = 0$ and r_i even, for $i = 1, 2$, and that r_i is odd for $i > 2$. If $r_i > 2$, for $i = 1$ or 2, then $r_i/2$ is divisible by a prime $r \neq 3$ with $2r \leqslant r_i \leqslant n-2$, and so $r < n/2$ and $c_V^{(b)}(t)$ is divisible by $(t^{2r}-b^{2r})/(t-b)$, which contradicts condition (b). Hence $r_1 = r_2 = 2$, and so g is a pre-double-transposition. □

6.1. Finding the scalar: theory

For almost all values of n and q, it turns out that, for all pre-3-cycles and pre-double-transpositions g, we can determine the scalar $b \in \mathbb{F}^{\#}$ from the characteristic polynomial $c_V^{(b)}(t)$ of bg on V. First we deal with pre-3-cycles. In this case, the scalar b can be identified for all $n \geqslant 5$ except the case $(n, p) = (5, 5)$, $q \equiv 1 \pmod 3$, in which case we can only find b^3.

Proposition 6.3. *Let $g \in H_0$ with cycle lengths $m_i = r_i p^{a_i}$ $(1 \leqslant i \leqslant l)$, $b \in \mathbb{F}^{\#}$, and $c_V^{(b)}(t)$ be as Notation* 4.1. *Suppose that $p \neq 3$, $n \geqslant 5$, and g is a pre-*3*-cycle with $m_1 = 3$.*

(1) *If $q \equiv 2 \pmod 3$, and $\mathcal{C}$ is the set of all $c \in \mathbb{F}^{\#}$ such that t^3-c^3 divides $c_V^{(b)}(t)$, then* $\mathrm{mult}^{(b)}(t^2+bt+b^2) = 1$ *and either*
   (a) $\mathcal{C} = \{b\}$; *or*
   (b) $\mathcal{C} = \emptyset$, $\delta = 2$, $n \not\equiv 0 \pmod 3$,

$$c_V^{(b)}(t) = \left(t^2+bt+b^2\right)\left(t^{n-4}+bt^{n-5}+\cdots+b^{n-4}\right)$$

and if p is odd, then the coefficient of t^{n-3} in $c_V^{(b)}(t)$ is $2b$, while if $p=2$, then the coefficient of t^{n-4} in $c_V^{(b)}(t)$ is b^2 (which determines b uniquely).

(2) *If $q \equiv 1 \pmod 3$, and $\mathcal{C}$ is the set of all $c \in \mathbb{F}^{\#}$ such that $t - cy^i$ divides $c_V^{(b)}(t)$ for $i = 0$ and for at least one of $i = 1$ and $i = 2$ (where $y \in \mathbb{F}^{\#}$, $|y| = 3$), then precisely one of* (a)–(c) *holds*:

(a) $\mathcal{C} = \{by, by^2\}$, $\operatorname{mult}^{(b)}(t - by) = \operatorname{mult}^{(b)}(t - by^2) = 1$, *and* $\operatorname{mult}^{(b)}(t - b) = 0$; *or*

(b) $\mathcal{C} = \{b, by, by^2\}$, $\operatorname{mult}^{(b)}(t - by) = \operatorname{mult}^{(b)}(t - by^2) = 1$, *and* $\operatorname{mult}^{(b)}(t - b) > 1$;

(c) $\mathcal{C} = \{b, by, by^2\}$, $\operatorname{mult}^{(b)}(t - by^i) = 1$ *for each i. In this case, $p \geqslant 5$ if $\delta = 2$. Moreover, exactly one of* (i)–(iii) *holds*:

(i) $l = \delta + 1$, $n \not\equiv \delta - 1 \pmod 3$, *and b is the coefficient of $t^{n-\delta-1}$ in $c_V^{(b)}(t) = (t^2 + bt + b^2)(t^{n-\delta-2} - b^{n-\delta-2})$, which equals*

$$t^{n-\delta} + bt^{n-\delta-1} + b^2t^{n-\delta-2} - b^{n-\delta-2}t^2 - b^{n-\delta-1}t - b^{n-\delta};$$

(ii) $\delta = 2$, $l = 3$, $n = p = 5$, *and $c_V^{(b)}(t) = t^3 - b^3$ (yielding only b^3); or*

(iii) $\delta = 2$, $l = 3$, *each $m_i = r_i \geqslant 2$, and $c_V^{(b)}(t)$ is*

$$\left(t^2 + bt + b^2\right)\left(t^{r_2-1} + bt^{r_2-2} + \cdots + b^{r_2-1}\right)\left(t^{r_3} - b^{r_3}\right),$$

so that $2b$ is the coefficient of t^{n-3} (yielding b since $p \geqslant 5$), the constant term is $-b^{n-2}$ and the coefficient of t is $-2b^{n-3}$.

Proof. Since g is a pre-3-cycle, $m \geqslant l \geqslant 2$ and 3 does not divide r_i for any $i > 1$. First we show that $\mathcal{C} \subseteq \{b\}$ if $q \equiv 2 \pmod 3$, and $\mathcal{C} \subseteq \{b, by, by^2\}$ if $q \equiv 1 \pmod 3$. Suppose that this is not the case. Then $\mathcal{C}$ contains an element c such that $(cb^{-1})^3 \neq 1$. This implies that $t - c$ divides $t^{r_i} - b^{r_i}$ for some $i \geqslant 2$. By Lemma 4.4, $|cb^{-1}|$ divides r_i and hence 3 does not divide $|cb^{-1}|$. If $q \equiv 2 \pmod 3$, then $f(t) = t^2 + ct + c^2$ is irreducible and it follows from Lemma 4.4 that $\operatorname{mult}^{(b)}(f) = 0$, contradicting the fact that $c \in \mathcal{C}$. Similarly if $q \equiv 1 \pmod 3$, then, again using Lemma 4.4, we deduce that $\operatorname{mult}^{(b)}(t - cy) = \operatorname{mult}^{(b)}(t - cy^2) = 0$ since $|cyb^{-1}| = |cy^2b^{-1}| = 3|cb^{-1}|$ does not divide r_i for any i. Hence $c \notin \mathcal{C}$, which is a contradiction.

Suppose that $q \equiv 2 \pmod 3$. Then by Proposition 6.2, $\operatorname{mult}^{(b)}(t^2 + bt + b^2) = 1$ and so part 1(a) holds if $t - b$ divides $c_V^{(b)}(t)$. So assume that $\operatorname{mult}^{(b)}(t - b) = m - \delta = 0$. Then $\mathcal{C} = \emptyset$, $m = \delta = 2 = l$, and hence $n = 3 + r_2 \not\equiv 0 \pmod 3$ and $c_V^{(b)}(t)$ is as in part 1(b). Thus

$$\begin{aligned} c_V^{(b)}(t) &= \frac{(t^3 - b^3)(t^{r_2} - b^{r_2})}{(t - b)^2} \\ &= t^{n-2} + 2bt^{n-3} + 3b^2t^{n-4} + \cdots + 3b^{n-4}t^2 + 2b^{n-3}t + b^{n-2} \end{aligned}$$

and 1(b) holds. Note that if $p = 2$, then $n \geqslant 6$ since $\delta = 2$ implies that p divides n, and thus the coefficient of t^{n-4} is b^2.

Now suppose that $q \equiv 1 \pmod 3$. By Proposition 6.2, $\operatorname{mult}^{(b)}(t - by) = \operatorname{mult}^{(b)}(t - by^2) = 1$ so $by, by^2 \in \mathcal{C}$. Also, since $\operatorname{mult}^{(b)}(t - b) = m - \delta$, it follows that $b \in \mathcal{C}$ if and

only if $m > \delta$. Thus either 2(a) or 2(b) holds, or $\mathcal{C} = \{b, by, by^2\}$ and $\mathrm{mult}^{(b)}(t - by^i) = 1 = m - \delta$ for each i. Assume the latter. If $\delta = 1$, then $m = l = 2$ so $n = 3 + r_2 \not\equiv 0 \pmod 3$, and part 2(c)(i) holds. Now assume that $\delta = 2$. Then $m = 3$ and so l is 2 or 3. Suppose that $l = 2$. Then $m = p^{a_1} + p^{a_2} = 1 + p^{a_2}$ so $p = 2$ and $a_2 = 1$. In particular, $m_2 = n - 3 = r_2 p^{a_2}$ is even. However, since $\delta = 2$, $p = 2$ divides n and hence $n - 3$ is odd, which is a contradiction. Hence $l = 3 = m$ so each $a_i = 0$. If $p = 2$, this means that each m_i is odd and hence $n = m_1 + m_2 + m_3$ is odd; but $\delta = 2$, so $p = 2$ divides n, a contradiction. Thus p is odd and as $p \neq 3$, it follows that $p \geqslant 5$. The cycle lengths of g are $3, r_2, r_3$ where $1 \leqslant r_2 \leqslant r_3$, and because g is a pre-3-cycle, we have that 3 does not divide r_2 or r_3. In particular,

$$c_V^{(b)}(t) = \frac{(t^3 - b^3)(t^{r_2} - b^{r_2})(t^{r_3} - b^{r_3})}{(t - b)^2}.$$

If $n = 5$, then $r_2 = r_3 = 1$ and, since p divides n, p must be 5, so part 2(c)(ii) holds. If $n \geqslant 6$ and $r_2 = 1$, then $n = 4 + r_3 \not\equiv 1 \pmod 3$ and hence part 2(c)(i) holds. Finally, if $r_2 > 1$, then part 2(c)(iii) holds. □

Now we deal with pre-double-transpositions. In this case the scalar b can be identified if $n \geqslant 5$ unless $n = 4 + \delta$, where sometimes we can only identify $\{b, -b\}$.

Proposition 6.4. *Suppose that $p = 3$, $n \geqslant 5$, and that $g \in H_0$ with cycle lengths $m_i = r_i p^{a_i}$ $(1 \leqslant i \leqslant l)$, $b \in \mathbb{F}^{\#}$, and $c_V^{(b)}(t)$ are as Notation* 4.1. *Suppose further that g is a pre-double-transposition with $m_1 = m_2 = 2$. Let $\mathcal{C}$ be the set of all $c \in \mathbb{F}^{\#}$ such that $t^2 - c^2$ divides $c_V^{(b)}(t)$. Then $\mathcal{C} = \{b, -b\}$, $\mathrm{mult}^{(b)}(t + b) = 2$, and one of the following holds*:

(a) $\mathrm{mult}^{(b)}(t - b) \neq 2$; *or*
(b) $\mathrm{mult}^{(b)}(t - b) = 2$, $\delta = 1$, $n \equiv 5 \pmod 6$,

$$c_V^{(b)}(t) = (t - b)(t + b)^2\big(t^{n-4} - b^{n-4}\big),$$

and if $n > 5$, then b is the coefficient of t^{n-2} in $c_V^{(b)}(t)$; or
(c) $\mathrm{mult}^{(b)}(t - b) = 2$, $\delta = 2$, $n \equiv 0 \pmod 6$, $m = l = 4$, *and one of* (i)–(iii) *holds*:
  (i) $n = 6$, $c_V^{(b)}(t) = t^4 + b^2 t^2 + b^4$;
  (ii) $n > 6$ *and $b, -b^2, -b^3$ are the coefficients of $t^{n-3}, t^{n-4}, t^{n-5}$ respectively in*

$$c_V^{(b)}(t) = (t - b)(t + b)^2\big(t^{n-5} - b^{n-5}\big);$$

  (iii) $n > 6$ *and $-b, b^2, 0$ are the coefficients of $t^{n-3}, t^{n-4}, t^{n-5}$ respectively in*

$$c_V^{(b)}(t) = (t + b)^2\big(t^{r_3} - b^{r_3}\big)\big(t^{r_4} - b^{r_4}\big),$$

where r_3, r_4 are odd and at least 5.

Proof. Since g is a pre-double-transposition and $n \geqslant 5$, we have $m \geqslant l \geqslant 3$ and r_i is odd for all $i > 2$. Thus $\mathrm{mult}^{(b)}(t-b) = m - \delta > 0$. By Proposition 6.2, $\mathrm{mult}^{(b)}(t+b) = 2$, and hence both b and $-b$ lie in $\mathcal{C}$. Suppose that $c \neq \pm b$ and $c \in \mathcal{C}$. Then at least one of cb^{-1} and $-cb^{-1}$ has order $2s$ for some $s > 1$. It follows from Lemma 4.4 that $2s$ divides r_i for some $i \geqslant 3$ which is a contradiction. Thus $\mathcal{C} = \{b, -b\}$.

Suppose now that $\mathrm{mult}^{(b)}(t-b) = 2$, that is, $m = \delta + 2$. If $\delta = 1$, then $m = l = 3$ so $n = 4 + r_3$ is odd and $a_3 = 0$; thus n and $n-4$ are coprime to 6, and it follows that $n \equiv 5 \pmod 6$, and $c_V^{(b)}(t)$, b are as in (b). If $\delta = 2$, then $m = l = 4$, $a_3 = a_4 = 0$, so $n = 4 + r_3 + r_4$ is even and hence $n \equiv 0 \pmod 6$. Let us suppose that $r_3 \leqslant r_4$. If $n = 6$, then $c_V^{(b)}(t) = (t-b)^2(t+b)^2$ as in (c)(i), so assume that $n > 6$. If $r_3 = 1$, then $c_V^{(b)}(t) = (t-b)(t+b)^2(t^{n-5} - b^{n-5})$ and $b, -b^2, -b^3$ are the coefficients of $t^{n-3}, t^{n-4}, t^{n-5}$, respectively, as in (c)(ii). If $r_3 > 1$, then since r_3, r_4 are coprime to 6, they are both at least 5, and $c_V^{(b)}(t)$ is as in (c)(iii). □

6.2. Proportions

In our algorithms we will construct a 3-cycle or double transposition from a pre-3-cycle or pre-double-transposition respectively that has reasonably small order, namely order at most $n^{18\log n}$. We give here estimates for the proportions of such elements in A_n and S_n.

Definition 6.5. Let $p_{\mathrm{pre3}}^{\mathrm{small}\,A}(n)$, $p_{\mathrm{pre3}}^{\mathrm{small}\,S}(n)$ denote the proportions of pre-3-cycles in A_n and S_n, respectively, that have order less than $n^{18\log n}$. Let $p_{\mathrm{pre22}}^{S}(n)$, $p_{\mathrm{pre22}}^{A}(n)$ denote the proportions of pre-double-transpositions in S_n and A_n, respectively, and let $p_{\mathrm{pre22}}^{\mathrm{small}\,A}(n)$ and $p_{\mathrm{pre22}}^{\mathrm{small}\,S}(n)$ denote the proportions of such elements that have order less than $n^{18\log n}$.

Lemma 6.6.

(a) [4, Theorem 5.2] *For* $n \geqslant 5$, $p_{\mathrm{pre3}}^{\mathrm{small}\,A}(n) > 0.140n^{-1/3}$ *and* $p_{\mathrm{pre3}}^{\mathrm{small}\,S}(n) > 0.282n^{-1/3}$.
(b) $p_{\mathrm{pre22}}^{S}(n) = \frac{1}{\sqrt{32\pi n}} + O(n^{-3/2})$ *and* $p_{\mathrm{pre22}}^{A}(n) = 2p_{\mathrm{pre22}}^{S}(n)$.
(c) *For* $n \geqslant 5$, $p_{\mathrm{pre22}}^{\mathrm{small}\,A}(n) > 0.0997n^{-1/2}$ *and* $p_{\mathrm{pre22}}^{\mathrm{small}\,S}(n) > 0.0498n^{-1/2}$.

Proof. (b) A pre-double-transposition $g \in S_n$ is of the form $g = (i,j)(k,l)h$, where i, j, k, l are distinct points fixed by h, and $|h|$ is odd. There are $3\binom{n}{4}$ possibilities for choosing a double transposition $(i,j)(k,l)$, and for a given choice there are $(n-4)!s_{\neg 2}(n-4)$ elements h of odd order on the remaining points, where $s_{\neg 2}(n)$ denotes the proportion of elements of S_n of odd order. Hence $p_{\mathrm{pre22}}^{S}(n) = s_{\neg 2}(n-4)/8$, and so by [4, Theorem 2.3(c)], $p_{\mathrm{pre22}}^{S}(n) = c(2)(n-4)^{-1/2}/8 + O(n^{-3/2}) = c(2)n^{-1/2}/8 + O(n^{-3/2})$, where $c(2) = (\pi/2)^{-1/2} \cong 0.798$, as claimed. Since all pre-double-transpositions are even permutations it follows that $p_{\mathrm{pre22}}^{A}(n) = 2p_{\mathrm{pre22}}^{S}(n)$.

(c) By [4, Theorem 4.1], the proportion $p_{\mathrm{small}}(n)$ of elements of S_n of order greater than $n^{18\log n}$ is less than n^{-7}. Also by [4, Theorem 2.3(a) and (b)], $p_{\mathrm{pre22}}^{S}(n) = s_{\neg 2}(n-4)/8 \geqslant s_{\neg 2}(n)/8 \geqslant c(2)n^{-1/2}(1-n^{-1})/8$, for all $n \geqslant 5$. Thus

$$p_{\text{pre22}}^{\text{small}\,S}(n) \geqslant p_{\text{pre22}}^{S}(n) - p_{\text{small}}(n)$$
$$> \frac{c(2)}{8n^{1/2}}\left(1-\frac{1}{n}\right) - \frac{1}{n^7} > \frac{c(2)}{16n^{1/2}} > 0.0498n^{-1/2}.$$

Similarly

$$p_{\text{pre22}}^{\text{small}\,A}(n) \geqslant p_{\text{pre22}}^{A}(n) - 2p_{\text{small}}(n)$$
$$> \frac{c(2)}{4n^{1/2}}\left(1-\frac{1}{n}\right) - \frac{2}{n^7} > \frac{c(2)}{8n^{1/2}} > 0.0997n^{-1/2}. \qquad \square$$

6.3. Procedures

In this subsection we give procedures for finding 3-cycles and double-transpositions based on the results above. For polynomials $f(t), c(t)$ over $\mathbb{F}$ with $f(t)$ irreducible, we denote by $\text{mult}_{c(t)}(f)$ the multiplicity of $f(t)$ in $c(t)$. First we find a 3-cycle in the case where the characteristic is not 3. If $p = 3$, then this procedure will not work, and in this case we use the similar Procedure 6.9 to find a double transposition.

Procedure 6.7 (FIND3CYCLE). *We are given a positive constant ε, and a subgroup $G \leqslant \mathrm{GL}(V) = \mathrm{GL}(d,q)$ (where q is a power of p and $p \neq 3$) such that $H' \leqslant G \leqslant Z_V \times H$ where H is conjugate to $H_0 \cong S_n$ and $n \geqslant 5$, $(n,p) \neq (5,5)$.*

1. *Select up to $\lceil \log(\varepsilon^{-1})n^{1/3}/0.07 \rceil$ random elements $x \in G$, and perform the following steps for each.*
2. *Find the characteristic polynomial $c(t)$ of x.*

3.1. *For $q \equiv 2 \pmod 3$, determine the subset $\mathcal{C}$ of elements $c \in \mathbb{F}^{\#}$ such that $t^3 - c^3$ divides $c(t)$.*
   (i) *If $|\mathcal{C}| \geqslant 2$, then return to Step* 1.
   (ii) *If $\mathcal{C} = \{c\}$, then if $\text{mult}_{c(t)}(t^2 + ct + c^2) = 1$ let $b = c$, and otherwise return to Step* 1.
   (iii) *If $\mathcal{C} = \emptyset$, then return to Step* 1 *unless $\delta = 2$ and $n \not\equiv 0 \pmod 3$. Let $c \in \mathbb{F}$ be such that the coefficient of t^{n-3} is $2c$ if p is odd, or the coefficient of t^{n-4} is c^2 if $p = 2$. If*

$$c(t) = \left(t^2 + ct + c^2\right)\left(t^{n-3} - c^{n-3}\right)/(t-c),$$

   then set $b = c$, and otherwise return to Step 1.

3.2. *For $q \equiv 1 \pmod 3$, determine the subset $\mathcal{C}$ of elements $c \in \mathbb{F}^{\#}$ such that $\text{mult}_{c(t)}(t-c) > 0$, and also $\text{mult}_{c(t)}(t - cy^i) > 0$ for at least one $i \in \{1,2\}$, where $|y| = 3$.*
   (i) *If $|\mathcal{C}| \notin \{2,3\}$, then return to Step* 1.
   (ii) *If $\mathcal{C} = \{c_1, c_2\}$, then if $|c_1 c_2^{-1}| = 3$, and $\text{mult}_{c(t)}(t - c_i) = 1$ for $i = 1,2$, set $b = c_1^2 c_2^{-1}$, and otherwise return to Step* 1.

(iii) *If* $|\mathcal{C}| = 3$, *then return to Step* 1 *unless* $y\mathcal{C} = \mathcal{C}$, *and* $\mathcal{C} = \{c_1, c_2, c_3\}$ *with* $\mathrm{mult}_{c(t)}(t - c_1) = \mathrm{mult}_{c(t)}(t - c_2) = 1 \leqslant \mathrm{mult}_{c(t)}(t - c_3)$.
For the case $\mathrm{mult}_{c(t)}(t - c_3) > 1$: *set* $b = c_3$.
For the case $\mathrm{mult}_{c(t)}(t - c_3) = 1$ *and* $\delta = 1$: *if* $n \not\equiv 0 \pmod 3$, *the coefficient* c *of* t^{n-2} *in* $c(t)$ *lies in* $\mathcal{C}$, *and*

$$c(t) = \bigl(t^2 + ct + c^2\bigr)\bigl(t^{n-3} - c^{n-3}\bigr),$$

then set $b = c$, *and otherwise return to Step* 1.
For the case $\mathrm{mult}_{c(t)}(t - c_3) = 1$ *and* $\delta = 2$: *return to Step* 1 *unless* $p \geqslant 5$. *Let* $c \in \mathbb{F}$ *be the coefficient of* t^{n-3} *in* $c(t)$. *If* $n \not\equiv 1 \pmod 3$, $c \in \mathcal{C}$, *and*

$$c(t) = \bigl(t^2 + ct + c^2\bigr)\bigl(t^{n-4} - c^{n-4}\bigr),$$

then set $b = c$. *Otherwise if* $a \in \mathbb{F}^{\#}$ *is such that* $c = 2a$, *and we have* $a \in \mathcal{C}$, *the constant term in* $c(t)$ *is* $-a^{n-2}$ *and the coefficient of* t *is* $-2a^{n-3}$, *then set* $b = a$ *and otherwise return to Step* 1.

4. *If there exists a prime* r *such that* $r \leqslant n/3$, $r \neq p$, *and* $t^{2r} + b^r t^r + b^{2r}$ *divides* $c(t)$, *then return to Step* 1.
5. *Set* $g = b^{-1}x$; *by the procedure in Proposition* 5.1, *determine whether the* p'*-part of* $|g|$ *is at most* $n^{18\log n}$ *and if so find* $|g| = Rp^v$. *If either* $|g| > n^{18\log n}$, *or* $|g| \leqslant n^{18\log n}$ *and* 3 *does not divide* R, *then return to Step* 1. *Otherwise compute* $g^{Rp^v/3}$ *and return this element.*
6. *If no element is returned at Step* 5 *for any of the random elements* x, *then report* FAILURE.

We prove that this procedure is valid and estimate its complexity. Recall that ξ is an upper bound for the cost of constructing a random element, $O(\rho_F n^\omega)$ is taken as the cost of multiplying two $n \times n$ matrices over $\mathbb{F}$, and ρ_F is an upper bound on the cost of a field operation in $\mathbb{F}$.

Lemma 6.8. *Suppose that* $n \geqslant 5$ *and* $(n, p) \neq (5, 5)$. *Then, with probability at least* $1 - \varepsilon$, *Procedure* 6.7 (FIND3CYCLE) *returns an element of* H *conjugate to a* 3*-cycle in* H_0. *It is a Las Vegas algorithm and runs at a cost of* $O((\log \varepsilon^{-1})(\xi n^{1/3} + \rho_F n^{1/3} \log^2 n(n^\omega + n \log(nq) \log\log n)))$. *This is* $O((\log \varepsilon^{-1})(\xi n^{1/3} + \rho_F n^{\omega+1/3} \log^2 n \log q))$.

Proof. It follows from Lemma 6.6 that the proportion of elements $bg \in G$ such that $b \in \mathbb{F}^{\#}$ and the permutation corresponding to g is a pre-3-cycle of order at most $n^{18\log n}$ is greater than $0.14n^{-1/3}$. For each random element we apply the Las Vegas algorithm in Lemma 4.6 with probability of failure at most $0.07n^{-1/3}$ to find the distinct linear factors of its characteristic polynomial. Thus, for each random element, the probability that it is a pre-3-cycle of order at most $n^{18\log n}$, and that in addition we succeed in finding its distinct linear factors, is at least $0.07n^{-1/3}$. It follows that the probability of failing to find such an element, and its linear factors, after N independent random selections from G is less than $(1 - 0.07n^{-1/3})^N$ and this quantity is less than ε provided that $N \geqslant \log(\varepsilon^{-1})n^{1/3}/0.07$.

Thus by taking $N = \lceil \log(\varepsilon^{-1}) n^{1/3}/0.07 \rceil$, the procedure will find such an element, and its linear factors, with probability greater than $1 - \varepsilon$.

Moreover, if bg is selected, where $b \in \mathbb{F}^{\#}$ and g is a pre-3-cycle of order at most $n^{18 \log n}$, then it follows from Proposition 6.3 that b is identified correctly in Step 3, and from Propositions 6.2 and 5.1 that Steps 4 and 5 respectively are completed successfully. The procedure therefore returns a 3-cycle in $\langle g \rangle$ in this case. Thus we have proved that the probability the procedure reports FAILURE is less than ε. To complete the proof that this is a Las Vegas algorithm we need to prove that whenever an answer is returned it is correct, that is to say, we must prove that any element returned by the procedure is indeed a 3-cycle.

Suppose that an element is returned after testing $x \in G$, where $x = dg$ with $d \in \mathbb{F}^{\#}$ and $g \in H$. Suppose first that Step 3 correctly identifies the scalar $b = d$. In each case, from the definition of b it follows that condition (a) of Proposition 6.2 holds. Also condition (b) of Proposition 6.2 follows from Step 4. Hence g is a pre-3-cycle, and in this case we saw in the previous paragraph that the element returned is a 3-cycle.

It remains to prove that, whenever an element $x = dg$ is returned, then Step 3 correctly defines b as d. What we prove is that, if in processing an element $x = dg$ a scalar b is defined at Step 3, then either $b = d$, or, if not, then the element fails the tests in Step 4. Suppose then that in Step 3 the scalar b is defined and $b \neq d$. Set $z = bd^{-1}$ and $s = |z| > 1$, let the cycle lengths of g be $m_i = r_i p^{a_i}$, for $1 \leqslant i \leqslant l$, and $m = \sum_i p^{a_i}$, as in Notation 4.1.

Suppose first that $q \equiv 2 \pmod 3$. Then whether b is defined in Step 3.1(ii) or Step 3.1(iii), the irreducible polynomial $t^2 + bt + b^2$ divides $c(t)$. Thus, by Lemma 4.4(b), $3s$ divides r_i for some i. Now $3|(bz)d^{-1}| = 3|z^2|$ divides $3|z| = 3s$ which in turn divides r_i. Hence by Lemma 4.4(b), $\mathrm{mult}_{c(t)}(t^2 + (bz)t + (bz)^2) > 0$. Also by Lemma 4.4(a), $\mathrm{mult}_{c(t)}(t - bz) > 0$ either if $bz \neq d$, or if $bz = d$ and $m > \delta$. However, if the latter multiplicity is positive, then $bz \in \mathcal{C}$, which is a contradiction (whether b was defined in Step 3.1(ii) or (iii)). Thus $m = \delta$, and $bz = d$ which implies that $z = b^{-1}d = z^{-1}$. Hence $z^2 = 1$ and since $z \neq 1$, we conclude that $z = -1$ and p is odd, $s = 2$, and $b = -d$. This means that $3s = 6$ divides r_i, and so $c(t)$ is divisible by $(t^6 - d^6)/(t^2 - d^2) = t^4 + d^2t^2 + d^4$ (see Lemma 4.4(b) again). Also, $n \geqslant r_i \geqslant 6$. Thus the prime $r = 2$ satisfies $r \leqslant n/3$, $r \neq p$, and so this element would not pass the test of Step 4, and hence such an element x is never returned.

Therefore $q \equiv 1 \pmod 3$. For each of the possibilities in this case we have $by, by^2 \in \mathcal{C} \subseteq \{b, by, by^2\}$ and $\mathrm{mult}_{c(t)}(t - by) = \mathrm{mult}_{c(t)}(t - by^2) = 1$. Therefore by Lemma 4.4(a), each of $|zy|$ and $|zy^2|$ divides r_i for some i. Therefore, the orders of $(byz)d^{-1} = (zy^2)^2$ and $(by^2z)d^{-1} = (zy)^2$ also divide r_i, and so for $j = 1, 2$, by Lemma 4.4, $\mathrm{mult}_{c(t)}(t - by^j z) > 0$ either if $by^j z \neq d$ or if $by^j z = d$ and $m - \delta > 0$.

Claim. *$d = by$ or by^2, so that $d \in \mathcal{C}$, and $s = 3$ divides r_i for some i.*

If $\mathrm{mult}_{c(t)}(t - by^j z) > 0$ for both $j = 1$ and $j = 2$, then $bzy, bzy^2 \in \mathcal{C}$. Since $z \neq 1$, this implies that $z = y$ or y^2, and so $s = |z| = 3$ and $d = by$ or by^2. In particular $d \in \mathcal{C}$, and $s = 3$ divides r_i for some i. Thus the claim is proved in this case. On the other hand, if for $j = 1$ or 2 we have $\mathrm{mult}_{c(t)}(t - by^j z) = 0$, then (by the observation at the end of

the previous paragraph) $m = \delta$ and $by^j z = d$. Then since $z = bd^{-1}$, this implies that $y^j = b^{-1}dz^{-1} = z^{-2}$, and so $d = by^j z = bz^{-1}$ and $s = |z| = 3$ or 6. If $s = 3$ then $z = z^{-2} = y^j$ and so $d = by^{2j}$ and as in the previous case, $d \in \mathcal{C}$ and the claim is proved. Suppose then that $s = |z| = 6$, so that in particular p is odd. Then $|zy^j| = |z^{-1}| = 6$, and we showed in the previous paragraph that this divides r_i for some i. Thus $n \geqslant r_i \geqslant 6$, and $c(t)$ is divisible by

$$\frac{t^6 - d^6}{t - d} = (t - dz)\bigl(t - dz^2\bigr)\bigl(t - dz^3\bigr)\bigl(t - dz^4\bigr)\bigl(t - dz^5\bigr).$$

However, in this case all five of the elements dz^ℓ, with $1 \leqslant \ell \leqslant 5$, satisfy the condition for membership of the set $\mathcal{C}$, and in such a case, the element would have failed the test at Step 3.2(i), and the scalar b would not have been defined. Thus the claim is proved in all cases.

Since we are assuming that x is returned, we have now that $d = by$ or by^2, that $d \in \mathcal{C}$, and that 3 divides r_i for some i. The last condition implies that $t^2 + dt + d^2 = (t - dy) \times (t - dy^2)$ divides $c(t)$, and so $t - b$ divides $c(t)$. Hence $b \in \mathcal{C}$. Thus $\mathcal{C} = \{b, by, by^2\} = \{b, c, d\}$, where $t^2 + dt + d^2 = (t - b)(t - c)$. By Lemma 4.4, $\mathrm{mult}_{c(t)}(t - b) = \mathrm{mult}_{c(t)}(t - c) = \sum_{3|r_i} p^{a_i}$, and it follows from Step 3.2 that this multiplicity must be 1.

Also by Step 3.2, and since $d \neq b$, it follows that $\mathrm{mult}_{c(t)}(t - d) = 1$ and so by Lemma 4.4, $m - \delta = 1$. If $\delta = 1$, then $m = 2$, $a_1 = a_2 = 0$, and so $c(t) = (t^{r_1} - d^{r_1}) \times (t^{r_2} - d^{r_2})/(t - d)$, and the coefficient of t^{n-2} is 0 if $\min\{r_1, r_2\} = 1$, and d otherwise. However, by the definition of b in Step 3.2(iii), the coefficient of t^{n-2} is b, which is a contradiction. Thus $\delta = 2$, $m = 3$, and all the $a_i = 0$. By Step 3.2(iii), $p \geqslant 5$ and since p divides n, also $n \geqslant 5$. At this stage we have

$$c(t) = \bigl(t^{r_1} - d^{r_1}\bigr)\bigl(t^{r_2} - d^{r_2}\bigr)\bigl(t^{r_3} - d^{r_3}\bigr)/(t - d)^2,$$

and 3 divides r_1, say, and $1 \leqslant r_2 \leqslant r_3$. Also by Step 3.2, the coefficient of t^{n-3} is b or $2b$ (and in particular is non-zero). The coefficient of t^{n-3} in $c(t)$ above is $2d$ (if $r_2 > 1$), d (if $r_2 = 1 < r_3$), or 0 (if $r_2 = r_3 = 1$). Since $b \neq d$ and the coefficient of t^{n-3} is b or $2b$, it follows that $r_3 > 1$, and either $r_2 > 1$ and $b = 2d$, or $r_2 = 1$ and $2b = d$. However $d = by$ or by^2, and hence $b^3 = d^3$. This implies that $p = 7$, and hence $n \geqslant 7$ (since $\delta = 2$). Suppose that $r_2 = 1$ and $2b = d$, so that

$$c(t) = \bigl(t^{r_1} - d^{r_1}\bigr)\bigl(t^{r_3-1} + dt^{r_3-2} + \cdots + d^{r_3-1}\bigr).$$

By Step 3.2, the constant term is $-b^{n-2}$ and the coefficient of t is $-2b^{n-3}$. Comparing the constant term and coefficient of t in $c(t)$ above we get that $-d^{n-2} = -b^{n-2}$ and $-d^{n-3} = -2b^{n-3}$, respectively. Substituting $d = 2b$ in these equations, we find that 7 divides both $2^{n-2} - 1$ and $2^{n-4} - 1$ which is impossible. Thus $r_2 > 1$, $b = 2d$, and

$$c(t) = \bigl(t^{r_1} - d^{r_1}\bigr)\bigl(t^{r_2-1} + dt^{r_2-2} + \cdots + d^{r_2-1}\bigr)\bigl(t^{r_3-1} + dt^{r_3-2} + \cdots + d^{r_3-1}\bigr).$$

By Step 3.2(iii), $c(t) = (t^2 + bt + b^2)(t^{n-4} - b^{n-4})$. Comparing the constant term and coefficient of t in these two expressions for $c(t)$, we get that $-d^{n-2} = -b^{n-2}$ and

$-2d^{n-3} = -b^{n-3}$, respectively. Substituting $b = 2d$ in these equations, we find that 7 divides both $2^{n-2} - 1$ and $2^{n-4} - 1$ which is impossible.

Thus the procedure correctly identifies $b = d$ and we have completed the proof that each element returned is a 3-cycle.

Finally we determine the cost. The cost of processing each of up to $\log(\varepsilon^{-1})n^{1/3}/0.07$ random elements x is as follows. First we compute $c(t)$ at a cost of $O(\rho_F n^\omega \log n)$, see Lemma 4.5. Next we determine the set $\mathcal{C}$: first we compute the set of distinct linear factors of $c(t)$ at a cost of $O(\rho_F n \log^2 n \log(nq) \log\log n)$, see Lemma 4.6. If $q \equiv 2 \pmod 3$, then we either compute the set of quadratic irreducible factors of $c(t)$, and from this determine $\mathcal{C}$, or we determine by division the factors $t^3 - c^3$ of $c(t)$, for which $t - c$ is a linear factor. If $q \equiv 1 \pmod 3$, then we can easily determine $\mathcal{C}$ from the set of linear factors. In either case the cost is less than $O(\rho_F n^\omega)$. (Note that the factorisation algorithm employed is Las Vegas and may fail, as discussed in the first paragraph of the proof.)

If we have been successful in determining $\mathcal{C}$, then determining b costs at most $O(n)$ field operations (as we may need to multiply two polynomials, one of degree 2 and the other of degree at most $n - 3$). To perform Step 4, for each of the $O(n/\log n)$ primes $r \leqslant n/3, r \neq p$, we require $O(\log r)$ field operations to determine $t^{2r} + b^r t^r + b^{2r}$, and then $O(n)$ field operations to check whether it divides $c(t)$, a total of $O(n^2/\log n)$ field operations. Determining g costs $O(n^2)$ field operations, and deciding whether g has order less than $n^{18\log n}$, and if so finding $|g|$ costs $O(n^\omega \log^2 n)$ field operations by Proposition 5.1. Finally extracting the 3-cycle costs another $O(n^\omega \log^2 n)$ field operations. □

It is unfortunate that the procedure above fails when $p = 3$. In this case we have an analogous method based on pre-double-transpositions to construct a double-transposition. This method only fails when $p = 2$, but for simplicity we present it only for $p = 3$. It is based on Propositions 6.2 and 6.4.

Procedure 6.9 (FINDDOUBLETRANSPOSITION). *We are given a positive constant ε, and a subgroup $G \leqslant \mathrm{GL}(V) = \mathrm{GL}(d, q)$ (where $p = 3$) such that $H' \leqslant G \leqslant Z_V \times H$ where H is conjugate to $H_0 \cong S_n$ and $n \geqslant 7$.*

1. *Select up to $\lceil \log(\varepsilon^{-1})n^{1/2}/0.0249 \rceil$ random elements $x \in G$, and perform the following steps for each.*
2. *Find the characteristic polynomial $c(t)$ of x, and the set $\mathcal{C}$ of elements $c \in \mathbb{F}^\#$ such that $t^2 - c^2$ divides $c(t)$.*
3. *If $\mathcal{C}$ is not of the form $\mathcal{C} = \{c, -c\}$, with $\mathrm{mult}_{c(t)}(t + c) = 2$, then return to Step* 1.
   (i) *If $\mathrm{mult}_{c(t)}(t - c) \neq 2$, then let $b = c$.*
   (ii) *If $\mathrm{mult}_{c(t)}(t - c) = 2$ and $\delta = 1$, then if $n \equiv 5 \pmod 6$ and the coefficient d of t^{n-2} in $c(t)$ lies in $\mathcal{C}$ and $c(t) = (t - d)(t + d)^2(t^{n-4} - d^{n-4})$, then set $b = d$; otherwise return to Step* 1.
   (iii) *If $\mathrm{mult}_{c(t)}(t - c) = 2$ and $\delta = 2$, then return to Step* 1 *unless $n \equiv 0 \pmod 6$ and the coefficient d of t^{n-3} in $c(t)$ lies in $\mathcal{C}$.*
   If $c(t) = (t - d)(t + d)^2(t^{n-5} - d^{n-5})$, then set $b = d$.
   If the coefficients of t^{n-4}, t^{n-5} in $c(t)$ are $d^2, 0$ respectively, then set $b = -d$.
   Otherwise return to Step 1.

4. *If there exists a prime r such that $r = 2$ or $5 \leqslant r < n$, and $t^r + b^r$ divides $c(t)$, then return to Step* 1.
5. *Set $g = b^{-1}x$; using the procedure in Proposition* 5.1, *determine whether $|g| \leqslant n^{18\log n}$ and if so find $|g| = R3^v$. If either $|g| > n^{18\log n}$, or $|g| \leqslant n^{18\log n}$ and R is odd, then return to Step* 1. *Otherwise compute $g^{R3^v/2}$ and return this element.*
6. *If no element is returned at Step* 5 *for any of the random elements x then report* FAILURE.

We prove that this procedure is valid and estimate its complexity.

Lemma 6.10. *Suppose that $n \geqslant 7$. Then, with probability at least $1 - \varepsilon$, Procedure* 6.9 (FINDDOUBLETRANSPOSITION) *returns an element of H conjugate to a double-transposition in H_0. It is a Las Vegas algorithm and runs at a cost of at most $O((\log \varepsilon^{-1})(\xi n^{1/2} + \rho_F n^{1/2}\log^2 n(n^\omega + n\log(nq)\log\log n)))$. This is at most $O((\log \varepsilon^{-1})(\xi n^{1/2} + \rho_F n^{\omega+1/2}\log^2 n \log q))$.*

Proof. It follows from Lemma 6.6 that the proportion of elements $bg \in G$ such that $b \in \mathbb{F}^\#$ and the permutation corresponding to g is a pre-double-transposition of order at most $n^{18\log n}$ is greater than $0.0498n^{-1/2}$. As in the first paragraph of the proof of Lemma 6.8, we find the characteristic polynomial $c(t)$ of each random element and use a Las Vegas algorithm to find the distinct linear factors of $c(t)$ with probability of failure less than $0.0249n^{-1/2}$; then by analysing up to $N = \lceil \log(\varepsilon^{-1})n^{1/2}/0.0249 \rceil$ random elements, the procedure will find an element bg with g a pre-double-transposition of order less than $n^{18\log n}$, and will succeed in finding the linear factors of its characteristic polynomial, with probability greater than $1 - \varepsilon$. Moreover, for such an element bg, it follows from Proposition 6.4 that b is identified correctly in Step 3, and from Propositions 6.2 and 5.1 that Steps 4 and 5 respectively are completed successfully. The procedure therefore returns the double-transposition in $\langle g \rangle$ in this case. Thus the probability the procedure reports FAILURE is less than ε. The next step is to prove that any element returned by the procedure is indeed a double-transposition.

Suppose that an element is returned after testing $x \in G$, where $x = cg$ with $c \in \mathbb{F}^\#$ and $g \in H$. Suppose first that in Step 3 the scalar b is defined as $b = c$. Then conditions (a) and (b) of Proposition 6.2.2 follow from the tests in Steps 3 and 4, respectively, and so g is a pre-double-transposition. Thus the element returned by Step 5 is a double transposition.

It remains to prove that, whenever an element $x = cg$ is returned, then Step 3 correctly defines b as c. What we prove is that, if while processing an element $x = cg$ a scalar b is defined in Step 3, then either $b = c$, or, if not, then the element fails the tests in Step 4. Suppose then that in Step 3 the scalar b is defined and $b \neq c$. Set $z = bc^{-1}$ and $s = |z| > 1$, let the cycle lengths of g be $m_i = r_i p^{a_i}$, for $1 \leqslant i \leqslant l$, and $m = \sum_i p^{a_i}$, as in Notation 4.1. Note that if s is odd, then $|-z| = 2s$. By Step 3 and the definition of $\mathcal{C}$, we have $\mathcal{C} = \{b, -b\}$, $\mathrm{mult}_{c(t)}(t+b) = 2$, and $\mathrm{mult}_{c(t)}(t-b) > 0$. Then by Lemma 4.4(a), $s = |z|$ divides r_i for some i, and if s is odd, then also $2s = |-z|$ divides r_i for some i. Without loss of generality we may assume that $\mathrm{lcm}\{2, s\}$ divides r_1. Suppose first that $s = |z| > 2$, which means that $bz \neq c$. Now $(bz)c^{-1} = z^2$, so $|(bz)c^{-1}|$ divides r_1 and hence, by Lemma 4.4, $\mathrm{mult}_{c(t)}(t - bz) > 0$ (since $bz \neq c$). If s is odd, then $|-z^2| = 2|z| = 2s$ which divides r_1;

also if s is even and greater than 4, then $-bz \neq c$ and $1 \neq (-bz)c^{-1} = -z^2 \in \langle z \rangle$, and so $|-z^2|$ divides r_1. Thus in either of these cases $|(-bz)c^{-1}| > 1$ and divides r_1, and $-bz \neq c$, and hence, by Lemma 4.4, $\mathrm{mult}_{c(t)}(t + bz) > 0$. It follows that $bz, -bz \in \mathcal{C}$, which is a contradiction. Thus $s = 4$, and so $c = bz^{-1} = -bz$ and $b^4 = c^4$. Since 4 divides r_i, it follows that $t^4 - c^4 = t^4 - b^4$ divides $t^{r_1} - c^{r_1}$ and hence $t^2 + b^2$ divides $c(t)$, and so, if $s > 2$ then the element $x = cg$ fails the test in Step 4.

Assume now that $s = 2$. Then r_1 is even, and $b = -c$. Since $\mathrm{mult}_{c(t)}(t + b) = 2$, it follows from Lemma 4.4 that $m - \delta = 2$, so $m = \sum_i 3^{a_i} = 2 + \delta \leqslant 4$. We claim that all the $a_i = 0$. If this is not so then a unique $a_i = 1$; if $\delta = 1$ then the number of cycles $l = 1$ and so 3 divides $r_1 = n$ contradicting $\delta = 1$; similarly if $\delta = 2$ then $l = 2$, and exactly one of the two cycle lengths is divisible by 3, contradicting $\delta = 2$. Thus all the $a_i = 0$ and so $l = 2 + \delta$. Suppose that the element passes the tests at Step 4. Then $c(t)$ is not divisible by $t^2 + b^2 = t^2 + c^2$, or by $t^r + b^r = t^r - c^r$ for any odd prime r satisfying $5 \leqslant r < n$. Thus the only possible cycle lengths are 1 and 2, and hence $n \leqslant 2l = 4 + 2\delta$. Since $n \geqslant 7$ this is a contradiction. Thus our claim is proved, and the proof is complete that every element returned by the procedure is a double-transposition.

Finally we determine the cost. The cost of processing each of up to $\lceil \log(\varepsilon^{-1}) n^{1/2} / 0.0249 \rceil$ random elements x is as follows. First we compute $c(t)$ at a cost of $O(n^\omega \log n)$ field operations (see Lemma 4.5), and determine the set $\mathcal{C}$ at a cost of $O(\rho_F n \log^2 n \times \log(nq) \log\log n)$ (see Lemma 4.6). Step 3 requires a constant number of field operations to compute a polynomial for comparison with $c(t)$. To perform Step 4, for each of the $O(n/\log n)$ primes r such that $r = 2$ or $5 \leqslant r < n/2$, we require $O(n)$ field operations to check whether $t^r + b^r$ divides $c(t)$, a total of $O(n^2/\log n)$ field operations. Determining g costs $O(n^2)$ field operations, and deciding that g has order less than $n^{18 \log n}$, and if so finding $|g|$ costs $O(n^\omega \log^2 n)$ field operations by Proposition 5.1. Finally extracting the double transposition costs another $O(n^\omega \log^2 n)$ field operations. □

7. Constructing the first vector of $\mathcal{B}$

In order to identify G as conjugate to a subgroup of $Z_V \times H_0$, we need to find a linked basis $\mathcal{B}$ for V relative to H, as defined in Section 3.2. We construct a vector of this basis using a 3-cycle or double-transposition g constructed in the previous section. We need to find a conjugate g' of g such that $[g, g'] = g^{-1} g'^{-1} g g' \neq 1$. In [6, Lemma 5.4], a Monte Carlo algorithm called DOUBLEANDSHRINK is given for achieving this for a larger class of elements g. In our situation, where g is a 3-cycle or a double transposition, the algorithm can be modified to give a Las Vegas algorithm (by checking that $|gg'| = 5$ if g is a 3-cycle, or $|gg'| = 6$ if g is a double transposition, see Table 1).

Lemma 7.1 (DOUBLEANDSHRINK). *Given an element $g \in \mathrm{GL}(V)$ that is conjugate to a 3-cycle or a double-transposition in H_0, there is a Las Vegas algorithm that, with probability greater than $1/10$, constructs a conjugate g' of g such that there is exactly one point moved by both g and g' (and in particular $[g, g'] \neq 1$); the cost is $O(\log n(\xi + \rho_F n^\omega))$.*

Table 1
Pairs of double-transpositions

z	g'	gg'	$\lvert gg'\rvert$	$[g,g']$	$\lvert[g,g']\rvert$
1	(15)(67)	(125)(34)(67)	6	(152)	3
2	(15)(26)	(1625)(34)	4	(12)(56)	2
2	(13)(56)	(1234)(56)	4	(13)(24)	2
2	(15)(36)	(125)(346)	3	(152)(364)	3
3	(12)(35)	(345)	3	(354)	3
3	(13)(25)	(15234)	5	(12453)	5

The next two lemmas deal separately with the cases $p \neq 3$ and $p = 3$, and show how such elements g, g' can be used to identify a vector of the form $b(e_i - e_j) + (W \cap E)$ for some $i \neq j$; by relabeling the e_i and re-scaling if necessary, we may assume that such a vector lies in $\mathcal{B}$. Without loss of generality we may assume that $g, g' \in H_0$ and we identify these elements with their corresponding permutations.

Lemma 7.2. *Suppose that $p \neq 3$ and that $g, g' \in H_0$ correspond to $g = (123)$ and $g' = (145)$.*

(a) *The fixed point subspace $F_V(g)$ of g in V is*

$$F_V(g) = \langle v_1 + 2v_2 + 3v_3, v_4, \ldots, v_{n-\delta}\rangle,$$

$V = F_V(g) \oplus V(g)$ where $V(g) = \langle v_1, v_2\rangle$, and both $F_V(g)$ and $V(g)$ are $\langle g\rangle$-invariant.

(b) *For any $v \in V$, $v + v^g + v^{g^2} \in F_V(g)$ and $2v - v^g - v^{g^2} \in V(g)$, and in particular $v \in V(g)$ if and only if $v + v^g + v^{g^2} = 0$.*

(c) *Similarly $V = F_V([g,g']) \oplus V([g,g'])$, where $F_V([g,g'])$ is the fixed point subspace of $[g,g'] = (142)$ and $V([g,g']) = \langle v_1, v_2 + v_3\rangle$. Moreover $V(g) \cap V([g,g']) = \langle v_1\rangle$ is the span of an element of $\mathcal{B}$.*

Proof. The fixed point space of g in U is $F_U(g) = \langle e_1 + e_2 + e_3, e_4, \ldots, e_n\rangle$. Since $E \subseteq F_U(g) \not\subseteq W$, it follows that $\dim F_V(g) = \dim F_U(g) - \delta = n - 2 - \delta$. Now $F_U(g) \cap W = \langle (e_1 - e_2) + 2(e_2 - e_3) + 3(e_3 - e_4), e_4 - e_5, \ldots, e_{n-1} - e_n\rangle$, and $F_V(g)$ is the image of this subspace under the quotient map $W \to W/(W \cap E)$. Thus $F_V(g)$ is as claimed. Clearly $V = F_V(g) \oplus V(g)$ and both $F_V(g)$ and $V(g)$ are $\langle g\rangle$-invariant.

Let $v \in V$. Then $v = x + av_1 + bv_2$ for some $a, b \in \mathbb{F}$ and $x \in F_V(g)$. We compute $v + v^g + v^{g^2}$ as

$$(x + av_1 + bv_2) + \big(x + av_2 - b(v_1 + v_2)\big) + \big(x - a(v_1 + v_2) + bv_1\big),$$

which equals $3x \in F_V(g)$. Thus, $v + v^g + v^{g^2} \in F_V(g)$, and $2v - v^g - v^{g^2} = 3(v - x) = 3av_1 + 3bv_2 \in V(g)$. In particular $v \in V(g)$ if and only if $v + v^g + v^{g^2} = 0$.

Now $gg' = (12345)$ and $[g,g'] = (142)$. Thus $g^{(324)} = [g,g']$ and therefore $V([g,g'])$ is the subspace spanned by $v_1^{(324)}$ and $v_2^{(324)}$. Now $v_1 = e_1 - e_2 + (W \cap E)$ is mapped by

(324) to $e_1 - e_4 + (W \cap E) = v_1 + v_2 + v_3$, and $v_2 = e_2 - e_3 + (W \cap E)$ is mapped by (324) to $e_4 - e_2 + (W \cap E) = -(v_2 + v_3)$. Hence $V([g, g']) = \langle v_1 + v_2 + v_3, -v_2 - v_3 \rangle = \langle v_1, v_2 + v_3 \rangle$ as claimed. Finally $V(g) \cap V([g, g']) = \langle v_1 \rangle$. □

Lemma 7.3. *Suppose that* $p = 3$ *and that* $g, g' \in H_0$ *correspond to* $g = (12)(34)$ *and* $g' = (15)(67)$.

(a) *The fixed point subspace* $F_V(g)$ *of* g *in* V *is*

$$F_V(g) = \langle v_1 + 2v_2 + v_3, v_3 + 2v_4, v_5, \ldots, v_{n-\delta} \rangle,$$

$V = F_V(g) \oplus V(g)$ *where* $V(g) = \langle v_1, v_3 \rangle$, *and both* $F_V(g)$ *and* $V(g)$ *are* $\langle g \rangle$-*invariant.*

(b) *For any* $v \in V$, $v + v^g \in F_V(g)$ *and* $2v + v^g \in V(g)$. *In particular* $v \in V(g)$ *if and only if* $v + v^g = 0$.

(c) *Similarly* $V = F_V(g^{g'}) \oplus V(g^{g'})$, *where* $F_V(g^{g'})$ *is the fixed point subspace of* $g^{g'} = (25)(34)$ *and* $V(g^{g'}) = \langle v_2 + v_3 + v_4, v_3 \rangle$. *Moreover,* $V(g) \cap V(g^{g'}) = \langle v_3 \rangle$ *is the span of an element of* $\mathcal{B}$.

Proof. The fixed point space of g in U is $F_U(g) = \langle e_1 + e_2, e_3 + e_4, e_5, \ldots, e_n \rangle$. By a similar argument to the proof of Lemma 7.2, we have $F_U(g) \cap W = \langle (e_1 - e_2) + 2(e_2 - e_3) + (e_3 - e_4), (e_3 - e_4) + 2(e_4 - e_5), e_5 - e_6, \ldots, e_{n-1} - e_n \rangle$, and $F_V(g)$ is the image of this subspace under the quotient map $W \to W/(W \cap E)$. Thus $F_V(g)$ is as claimed. Clearly $V = F_V(g) \oplus V(g)$ and both $F_V(g)$ and $V(g)$ are $\langle g \rangle$-invariant.

Let $v \in V$. Then $v = x + av_1 + bv_3$ for some $a, b \in \mathbb{F}$ and $x \in F_V(g)$. We compute $v + v^g$ as $(x + av_1 + bv_3) + (x - av_1 - bv_3) = 2x \in F_V(g)$, and so $2v + v^g = v + 2x = av_1 + bv_3 \in V(g)$. In particular $v \in V(g)$ if and only if $v + v^g = 0$. Clearly $V(g) \cap V(g^{g'}) = \langle v_3 \rangle$.

For part (c), set $h = g^{g'}$. Then $V(h) = V(g)^{g'}$ is the subspace spanned by $v_1^{g'}$ and $v_3^{g'}$. Now $v_1^{g'} = e_5 - e_2 + (W \cap E) = -v_2 - v_3 - v_4$, and $v_3^{g'} = v_3$. Thus $V(h)$ is as claimed, and $V(g) \cap V(h) = \langle v_3 \rangle$. □

We formalise our procedure to construct the first basis element of $\mathcal{B}$ in the following procedure. When $p \neq 3$ we construct a conjugate g' of a 3-cycle g such that g, g' move exactly one common point. We can recognise that a conjugate g' has this property by checking that $|gg'| = 5$. For the case $p = 3$ we work with a double transposition g. There are six A_n-conjugacy classes of pairs (g, g') such that g' is a conjugate of g and $[g, g'] \neq 1$. Taking $g = (12)(34)$, we list in Table 1 a representative for g' from each of these conjugacy classes, and record the number z of points moved by both g and g', the elements gg', $[g, g']$, and their orders. We may recognise a conjugate g' of g that moves exactly one of the points moved by g by checking that $|gg'| = 6$.

Procedure 7.4 (FINDBASISELEMENT). *We are given a positive constant* ε, *a subgroup* $G \leqslant \mathrm{GL}(V) = \mathrm{GL}(d, q)$ *such that* $H' \leqslant G \leqslant Z_V \times H$ *where* H *is conjugate to* H_0, $n \geqslant 7$,

and an element $g \in H'$ that is conjugate to a 3*-cycle if $p \neq 3$ or a double transposition if $p = 3$.*

1. *Run the procedure* DOUBLEANDSHRINK *on g up to $\log(\varepsilon^{-1})/\log(10/9)$ times until a conjugate g' of g is found such that $|gg'| = 5$ (if $p \neq 3$) or $|gg'| = 6$ (if $p = 3$); if no such element is obtained then return* FAILURE. *Otherwise set $h := [g, g']$ (if $p \neq 3$) or $g'gg'$ (if $p = 3$) and perform the following steps.*
2. *Compute $F_V(g)$ and $F_V(h)$.*
3. *Compute $V(g)$ as follows: choose $y \in V \setminus F_V(g)$ and set $u := 2y - y^g - y^{g^2}$ (if $p \neq 3$) or $2y + y^g$ (if $p = 3$); choose $y' \in V \setminus \langle F_V(g), u\rangle$ and set $u' := 2y' - y'^g - y'^{g^2}$ (if $p \neq 3$) or $2y' + y'^g$ (if $p = 3$); set $V(g) = \langle u, u'\rangle$.*
4. *Similarly compute $V(h)$ and return a non-zero vector $v \in V(g) \cap V(h)$.*

We prove that this procedure is valid and estimate its complexity. The complexity involves the quantities ρ_F, ω defined before Theorem 1.1.

Lemma 7.5. *With probability at least $1 - \varepsilon$, Procedure* 7.4 (FINDBASISELEMENT) *returns a vector bv with $v \in \mathcal{B}$ involving two of the points moved by g and in the same g-cycle, for some $b \in \mathbb{F}^{\#}$. It is a Las Vegas algorithm and runs at a cost of at most $O((\log \varepsilon^{-1})(\xi \log n + \rho_F n^{\omega} \log n))$.*

Proof. Since the procedure DOUBLEANDSHRINK returns a suitable g' with probability greater than $1/10$ on a single run, the probability that it fails to find such a g' after $N = \lceil \log(\varepsilon^{-1})/\log(10/9) \rceil$ runs is less than $t = (9/10)^N$ and since $\log(t^{-1}) = N \log(10/9) \geqslant \log(\varepsilon^{-1})$ we have $t \leqslant \varepsilon$. Thus the procedure reports FAILURE with probability less than ε. By our comments in the paragraph preceding Procedure 7.4, the order tests in Step 1 correctly recognise that g' and g move exactly one common point. Also by Lemmas 7.2 and 7.3, the element h is such that $V(g) \cap V(h)$ is a 1-dimensional subspace generated by an element of the required form. Also from these lemmas it follows that Step 3 correctly computes the subspaces $V(g)$ and $V(h)$, and hence the returned vector is of the form claimed, that is to say, the returned vector is a scalar multiple of $e_i - e_j + (W \cap E)$, for some distinct i, j lying in the same g-cycle.

It remains to determine the cost. As remarked in the first paragraph of this section, the cost of finding g' is $O(\log(\varepsilon^{-1})(\log n(\xi + \rho_F n^{\omega})))$, and h is computed with a further cost of $O(\rho_F n^{\omega})$.

Computing a basis for $F_V(g)$ and $F_V(h)$ can be done at a cost of $O(\rho_F n^{\omega})$ as follows: for $X = g - I$ or $h - I$, by [13, Theorem 2.2] there is a deterministic algorithm that computes $(n - \delta) \times (n - \delta)$ matrices L, Q, U, P at a cost of $O(\rho_F n^{\omega})$ such that $X = LQUP$, where L is a lower triangular matrix with 1's on the diagonal, Q, P are permutation matrices, and

$$U = \begin{bmatrix} U_1 \\ 0 \end{bmatrix},$$

where U_1 is an $s \times (n-\delta)$ upper triangular matrix with non-zero diagonal entries, where $s \leqslant n-\delta$ and $s = \mathrm{rank}(X)$. Consider the case where $X = g - I$. (The case $X = h - I$ is similar.) A vector $v \in F_V(g)$ if and only if $vX = 0$, and this holds if and only if $vLQU = 0$, since P is non-singular. Now $vLQU = 0$ holds if and only if the first s entries of vLQ are zero. Let I' denote the $(n-\delta-s) \times (n-\delta)$ matrix formed by removing the first s rows of the identity matrix I. Then $vLQU = 0$ holds if and only if vLQ lies in the row space of I', or equivalently (since Q, L are non-singular), if and only if v lies in the row space of $I'Q^{-1}L^{-1}$. Thus $F_V(g)$ is the row space of $I'Q^{-1}L^{-1}$ and the matrix $I'Q^{-1}L^{-1}$ can be computed as the last $n-\delta-s$ rows of $Q^{-1}L^{-1}$ at a further cost of $O(\rho_F n^\omega)$. Note that $n-\delta-s = \mathrm{rank}(I'Q^{-1}L^{-1}) = \dim(F_V(g)) = n-\delta-2$ so $s = 2$.

For the vector $y \in V \setminus F_V(g)$, we can choose the first row of $Q^{-1}L^{-1}$. Thus u can be found at a cost of $O(\rho_F n^2)$, given $Q^{-1}L^{-1}$. Next we echelonise u against $I'Q^{-1}L^{-1}$, at a cost of $O(\rho_F n^\omega)$, to find a basis for $\langle F_V(g), u\rangle$, and then choose a vector $y' \in V \setminus \langle F_V(g), u\rangle$. Thus the cost of computing u', given u, is $O(\rho_F n^\omega)$, and the basis u, u' for $V(g)$ has been found at a cost of $O(\rho_F n^\omega)$. The cost of finding a basis w, w' for $V(h)$ is the same. Finally to compute v we find a non-trivial solution for $au + a'u' = bw + b'w'$ for $a, a', b, b' \in \mathbb{F}$ at a cost of $O(\rho_F n)$. □

8. More vectors of $\mathcal{B}$: avoiding n-cycles

The algorithm presented in [5] to recognise A_n and S_n requires an n-cycle, which is found by random search, and requires the examination of $O(n)$ group elements. Finding an n-cycle by random search is more expensive than the algorithms presented above for finding the 3-cycle or double-transposition g. In this section we discuss a method that avoids n-cycles for constructing a linked basis $\mathcal{B}$. It uses an element $bh \in G$ such that h is conjugate to an element of H_0 involving a cycle C of prime length r greater than $3n/5$ with an additional property. If $p \neq 3$, then we require that the cycle C contain exactly two of the points moved by the 3-cycle g, while if $p = 3$, then we require that C contain both points from a specified transposition involved in the double-transposition g, and neither of the points from the other transposition. To ensure that we have sufficiently many such elements, we must be able to utilise elements of this type where the prime r is significantly smaller than n.

Lemma 8.1. *Let $n \geqslant 13$, let $g \in A_n$ be a 3-cycle or a double transposition, and in the latter case let (i, j) be one of the g-cycles. Then the proportion of elements h of either S_n or A_n such that*

(a) *h has a cycle C of length r for some prime r satisfying $0.6n + 0.4 < r < 0.95n - 0.85$, and*
(b) *C contains exactly two of the points moved by g, and if g is a double transposition then these points are i and j,*

is greater than $0.03/\log n$ if g is a 3-cycle, or $0.5 \times 10^{-3}/\log n$ if g is a double transposition.

Proof. Let $c(n)$ denote the number of primes r satisfying $0.6n + 0.4 < r < 0.95n - 0.85$. We claim that $c(n) > 0.23n/\log n$ for all $n \geqslant 13$ except $n = 31$, and that if $n = 31$ then $c(n) > 0.15n/\log n$. This claim can be checked by direct computation for $n < 1,000$. Suppose then that $n \geqslant 1,000$. By [24, Theorem 1], for $x \geqslant 1,000$, the number of primes $\pi(x)$ less than x satisfies

$$\frac{x}{\ln x} < \pi(x) < \frac{x}{\ln x}\left(1 + \frac{3}{2\ln x}\right)$$

and hence, since $(1 + 3/(2\ln(1,000))) < 1.218$, we have

$$c(n) > \frac{0.95n - 0.85}{\ln(0.95n - 0.85)} - 1.218\frac{0.6n + 0.4}{\ln(0.6n + 0.4)}.$$

Then, using the facts that cx (for $c > 0$) and $(\ln(0.6x))/\ln x$ are increasing functions for $x \geqslant 1,000$, we obtain the following:

$$0.95n - 0.85 \geqslant 0.949915n, \qquad 0.6n + 0.4 \leqslant 0.6004n,$$
$$\ln(0.95n - 0.85) < \ln n, \qquad \ln(0.6n + 0.4) > \frac{\ln(600)}{\ln(1,000)}\ln n,$$

and hence

$$\begin{aligned} c(n) &> \frac{0.949915 \times n}{\ln n} - \frac{1.218 \times 0.6004 \times n \times \ln(1,000)}{\ln(600) \times \ln n} \\ &> 0.16\frac{n}{\ln n} > 0.23\frac{n}{\log n}. \end{aligned}$$

Thus the claim is proved.

The proportion of elements of S_n having a cycle C of length r, for a given prime r as above, such that C contains exactly two points of a given 3-cycle is

$$\frac{3\binom{n-3}{r-2}(r-1)!(n-r)!}{n!} = \frac{3}{n}\frac{r-1}{n-1}\frac{n-r}{n-2}. \tag{3}$$

For fixed n, the right-hand side of (3) is a monotone decreasing function of r, for r satisfying $0.6n + 0.4 < r < 0.95n - 0.85$. Thus the proportion is greater than the value of the right-hand side of (3) at $r_0 := 0.95n - 0.85$. Now $(r_0 - 1)/(n - 1) \geqslant 0.875$ and $(n - r_0)/(n - 2) > 0.05$ for all $n \geqslant 13$, and so this proportion is greater than $0.13125/n$. For $n \neq 31$, we showed above that there are more than $0.23n/\log n$ such primes r, and hence the proportion of elements in S_n of the required type is greater than $0.03/\log n$. If $n = 31$, then the only prime in the interval is $r = 23$, so the right-hand side of (3) is $\frac{3\times 22\times 8}{n\times 30\times 29} > 0.6/n$. Since $c(31) > 0.15n/\log n$, the proportion of elements in this case is at least $0.09/\log n$. For all n, since exactly half of such elements are even permutations, the proportion in A_n is the same as the proportion in S_n.

Similarly the proportion of elements of S_n, or of A_n, having a cycle C of length r, for a given prime r as above, such that C contains the two points i, j and neither of the other two points moved by the double transposition g, is

$$\frac{\binom{n-4}{r-2}(r-1)!(n-r)!}{n!} = \frac{1}{n}\frac{r-1}{n-1}\frac{n-r}{n-2}\frac{n-r-1}{n-3}.$$

The right-hand side of this equality is $(n-r-1)/3(n-3) > 0.05/3$ times the right-hand side of (3). It follows that the proportion of elements of S_n of the required type is at least $0.03 \times (0.05/3)/\log n = 0.5 \times 10^{-3}/\log n$. □

In the previous section we showed how to construct a vector v of a linked basis $\mathcal{B}$ using either a (conjugate of a) 3-cycle or double-transposition g. Moreover, in the former case, v involved two of the points moved by g, while in the latter case, v involved two points forming one of the transpositions of g (see Lemma 7.5). Now we show how to use v to construct an element $bh \in G$, where $b \in \mathbb{F}^{\#}$ and $h \in H$, such that h satisfies Lemma 8.1(a) and (b). We simultaneously construct a linked sequence of vectors of length greater than $0.6n$. First we handle the case where $p \neq 3$.

Lemma 8.2. *Let $v = e_i - e_j + (W \cap E) \in \mathcal{B}_0$ for some distinct i, j, let $c \in \mathbb{F}^{\#}$, $h \in H_0$, and let r be an odd prime such that $n/2 < r \leqslant n$. Let $r(v, ch)$ denote the least positive integer k such that $v(ch)^k \in \langle v \rangle$. Then $r(v, ch) = r$ if and only if h has a (unique) cycle C of length r and either $\{i, j\} \subseteq C$, or $|\{i, j\} \cap C| = 1$ and h fixes $\{i, j\} \setminus C$.*

Proof. Set $x = ch$. Now $vh^k = e_s - e_t + (W \cap E)$ where $s = i^{h^k}$ and $t = j^{h^k}$. Thus $vx^k \in \langle v \rangle$ if and only if h^k fixes $\{i, j\}$ setwise. Suppose that $r(v, x) = r$. Then h^{2r} fixes i and j so the h-cycles containing i and j have lengths dividing $2r$. By the minimality of $r(v, x)$, at least one of these cycles has length a multiple of r, and since $r > n/2$ there must be a (unique) h-cycle C of length r and C must contain at least one of i and j, say $i \in C$. If $j \notin C$ then h^r fixes $\{i, j\}$ and i, and hence h^r also fixes j. Thus the h-cycle C' containing j has length dividing r, and since $C' \neq C$ and $r > n/2$, we have $|C'| = 1$. Conversely if h has a cycle C of length r and either $\{i, j\} \subseteq C$, or $|\{i, j\} \cap C| = 1$ and h fixes $\{i, j\} \setminus C$, then clearly $r(v, x) = r$. Thus the lemma is proved. □

Procedure 8.3 (MOREBASISVECTORS). *We are given a positive constant ε, a subgroup $G \leqslant \mathrm{GL}(V) = \mathrm{GL}(d, q)$ such that $H' \leqslant G \leqslant Z_V \times H$ where H is conjugate to H_0 and $n \geqslant 13$, an element $g \in H'$ conjugate to a 3-cycle if $p \neq 3$ or a double-transposition if $p = 3$, and a vector $v \in \mathcal{B}$ involving two of the points moved by g and in the same g-cycle.*

1. *Select up to $\frac{1}{c}\log(\varepsilon^{-1})\log n$ random elements $x \in G$, and perform the following steps for each, where $c = 0.03$ if $p \neq 3$ and $c = 0.5 \times 10^{-3}$ if $p = 3$.*
2. *For $i = 0$ if $p = 3$, or $i \in \{0, 1, 2\}$ if $p \neq 3$, compute the vectors $vg^i, vg^i x, \ldots, vg^i x^{n-1}$ and check whether $r_i := r(vg^i, x)$ is a prime r satisfying $0.6n + 0.4 < r < 0.95n - 0.85$. If this is not the case for any i, then return to Step 1.*

If $p \neq 3$, then r_i might be such a prime for more than one value of i. However in this case the value of the prime is the same for all such i.

Else if r is the unique such prime occurring then go to Step 3.

3. *For each i such that $r_i = r$, compute $vg^i x^\ell g$ for $0 \leqslant \ell \leqslant n-1$, and find $n_i := |\{\ell \mid 1 \leqslant \ell < r,\ vg^i x^\ell \neq vg^i x^\ell g\}|$.*

   # For any $x \in G$, the inequality $n_i \leqslant 2$ holds for at most i.

   If we find no i such that $n_i \leqslant 2$, then return to Step 1; *else find ℓ such that $1 \leqslant \ell < r/2$ and $u := vg^i x^\ell - vg^i x^\ell g \neq 0$.*
   Case $p \neq 3$: if $u = dvg^i$, then compute $h_0 := -d^{-1}x^\ell$; or if $u = dvg^{i+1}$, then compute $h_0 := d^{-1}x^\ell$; else return to Step 1.
   Case $p = 3$: if $u = dv$, then compute $h_0 := -d^{-1}x^\ell$; else return to Step 1.
   Compute and return h_0, and the vectors $(vg^i, vg^i h_0, \ldots, vg^i h_0^{(r-2)})$.
4. *If no elements and vectors are returned at Step* 3 *for any of the random elements then report* FAILURE.

Lemma 8.4. *Suppose that $n \geqslant 13$. Then Procedure* 8.3 (FINDMOREVECTORS) *is a Las Vegas algorithm that, with probability at least $1-\varepsilon$, and at a cost of $O((\log \varepsilon^{-1})(\xi \log n + \rho_F n^\omega \log^2 n))$, returns*

(a) *an element in H conjugate to an element of H_0 involving an r-cycle C, for some prime r such that $0.6n + 0.4 < r < 0.95n - 0.85$, and*
(b) *a linked sequence of $r-1$ vectors $(v_1, \ldots, v_{r-1})$ relative to H.*

Proof. Let q be the proportion of elements $h \in H$ such that the permutation corresponding to h has a cycle C of prime length r (where $0.6n + 0.4 < r < 0.95n - 0.85$), C contains exactly two of the points moved by g, and if $p = 3$ then these points are interchanged by g and are the two points involved in the vector v. It follows from Lemma 8.1 that $q > c/\log n$. Arguing as in the first paragraph of the proof of Lemma 6.8, we see that, with probability greater than $1-\varepsilon$, we select at Step 1 at least one element ch with $c \in \mathbb{F}^\#$ and h such an element.

Suppose that $x = ch$ is such an element. Suppose first that $p = 3$. Then $v = e_a - e_b + (W \cap E)$ where a and b lie in C and (a, b) is a transposition of g. In this case Step 2 will succeed by Lemma 8.2, and find $r_0 = r$. Then Step 3 will find $n_0 = 2$ since there are exactly two distinct values of ℓ such that $1 \leqslant \ell < r$ and h^ℓ maps either a to b or b to a; and hence exactly two ℓ such that $vx^\ell \neq vx^\ell g$. Exactly one of these two values of ℓ is less than $r/2$, and for this ℓ we have $u := vx^\ell - vgx^\ell g \neq 0$. Thus in Step 4 we have that $a^{h^\ell} = b$ or $b^{h^\ell} = a$, and in either case we have that $u = -c^\ell v$ and the procedure defines h_0 as $-(-c^\ell)^{-1}x^\ell = h^\ell$. We cannot tell whether h^ℓ maps a to b, or b to a, and so, relabelling the standard basis vectors so that $\{a, b\} = \{1, 2\}$ and $C = (1, 2, 3, \ldots, r)$, the procedure returns either $e_1 - e_2 + (W \cap E), e_2 - e_3 + (W \cap E), \ldots, e_{r-1} - e_r + (W \cap E)$, or $e_2 - e_1 + (W \cap E), e_3 - e_2 + (W \cap E), \ldots, e_r - e_{r-1} + (W \cap E)$, in either case a linked sequence as required.

Next suppose that $x = ch$, that C and r are as in the first paragraph, and that $p \neq 3$, so g is a 3-cycle. For exactly one value of $i \in \{0, 1, 2\}$, the vector vg^i will involve the two points of C moved by g, so we will have $vg^i = e_a - e_b + (W \cap E)$ where a and b lie in C and are moved by g. Thus, for this value of i, Step 2 will find $r_i = r$, by Lemma 8.2. If there is a second value of i, say i', for which $r_{i'}$ is a prime in the correct range then by Lemma 8.2, since a and b are the only points of C moved by g, $r_{i'} = r$ (note, the same value of r) and the third point j moved by g is fixed by h. In this case $r_{i'} = r$ for all three values of i'. So whether $r_{i'} = r$ for a unique $i' = i$, or for all three values of i', we will proceed to Step 3. In Step 3, we obtain $n_i = 2$ (where i is as above); and if h fixes the third point moved by g then $n_{i'} \geqslant r - 2 > 2$ for each $i' \neq i$. Thus, as in the previous paragraph, the procedure will seek (and find) a unique ℓ such that $1 \leqslant \ell < r/2$ and $vg^i x^\ell \neq vg^i x^\ell g$, and will define $u := vg^i x^\ell - vg^i x^\ell g \neq 0$. If $a^{h^\ell} = b$ then b^{h^ℓ} is fixed by g and so $u = c^\ell(vg^i h^\ell - vg^i h^\ell g) = c^\ell(e_b - e_j) + (W \cap E) = c^\ell vg^{i+1}$ and in this case the procedure defines h_0 as $(c^\ell)^{-1}x^\ell = h^\ell$. Alternatively if $b^{h^\ell} = a$ then $u = c^\ell(e_b - e_a) + (W \cap E) = -c^\ell v$, and the procedure defines h_0 as $-(-c^\ell)^{-1}x^\ell = h^\ell$. Thus relabelling the standard basis vectors if necessary, we may assume that $a = 1$, $b = 2$, $1^{h_0} = 2$, and $h_0 = h^\ell$ involves the r-cycle $(1, 2, \ldots, r)$, so that the returned vectors are $e_1 - e_2 + (W \cap E), e_2 - e_3 + (W \cap E), \ldots, e_{r-1} - e_r + (W \cap E)$.

Next, for any prime p, we prove that any vectors returned, for a random $x = ch$, form a linked sequence relative to H. Suppose that Step 2 succeeds for $x = ch$ (where $c \in \mathbb{F}^\#$, $h \in H$) with $r_i = r(vg^i, x) = r$, and suppose that $vg^i = e_a - e_b + (W \cap E)$. By Lemma 8.2, h has a cycle C of length r, C contains at least one of the points a, b, say $a \in C$, and either $b \in C$ or $b^h = b$. Note that the condition $r > n/2$ implies that r is the only prime in the correct interval that we can find for any $r_{i'}$. Suppose also that in Step 3 we find that $n_i \leqslant 2$. Now $vg^i x^\ell \neq vg^i x^\ell g$ if and only if g moves at least one of a^{h^ℓ}, b^{h^ℓ} (note here that, if $p = 2$, then g does not interchange a^{h^ℓ} and b^{h^ℓ} since in that case $|g| = 3$). Thus the fact that $n_i \leqslant 2$ implies that $b^h \neq b$ and hence that $\{a, b\} \subset C$.

We claim that the only points of C moved by g are the points a and b. Suppose to the contrary that g moves a point m of C where $m \notin \{a, b\}$. Then there are positive integers ℓ, ℓ', ℓ'' less than r such that $a^{h^\ell} = b$, $b^{h^{\ell'}} = m$, and $m^{h^{\ell''}} = a$ (so $\ell + \ell' + \ell'' = r$), and it follows that $\{\ell \mid 1 \leqslant \ell < r, vg^i x^\ell \neq vg^i x^\ell g\}$ contains these three integers, and also $r - \ell$, $r - \ell'$, and $r - \ell''$. However this set must have size at most two. Since r is prime, $\ell \neq r - \ell$, so each of these six integers is equal to ℓ or $r - \ell$. Since $\ell' + \ell'' = r - \ell$ it follows therefore that $\ell' = \ell'' = \ell$. However this implies that $r = 3\ell$ contradicting the fact that r is prime. Thus the claim is proved. Therefore the element $x = ch$ is of the type already considered, and for such elements we have proved that the procedure returns a linked sequence of vectors. Thus with probability greater than $1 - \epsilon$ the procedure will succeed and return an element and vectors as claimed.

This analysis has proved the two claims made as comments in Procedure 8.3, namely, in the case $p \neq 3$, at Step 2 there is at most one prime r found for the r_i, and at Step 3 there is at most one value of i such that $n_i \leqslant 2$.

For each of the (up to) $\frac{1}{c}\log(\varepsilon^{-1})\log n$ random elements the cost is as follows. First consider Step 2 where we compute the sequence $vg^i, vg^i x, \ldots, vg^i x^{n-1}$, for $i = 0$ or, if $p = 3$, for $i = 0, 1, 2$. To do this we compute vg^i (at a cost of $O(\rho_F n)$) and $x^2, x^{2^2}, \ldots, x^{2^k}$

where $2^k \leqslant n < 2^{k+1}$ (at a cost of $O(\rho_F n^\omega \log n)$); then, for each $j = 0, \ldots, k$, we multiply the $2^j \times (n-\delta)$ matrix with rows $vg^i, vg^i x, \ldots, vg^i x^{2^j-1}$ by the matrix x^{2^j} to determine $vg^i x^{2^j}, vg^i x^{2^j+1}, \ldots, vg^i x^{2^{j+1}-1}$ (at a cost of $O(\rho_F n^\omega \log n)$). Thus the cost of Step 2 is $O(\rho_F n^\omega \log n)$. For Step 3, computing $vg^i x^\ell g$ for $0 \leqslant \ell < n$ is done with a single matrix multiplication (since we have already computed $vg^i, vg^i x, \ldots, vg^i x^{n-1}$), and finding $\ell < r/2$ such that $u = vg^i x^\ell - vg^i x^\ell g \neq 0$ costs $O(\rho_F n^2)$. Computing x^ℓ requires $O(\log n)$ matrix multiplications (using the x^{2^j} already computed); and finally computing the vectors that are returned costs, as before, $O(\rho_F n^\omega \log n)$. □

9. Completing $\mathcal{B}$

The next step of our procedure is the most delicate. We have, from Procedure 8.3, a linked sequence of $r-1$ vectors that we take to be

$$v_1 = (e_1 - e_2) + (W \cap E), \quad \ldots, \quad v_{r-1} = (e_{r-1} - e_r) + (W \cap E).$$

We need to extend this sequence to a linked basis $\mathcal{B}$. We do this by studying the images of the v_i under random elements from G. The following result, which is a modification of the main result of [10] (see [25, Theorem 4.4.6]), tells us how many random elements will be needed.

For a sequence $H_k = (h_1, \ldots, h_k)$ of elements from a group G, the *cube* of H_k is defined recursively as the subset $C(H_k) = C(H_{k-1}) \cup C(H_{k-1})h_k$, where $H_{k-1} = (h_1, h_2, \ldots, h_{k-1})$, and $C(H_1) = \{1, h_1\}$. Also, for subsets $H \subseteq S_n$ and $R \subseteq \{1, 2, \ldots, n\}$, we denote the set of all points i^h, for $i \in R$ and $h \in H$, by R^H.

Lemma 9.1. *Let $n \geqslant 13$, let ε be a positive constant, and let R be a subset of $\{1, 2, \ldots, n\}$ such that $|R|$ is at least the smallest prime number r satisfying $r > 0.6n + 0.4$. If $H = (g_1, \ldots, g_m)$ is a sequence of uniformly distributed random elements of A_n or S_n with $m \geqslant \log n(\log(\varepsilon^{-1}) + \log n)$, then $R^{C(H)} = \{1, 2, \ldots, n\}$ with probability greater than $1 - \varepsilon$.*

Proof. For $0.6n + 0.4 \leqslant x \leqslant n$, define

$$f(x) = n - \left\lfloor \frac{(n-x)^2}{0.54n} \right\rfloor$$

and note that $f(x) > 0.7n$. By [25, Lemma 4.4.5], for a uniformly distributed random element $g \in A_n$ or S_n, if $S \subseteq \{1, 2, \ldots, n\}$ with $|S| > 3n/5$, then $|S \cup S^g| \geqslant f(|S|)$ with probability at least 0.46. Thus if, say, t uniformly distributed random elements g from A_n or S_n are selected, then the probability that $|S \cup S^g| \geqslant f(|S|)$ for at least one of these elements g is at least $1 - (0.54)^t$.

For a positive integer k, let $f^{\{k\}}(x)$ denote the kth iterated function $f^{\{k\}}(x) = f(f(f\ldots(f(x))\ldots))$. Define ℓ_0 to be the least positive integer k such that $f^{\{k\}}(x) = n$ for all x in the interval $r \leqslant x \leqslant n$. Also define t by

$$t := \left\lceil \frac{\log(\varepsilon^{-1})}{\log(0.54^{-1})} + \frac{\log(\ell_0)}{\log(0.54^{-1})} \right\rceil, \tag{4}$$

and set $m_0 := t\ell_0$. We claim that $m_0 < (\log n)(\log(\varepsilon^{-1}) + \log n)$.

Suppose first that $n > 2^8/0.54$. Then it was shown in the last paragraph of the proof of [25, Theorem 4.4.6] that, if $0.625n < y \leqslant n$ and $k > 0.93 + \log\log(0.54n)$, then $f^{\{k\}}(y) = n$. Now, as observed at the beginning of the proof, whenever $r \leqslant x \leqslant n$ we have $y = f(x) > 0.7n$, and hence $f^{\{k\}}(x) = n$ whenever $k > 1.93 + \log\log(0.54n)$. Thus $\ell_0 \leqslant \ell := \lceil 1.93 + \log\log(0.54n)\rceil$. Now (4) implies that $t < 1.13(\log(\varepsilon^{-1}) + \log\ell_0) + 1$, and, for $n > 2^8/0.54$, we have $\ell < (\log n)/1.13$, and hence $m_0 = t\ell_0 \leqslant t\ell < (\log n)(\log(\varepsilon^{-1}) + \log n)$, as claimed. For $13 \leqslant n \leqslant 2^8/0.54$, we compute the exact value of ℓ_0 and check, for each of these values of n, that $m_0 < (\log n)(\log(\varepsilon^{-1}) + \log n)$. Thus the claim is true for all $n \geqslant 13$.

Let $x_0 = |R|$ so $x_0 > 0.6n + 0.4$. Suppose that $H = (g_1, \ldots, g_m)$ are uniformly distributed random elements from A_n or S_n, where $m \geqslant \lceil(\log n)(\log(\varepsilon^{-1}) + \log n)\rceil$. Let k_1 be the least i such that $|R \cup R^{g_i}| \geqslant f(x_0)$, if such an integer exists, and set $H_1 = (g_{k_1})$, $R_1 = R \cup R^{g_{k_1}} = R^{C(H_1)}$, and $x_1 = |R_1|$ so that $x_1 \geqslant f(x_0)$. Suppose that $H_i = (g_{k_1}, \ldots, g_{k_i})$, $R_i = R^{C(H_i)}$, and $x_i = |R_i|$ have been defined with $k_1 < \cdots < k_i$ and $x_i \geqslant f^{\{i\}}(x_0)$. Let k_{i+1} be the least integer $j > k_i$ such that $|R_i \cup R_i^{g_j}| \geqslant f(x_i)$, if such an integer exists, and set $H_{i+1} = (g_{k_1}, \ldots, g_{k_i}, g_{k_{i+1}})$, $R_{i+1} = R^{C(H_{i+1})}$, and $x_{i+1} = |R_{i+1}|$. Note that $x_{i+1} \geqslant f(x_i) \geqslant f^{\{i+1\}}(x_0)$ since f is a monotonically increasing function. Continuing in this way, let ℓ' be the number of k_i that we obtain. If $\ell' \geqslant \ell_0$, then it follows from the definition of ℓ_0 that $R^{C(H_{\ell'})} = \{1, 2, \ldots, n\}$, and hence that $R^{C(H)} = \{1, 2, \ldots, n\}$.

Thus it is sufficient to prove that $\ell' \geqslant \ell_0$ with probability greater than $1 - \varepsilon$. By our claim proved above, $m \geqslant m_0 = t\ell_0$, and hence the probability that $\ell' \geqslant \ell_0$ is at least the probability that we obtain ℓ_0 integers k_i satisfying $k_{i+1} - k_i \leqslant t$, and by the first paragraph of the proof, this probability is at least $(1 - (0.54)^t)^{\ell_0} > 1 - \ell_0(0.54)^t$, which is at least $1 - \varepsilon$. The last inequality holds if and only if $\varepsilon \geqslant \ell_0(0.54)^t$, which is equivalent to $t \geqslant \frac{\log(\varepsilon^{-1})}{\log(0.54^{-1})} + \frac{\log(\ell_0)}{\log(0.54^{-1})}$, and this is true by the definition of t in (4). □

The following observation will be helpful for understanding the procedure for finding a linked basis for V. Suppose that $(v_1, \ldots, v_s)$ is a linked sequence of vectors with $v_i = (e_i - e_{i+1}) + (W \cap E)$ $(1 \leqslant i \leqslant s)$. Then for $h \in H$, $b \in \mathbb{F}^{\#}$, and $i \leqslant s$, $v_i(bh) \in \langle v_1, \ldots, v_s\rangle$ if and only if $\{i^h, (i+1)^h\} \subseteq \{1, 2, \ldots, s+1\}$, and in this case if $\{i^h, (i+1)^h\} = \{j, k\}$ with $j < k$, then $v_i(bh) = b' \sum_{\ell=j}^{k-1} v_\ell$, where $b' = b$ if $i^h = j$ and $b' = -b$ if $i^h = k$. Since we can identify that $v_i(bh)$ has this form as a linear combination of $v_1, \ldots, v_s$, we can therefore identify the unordered pair $\{i^h, (i+1)^h\}$ and the scalar $\pm b$ up to a sign. Further, if we can also identify i^h then we can determine the scalar b. Thus we shall look for vectors of the form $b' \sum_{\ell=j}^{k-1} v_\ell = b'(e_j - e_k + (W \cap E))$ for some $b' \in \mathbb{F}^{\#}$ and j, k such that $1 \leqslant j < k \leqslant s$; we say that such a vector is an *interval vector in* $\langle v_1, \ldots v_s\rangle$ *with*

support $\{j, k\}$. Interval vectors can be recognised as scalar multiples of sums of consecutive basis vectors from $v_1, \ldots, v_s$.

Now we give an algorithm to construct a linked basis for V that can be used as the standard basis $\mathcal{B}$.

Procedure 9.2 (FINDBASIS). *We are given a positive constant* ε; *a subgroup* G *of* $\mathrm{GL}(V)$ *such that* $H' \leqslant G \leqslant Z_V \times H$ *where* H *is conjugate to* H_0, *and* $n \geqslant 13$; *and a linked sequence of* $r-1$ *vectors* $v_1, \ldots, v_{r-1}$, *where* r *is a prime satisfying* $0.6n + 0.4 < r < 0.95n - 0.85$ *and* $r - 1 < n - \delta$.

Initially set $s := r - 1$, $u_1 := 0$, *and compute* $u_i = \sum_{j=1}^{i-1} v_j$ *for* $i = 2, \ldots, s+1$;

if each $v_i = (e_i - e_{i+1}) + (W \cap E)$, then $u_i = (e_1 - e_i) + (W \cap E)$;

For $m = 1, \ldots, \lceil \log n(\log(\varepsilon^{-1}) + \log n)\rceil$, *do the following*:

1. *Set* $V_0 := \langle v_1, \ldots, v_s\rangle$ *and extend* $v_1, \ldots, v_s$ *to an ordered basis* $\mathcal{B} = (v_1, \ldots, v_s, w_{s+1}, \ldots, w_{n-\delta})$ *for* V;
2. *select a random element* $x = bh \in G$, *where* $h \in H$ *and* $b \in \mathbb{F}^{\#}$; *re-write* x *with respect to the basis* $\mathcal{B}$;

   # the vectors $v_i x$, for $i = 1, \ldots, s$, written with respect to the basis $\mathcal{B}$, will then be the first s rows of the re-written matrix x.

3. *Find the scalar* b *and integers* i, j *such that* $i \leqslant s$ *and* $j = i^h$ *as follows*:

3.1. *if there exists an* $i \leqslant s$ *such that* $v_i x \in V_0$, *then choose such an* i; *note that* $v_i x$ *is then an interval vector in* V_0 *with* (*known*) *support* $\{i^h, (i+1)^h\}$;

   (i) *if* $i < s$ *then, for* $\ell = i+1, \ldots, s$ *or until* i^h *is found, if* $v_i x + v_{i+1}x + \cdots + v_\ell x \in V_0$, *then it is an interval vector in* V_0 *with support* $\{i^h, (\ell+1)^h\}$; *hence find* i^h *and* b, *and set* $j = i^h$;

   (ii) *if* i^h *is not determined in* (i), *then in particular* $i \geqslant 2$; *for* $\ell = i-1, \ldots, 1$ *or until* i^h *is found, if* $v_\ell x + \cdots + v_{i-1}x + v_i x \in V_0$, *then it is an interval vector in* V_0 *with support* $\{\ell^h, (i+1)^h\}$;
   hence find $(i+1)^h$ *and therefore also* i^h *and* b, *and set* $j = i^h$;

3.2. *if there is no* $i \leqslant s$ *such that* $v_i x \in V_0$, *then find* $i < s$ *such that* $v_i x + v_{i+1}x \in V_0$; *it will be an interval vector in* V_0 *with support* $\{i^h, (i+2)^h\}$;

   (i) *if* $i \leqslant s-2$, *then, for* $\ell = i+2, \ldots, s$ *or until* i^h *is found, if* $v_i x + v_{i+1}x + \cdots + v_\ell x \in V_0$, *then it is an interval vector in* V_0 *with support* $\{i^h, (\ell+1)^h\}$; *hence find* i^h *and* b, *and set* $j = i^h$;

   (ii) *if* i^h *is not determined in* (i), *then in particular* $i \geqslant 2$; *for* $\ell = i-1, \ldots, 1$ *or until* i^h *is found, if* $v_\ell x + \cdots + v_i x + v_{i+1}x \in V_0$, *then it is an interval vector in* V_0 *with support* $\{\ell^h, (i+2)^h\}$; *hence find* $(i+2)^h$ *and therefore also* i^h *and* b, *and set* $j = i^h$.

   # Next we find all points of $\{1, \ldots, s\}^h \setminus \{1, \ldots, s\}$, and construct the corresponding new basis vectors. We now know $h = b^{-1}x$.

4.1. *For each k such that $i < k \leqslant s$,*
   (i) *compute $v_i h + v_{i+1}h + \cdots + v_k h$;*
   (ii) *if $v_i h + v_{i+1}h + \cdots + v_k h$ is not an interval vector in V_0, then set $u_{s+2} := v_i h + v_{i+1}h + \cdots + v_k h + u_j$, $v_{s+1} := u_{s+2} - u_{s+1}$, and $s := s+1$; if $s = n - \delta$, then return $v_1, \ldots, v_s$ and $u_2, \ldots, u_{s+1}$;*

4.2. *and also for each k such that $1 \leqslant k < i$,*
   (i) *compute $v_k h + \cdots + v_{i-1}h$;*
   (ii) *if $v_k h + \cdots + v_{i-1}h$ is not an interval vector in V_0, then set $u_{s+2} := -(v_k h + \cdots + v_{i-1}h) + u_j$, $v_{s+1} := u_{s+2} - u_{s+1}$, and $s := s+1$; if $s = n - \delta$, then return $v_1, \ldots, v_s$ and $u_2, \ldots, u_{s+1}$.*

5. *Set $m := m + 1$.*

If no vectors are returned then report FAILURE.

Lemma 9.3. *Let $n \geqslant 13$. Procedure* 9.2 (FINDBASIS) *is a Las Vegas algorithm that, with probability at least $1 - \varepsilon$, when given a linked sequence relative to H of length greater than $0.6n - 0.6$ returns a linked basis for V, at a cost of $O((\log n \log(\varepsilon^{-1}) + \log^2 n)(\xi + \rho_F n^\omega))$.*

Proof. First we show that any sequence of vectors $v_1, \ldots, v_{n-\delta}$ returned by Procedure 9.2 is a linked basis. Suppose that at the start of some run of the 'for loop' the vectors v_i form a linked sequence and hence, without loss of generality, can be taken as

$$v_1 = (e_1 - e_2) + (W \cap E), \quad \ldots, \quad v_s = (e_s - e_{s+1}) + (W \cap E).$$

This is certainly true at the beginning of the first run of the 'for loop.' We will show that under this assumption we also have a linked sequence at the end of this run of the 'for loop.'

For $i = 1, \ldots, s+1$, we have $u_i = (e_1 - e_i) + (W \cap E)$. Let $S = \{1, 2, \ldots, s+1\}$. During this run of the 'for loop,' suppose that the matrices and vectors are re-written in terms of the ordered basis $\mathcal{B} = (v_1, \ldots, v_s, w_{s+1}, \ldots, w_{n-\delta})$. This means, in particular, that the interval vector $b(e_j - e_k) + (W \cap E)$, where $b \in \mathbb{F}^\#$ and $1 \leqslant j < k \leqslant s+1$, which is equal to $b \sum_{\ell=j}^{k-1} v_\ell$, is represented as the $(n-\delta)$-tuple with ith-entry equal to b if $j \leqslant i < k$, and 0 otherwise.

Now $|S| = s + 1 \geqslant r > 0.6 + 0.4$. Let $T = S \cap S^{h^{-1}} = \{i \in S \mid i^h \in S\}$, and let $t = |T \cap \{s, s+1\}|$. Then $|T| \geqslant 2|S| - n > 0.2n + 0.8$, which is at least 3 for $n \geqslant 13$. Assume that, for each $i \in T$, we have $i+1 \notin T$ and $i+2 \notin T$. This means in particular that, if $s \in T$, then $s+1 \notin T$, and so $t \leqslant 1$. Also this assumption implies that $|S| \geqslant 3(|T| - t) + t = 3|T| - 2t \geqslant 3(2|S| - n) - 2$, and hence $|S| \leqslant (3n+2)/5$, which is a contradiction. Thus there exists $i \in T$ such that either $(i+1)^h \in S$ or $(i+2)^h \in S$, or equivalently, such that $v_i h \in V_0$ or $v_i h + v_{i+1}h \in V_0$, respectively. Thus the steps in either 3.1 or 3.2 will be attempted so i will be defined. Note that, whenever $1 \leqslant c < d \leqslant n - \delta$, $\sum_{\ell=c}^{d-1} v_\ell(bh) = b(e_{c^h} - e_{d^h}) + (W \cap E)$. Since $|S \cap S^{h^{-1}}| \geqslant 3$, it follows that the steps in 3.1 or 3.2 will succeed in correctly determining the value $j = i^h$.

Next suppose that in Step 4.1(ii) we find that $\sum_{\ell=i}^{k} v_\ell h = (e_{i^h} - e_{(k+1)^h}) + (W \cap E) \notin V_0$. Since $i^h \in S$, this means that $(k+1)^h \notin S$. Relabelling $(k+1)^h$ as $s+2$,

we see that the new vectors defined in this step are $u_{s+2} = (e_1 - e_{s+2}) + (W \cap E)$ and $v_{s+1} = (e_{s+1} - e_{s+2}) + (W \cap E)$. Thus the extended sequence $v_1, \ldots, v_{s+1}$ is also linked. Similarly if, in Step 4.2(ii) we find that $\sum_{\ell=k}^{i-1} v_\ell h = (e_{k^h} - e_{i^h}) + (W \cap E) \notin V_0$, then $k^h \notin S$, and relabelling k^h as $s+2$, we again find that $u_{s+2} = (e_1 - e_{s+2}) + (W \cap E)$, $v_{s+1} = (e_{s+1} - e_{s+2}) + (W \cap E)$, and so $v_1, \ldots, v_{s+1}$ is again linked. It follows that any sequence of vectors returned by the procedure will be a linked basis for V. Observe that, when processing a fixed random h, if $k, k' \leqslant s$ then $k^h \neq (k')^h$, and hence we do not need to recompute V_0 when s increases during a single run of the 'for loop.'

To see that the procedure succeeds with probability at least $1 - \varepsilon$, observe that, at the end of Step 4 we have extended the linked sequence of vectors to $v_1, \ldots, v_{s'}$ where, in the notation used in the previous paragraph, $\{1, \ldots, s'\} = S \cup S^h$. If $R = \{1, \ldots, r\}$, the initial value for the set S, and if $H = (h_1, \ldots, h_m)$ is the sequence of random elements selected, then the value of S after m runs of the for-loop will be the image $R^{C(H)}$ of R under the cube $C(H)$ of H, where $C(H)$ is as defined just before Lemma 9.1. If $m = 1, \ldots, \lceil \log n(\log(\varepsilon^{-1}) + \log n) \rceil$, then by Lemma 9.1, $R^{C(H)} = \{1, \ldots, n\}$ with probability at least $1 - \varepsilon$. Thus with probability at least $1 - \varepsilon$ the procedure returns a linked basis.

Finally consider the cost of the procedure. At the beginning of each run of the for-loop we form a $(n - \delta + s) \times (n - \delta)$ matrix with rows $v_1, \ldots, v_s$ followed by the identity matrix of order $n - \delta$, and find the lexicographically least maximal linearly independent set of rows at a cost of $O(\rho_F n^\omega)$ (see [13]); this sequence of rows will be our basis $\mathcal{B}$. Let P be the matrix with the vectors of $\mathcal{B}$ as rows. Then the matrix representing the random element $x = bh$ with respect to $\mathcal{B}$ is PxP^{-1}, and this can be found at a cost of $O(\rho_F n^\omega)$. All the vectors $v_i x$, for $1 \leqslant i \leqslant s$ can be computed at a cost of one matrix multiplication, that is $O(\rho_F n^\omega)$. An interval vector in V_0 can be found among the $v_i x$, if one exists, by inspecting these vectors to determine if there is one for which the non-zero entries are all equal and occur as a consecutive sequence in the first s positions; similarly in Step 3.2, the $v_i x + v_{i+1} x$ can be found at a cost of $O(\rho_F n^2)$, and then an interval vector in V_0 can be identified by inspecting the entries. Completing parts (i) and (ii) of Step 3.1 or 3.2 requires up to n vector additions and inspections, and this can be done at a cost of $O(\rho_F n^2)$. Similarly Step 4 requires $O(n)$ vector additions and inspections, at a cost of $O(\rho_F n^2)$. Thus the procedure costs $O((\log n \log(\varepsilon^{-1}) + \log^2 n)(\xi + \rho_F n^\omega))$. □

10. Identifying scalars and permutations

Now that we have constructed a linked basis $\mathcal{B}$ relative to H, we complete our procedure by showing, for a given $bx \in G$ with $b \in \mathbb{F}^\#$ and $x \in H$, how to identify the scalar b and find the permutation in S_n corresponding to x. We may assume without loss of generality that $\mathcal{B}$ is the standard basis $\mathcal{B} = (v_1, \ldots, v_{n-\delta})$, where $v_i = e_i - e_{i+1} + (W \cap E)$ for $1 \leqslant i \leqslant n - \delta$. If $\delta = 2$ we also set

$$\begin{aligned} v_{n-1} &:= \sum_{i=1}^{n-2} i\, v_i = e_1 + \cdots + e_{n-2} - (n-2)e_{n-1} + (W \cap E) \\ &= e_{n-1} - e_n + (W \cap E) \end{aligned}$$

noting that $\sum_{i=1}^{n} e_i \in W \cap E$ and $ne_{n-1} = 0$.

First we give a brief informal discussion of the case where the characteristic of $\mathbb{F}$ does not divide n, in order to give an understanding of how we may identify the permutation corresponding to x. In this case V has dimension $n-1$. Let g be the matrix representing bx with respect to the standard basis $\mathcal{B}$. It is easy to see that the non-zero entries of any row of g now consist of a consecutive sequence of equal values. We first determine the permutation x. If the above sequence for the ith row of g (corresponding to the basis vector $e_i - e_{i+1}$) starts in the jth column and ends in the kth column, then either $i^x = j$ and $(i+1)^x = k$ or $i^x = k$ and $(i+1)^x = j$. Thus looking at the first two rows gives the pairs $\{1^x, 2^x\}$ and $\{2^x, 3^x\}$. The common value in these two pairs must be the value of 2^x, thus determining also 1^x and 3^x as well. So now when analysing the ith row of the matrix we may assume that i^x is already known, so no ambiguity occurs in computing $(i+1)^x$. Thus x can be computed. Now b can be computed by looking at any row. If $i^x < (i+1)^x$ then b is the constant value of the non-zero elements of the row; else b is minus this value.

This is essentially our approach in determining x and b. However, because of the increased complexities of the case where p divides n, and because it is a little simpler to analyse a sparser matrix, for the purposes of this section we will work with the alternative basis $\mathcal{B}' = (u_2, \ldots, u_{n-\delta+1})$ where

$$u_i = \sum_{1 \leqslant j < i} v_j = e_1 - e_i + (W \cap E)$$

for $2 \leqslant i \leqslant n$ and we set $u_1 = 0$. Thus $v_i = u_{i+1} - u_i$ for $1 \leqslant i \leqslant n-1$. Note that the u_i are constructed in Procedure 9.2. If $\delta = 2$, we will need an expression for u_n as a linear combination of the vectors in $\mathcal{B}'$. We find this from the definition of u_n as follows. Note that in this case p divides n and so $ne_1 = 0$ and $e = \sum_{i=1}^{n} e_i \in W \cap E$, and hence $\sum_{i=1}^{n} u_i = ne_1 - \sum_{i=1}^{n} e_i + (W \cap E) = 0$. Therefore, since $u_1 = 0$,

$$u_n = -\sum_{i=2}^{n-1} u_i.$$

Lemma 10.1. *Let $bx \in G$ with $b \in \mathbb{F}^{\#}$ and $x \in H$, and let $i \in \{2, \ldots, n\}$. Also let $w'(i)$ denote the number of non-zero coefficients in the expression for $(u_i)(bx)$ as a linear combination of the vectors in $\mathcal{B}'$.*

(a) *If $\delta = 1$, or if $\delta = 2$ and $n \notin \{1^x, i^x\}$, then $(u_i)(bx) = b(u_{i^x} - u_{1^x})$ and $w'(i) = 2$ if $1 \notin \{1^x, i^x\}$, and is 1 otherwise.*
(b) *If $\delta = 2$ and $\{1^x, i^x\} = \{j, n\}$, where $j < n$, then*

$$(u_i)(bx) = \pm b\left(u_j + \sum_{\ell=2}^{n-1} u_\ell\right),$$

and so $w'(i) = n-3$ if $j > 1$ and $p = 2$, and is $n-2$ otherwise.

(c) *Thus if $\delta = 1$, then either $1^x = 1$ and $w'(i) = 1$ for all i, or $1^x > 1$ and $w'(i) = 1$ if $i^x = 1$ and is 2 otherwise. If $\delta = 2$, then one of the following holds*:
(i) $1^x = 1$ *and* $w'(i) = n-2$ *if* $i^x = n$ *and is* 1 *otherwise*;
(ii) $1 < 1^x < n$, *and*

$$w'(i) = \begin{cases} n-3 & \text{if } i^x = n \text{ and } p = 2, \\ n-2 & \text{if } i^x = n \text{ and } p > 2, \\ 2 & \text{if } i^x \neq n, 1, \\ 1 & \text{if } i^x = 1; \end{cases}$$

(iii) $1^x = n$, *and* $w'(i) = n-3$ *if* $i^x > 1$ *and* $p = 2$, *and is* $n-2$ *otherwise.*

Proof. Part (a) is easily checked and (c) follows from (a) and (b). To prove (b), assume that $\delta = 2$ and $\{1^x, i^x\} = \{j, n\}$. Then

$$\begin{aligned} \pm b^{-1}(u_i)(bx) &= \pm(e_{1^x} - e_{i^x}) + (W \cap E) = \pm(e_j - e_n) + (W \cap E) \\ &= \pm\big(e_j + (e_1 + \cdots + e_{n-1})\big) + (W \cap E) = \pm\left(-u_j - \sum_{\ell=2}^{n-1} u_\ell\right), \end{aligned}$$

since $ne_1 = 0$. □

Let bx be as in Lemma 10.1. If $w'(i) = 2$ then from $(u_i)(bx) = b(u_{i^x} - u_{1^x})$ we can determine $\pm b$ and $\{i^x, 1^x\}$. Similarly if $w'(i) = 1$ then $(u_i)(bx) = \pm b u_j$ for some j with $1 < j \leqslant n - \delta + 1$, and we can find $\pm b$ and $\{i^x, 1^x\} = \{1, j\}$. Thus if we find distinct i_1, i_2 greater than 1 such that $\{w'(i_1), w'(i_2)\} \subseteq \{1, 2\}$, then we can find $\pm b$, $\{i_1^x, 1^x\}$, $\{i_2^x, 1^x\}$, and hence also $1^x, i_1^x, i_2^x$ and b. After this, for any ℓ such that $w'(\ell) = 1$ or 2 we can determine ℓ^x since we already know 1^x. These observations form the basis of our procedure below.

Procedure 10.2 (FINDPERMUTATION). *We are given $bx \in Z_V \times H$ where $b \in \mathbb{F}^\#$, $x \in H$, and H is conjugate to $H_0 \cong S_n$ with $n \geqslant 6$; and the basis $\mathcal{B}'$ defined above.*

1. *Compute $(u_i)(bx)$ as a linear combination from $\mathcal{B}'$, and determine $w'(i)$, for $2 \leqslant i \leqslant n$.*
2. *Case $\delta = 1$.*
   Here $\{w'(2), w'(3)\} \subseteq \{1, 2\}$, so we can determine $b, 1^x, 2^x, 3^x$ as above, and then determine i^x from $(u_i)(bx)$ for $4 \leqslant i \leqslant n$.
3. *Case $\delta = 2$.*
   If $\{w'(2), w'(3)\} \subseteq \{1, 2\}$, then we determine $b, 1^x, 2^x, 3^x$ as above; next we determine i^x from $(u_i)(bx)$ for each $i > 3$ such that $w'(i) \leqslant 2$; and finally, for the unique i such that $w'(i) \geqslant n-3$, we have $i^x = n$.
   If $\{w'(2), w'(3)\} = \{1, n-2\}$ or $\{2, n-2\}$ in the case $p > 2$, or if $\{w'(2), w'(3)\} = \{1, n-3\}$ or $\{2, n-3\}$ in the case $p = 2$, then we must have $\{w'(4), w'(5)\} \subseteq \{1, 2\}$, and we can determine $b, 1^x, 4^x, 5^x$ as above; next we determine i^x from $(u_i)(bx)$ for

each $i \neq 1, 4, 5$ such that $w'(i) \leqslant 2$; and finally, for the unique i such that $w'(i) \geqslant n-3$, we have $i^x = n$.
Else $\{w'(2), w'(3)\} \subseteq \{n-3, n-2\}$. In this case $1^x = n$, and for each $i \geqslant 3$, $u'_i := (u_i)(bx) - (u_2)(bx) = b(u_{i^x} - u_{2^x})$; from u'_3 and u'_4 therefore we can determine $b, 2^x, 3^x, 4^x$; finally we determine i^x from u'_i for each $i > 4$.

Lemma 10.3. *Procedure* 10.2 (FINDPERM) *is a deterministic algorithm that, given $bx \in Z_V \times H$ where $b \in \mathbb{F}^{\#}$, $x \in H$, and H is conjugate to $H_0 \cong S_n$ with $n \geqslant 6$, and given the basis $\mathcal{B}'$ defined above, determines the scalar b and the permutation corresponding to x at a cost of $O(\rho_F n^{\omega})$.*

Proof. The correctness of Procedure 10.2 follows from Lemma 10.1 and the discussion following it. Now we determine the cost. Let P be the matrix with rows $u_2, \ldots, u_{n-\delta+1}$. Then the matrix for bx with respect to the basis $\mathcal{B}' = (u_2, \ldots, u_{n-\delta+1})$ is $P(bx)P^{-1}$, and can be found at a cost of $O(\rho_F n^{\omega})$; the rows of this matrix are $u_2(bx), \ldots, u_{n-\delta+1}(bx)$, written in terms of $\mathcal{B}'$. If $\delta = 2$ then we note that $u_n = -\sum_{2 \leqslant i \leqslant n-1} u_i$, and hence that $u_n(bx) = -\sum_{2 \leqslant i \leqslant n-1} u_i(bx)$. Thus, if $2 \leqslant i \leqslant n-\delta+1$, then $w'(i)$ is the number of non-zeros of the ith row vector of $P(bx)P^{-1}$, and if $\delta = 2$, then $w'(n)$ is the number of non-zero entries in the row vector obtained by adding together all the rows of $P(bx)P^{-1}$. Thus, given $P(bx)P^{-1}$, determining the $w'(i)$ costs at most $O(\rho_F n^2)$. From now on the computation of b and the permutation corresponding to x, is achieved at a cost of inspecting the entries of $O(n)$ vectors. □

10.1. The proof of Theorem 1.1

Finally we prove Theorem 1.1 by drawing together the procedures we have presented. Suppose, as in Section 3.2, that $H' \leqslant G = \langle X \rangle \leqslant Z_V \times H < \mathrm{GL}(V)$ where H is conjugate to $H_0 \cong S_n$, and that ε is given, with $0 < \varepsilon < 1$. Let $\varepsilon_0 = \varepsilon/4$. Also let $\alpha = 1/3$ if $p \neq 3$ and $\alpha = 1/2$ if $p = 3$.

Using Procedure 6.7 if $p \neq 3$, or Procedure 6.9 if $p = 3$, we construct with probability at least $1 - \varepsilon_0$, an element $g \in H'$ that is conjugate to a 3-cycle or double-transposition in H_0 (that is, g has type $1^{n-3}3^1$ or $1^{n-4}2^2$) respectively, at a cost of

$$O\big(\log(\varepsilon_0^{-1})n^{\alpha}\big(\xi + \rho_F \log^2 n\big(n^{\omega} + n \log nq \log\log n\big)\big)\big).$$

If this is successful, then we use g in Procedure 7.4 to construct, with probability at least $1 - \varepsilon_0$, a vector bv with $v \in \mathcal{B}$ involving two points moved by g and in the same g-cycle, and $b \in \mathbb{F}^{\#}$, at a cost of $O(\log(\varepsilon_0^{-1}) \log n(\xi + \rho_F n^{\omega}))$. If this is successful, then with this vector v we apply Procedures 8.3 and 9.2, each with probability at least $1 - \varepsilon_0$, to construct a linked basis $\mathcal{B}$ at a cost of

$$O\big(\big(\log n \log\big(\varepsilon_0^{-1}\big) + \log^2 n\big)\xi + \rho_F n^{\omega} \log\big(\varepsilon_0^{-1}\big) \log^2 n\big).$$

If $\mathcal{B} = (v'_1, \ldots, v'_{n-\delta})$, then the map $S : v_i \mapsto v'_i$ defines an element of $\mathrm{GL}(V)$ that conjugates H_0 to H. The total cost of these procedures is therefore

$$O\big(\log\big(\varepsilon^{-1}\big)n^{\alpha}\big(\xi+\rho_F\log^2 n\big(n^{\omega}+n\log nq\log\log n\big)\big)\big),$$

as required. Finally Procedure 10.2 evaluates the scalar b and the permutation $\lambda(x)$ corresponding to a given element $bx \in Z_V \times H$ (relative to $\mathcal{B}$) at a cost of $O(\rho_F n^{\omega})$. Thus to evaluate the scalar b, and the permutation corresponding to each of the generators $bx \in X$ (and thereby define the homomorphism $\lambda : G \to Z_{q-1} \times S_n$) costs a further $O(|X|\rho_F n^{\omega})$. Evaluating λ^{-1} on a pair $(b, x) \in Z_{q-1} \times S_n$ can be done by assembling the matrix A representing x in H_0 with respect to the basis $\mathcal{B}' = \{u_2, \ldots, u_{n-\delta+1}\}$, and then conjugating bA by the appropriate change of basis matrix. This costs $O(n^{\omega}\rho_F)$. This completes the proof of Theorem 1.1.

References

[1] M. Aschbacher, On the maximal subgroups of the finite classical groups, Invent. Math. 76 (1984) 469–514.
[2] L. Babai, R.M. Beals, A polynomial-time theory of black-box groups, I, in: C. Campbell, E. Robertson, N. Ruskuc, G. Smith (Eds.), Groups St Andrews 1997 in Bath I, in: London Math. Soc. Lecture Note Ser., vol. 260, Cambridge Univ. Press, Cambridge, 1999, pp. 30–64.
[3] L. Babai, G. Cooperman, L. Finkelstein, E. Luks, Á. Seress, Fast Monte Carlo algorithms for permutation groups, J. Comput. System Sci. 50 (1995) 296–308.
[4] R.M. Beals, C.R. Leedham-Green, A.C. Niemeyer, C.E. Praeger, Á. Seress, Permutations with restricted cycle structure and an algorithmic application, Combin. Probab. Comput. 11 (2002) 447–464.
[5] R.M. Beals, C.R. Leedham-Green, A.C. Niemeyer, C.E. Praeger, Á. Seress, A black-box group algorithm for recognizing finite symmetric and alternating groups, I, Trans. Amer. Math. Soc. 355 (2003) 2097–2113.
[6] R.M. Beals, C.R. Leedham-Green, A.C. Niemeyer, C.E. Praeger, Á. Seress, A black-box group algorithm for recognizing finite symmetric and alternating groups, II, 2004, in preparation.
[7] S. Bratus, I. Pak, Fast constructive recognition of a black box group isomorphic to S_n or A_n using Goldbach's conjecture, J. Symbolic Comput. 29 (1) (2000) 33–57.
[8] P. Bürgisser, M. Clausen, M. Amin Shokrollahi, Algebraic Complexity Theory, Springer-Verlag, Berlin, 1997.
[9] F. Celler, C.R. Leedham-Green, S.H. Murray, A.C. Niemeyer , E.A. O'Brien, Generating random elements of a finite group, Comm. Algebra 23 (1995) 4931–4948.
[10] G. Cooperman, L. Finkelstein, N. Sarawagi, A random base change algorithm for permutation groups, in: Proceedings of International Symposium on Symbolic and Algebraic Computation, ISSAC'90, ACM Press, 1990, pp. 161–168.
[11] D. Coppersmith, S. Winograd, Matrix multiplication via arithmetic progressions, J. Symbolic Comput. 9 (3) (1990) 251–280.
[12] The GAP Group, GAP – Groups, Algorithms, and Programming, Version 4.4, http://www.gap-system.org, 2004.
[13] O.H. Ibarra, S. Moran, R. Hui, A generalisation of the fast LUP matrix decomposition algorithm and applications, J. Algorithms 3 (1982) 45–56.
[14] G.D. James, On the minimal dimensions of irreducible representations of symmetric groups, Math. Proc. Cambridge Philos. Soc. 94 (3) (1983) 417–424.
[15] E. Kaltofen, V. Shoup, Subquadratic-time factoring of polynomials over finite fields, Math. Comp. 67 (223) (1998) 1179–1197.
[16] W. Kantor, Á. Seress, Computing with matrix groups, in: A. Ivanov, M. Liebeck, J. Saxl (Eds.), Groups, Combinatorics, and Geometry, Durham 2001, World Scientific, 2003, pp. 123–137.
[17] P. Kleidman, M. Liebeck, The Subgroup Structure of the Finite Classical Groups, London Math. Soc. Lecture Note Ser., vol. 129, Cambridge Univ. Press, Cambridge, 1990.
[18] E. Landau, Handbuch der Lehre vor der Verteilung der Primzahlen, Bd. I, Teubner, Leipzig, 1909.
[19] C.R. Leedham-Green, The computational matrix group project, in: W.M. Kantor, Á. Seress (Eds.), Groups and Computation III, in: OSU Math. Res. Inst. Publ., vol. 8, de Gruyter, 2001, pp. 85–101.

[20] C.R. Leedham-Green, A.C. Niemeyer, E.A. O'Brien, C.E. Praeger, Recognising matrix groups over finite fields, in: V. Weispfenning, J. Grabmeier, E. Kaltofen (Eds.), Computer Algebra Handbook, Foundations, Applications, Systems, Springer-Verlag, Berlin, 2003, pp. 459–460.
[21] R. Lidl, H. Niederreiter, Finite Fields, second ed., Cambridge Univ. Press, Cambridge, 1997.
[22] M.W. Liebeck, On the orders of maximal subgroups of the finite classical groups, Proc. London Math. Soc. (3) 50 (1985) 426–446.
[23] I. Niven, H.S. Zuckerman, H.L. Montgomery, An Introduction to the Theory of Numbers, fifth ed., Wiley, New York, 1991.
[24] J. Barkley Rosser, L. Schoenfeld, Approximate formulas for some functions of prime numbers, Illinois J. Math. 6 (1962) 64–97.
[25] Á. Seress, Permutation Group Algorithms, Cambridge Univ. Press, Cambridge, 2003.
[26] J. von zur Gathen, J. Gerhard, Modern Computer Algebra, Cambridge Univ. Press, Cambridge, 1999.
[27] A. Wagner, The faithful linear representation of least degree of S_n and A_n over a field of characteristic 2, Math. Z. 151 (2) (1976) 127–137.
[28] A. Wagner, The faithful linear representations of least degree of S_n and A_n over a field of odd characteristic, Math. Z. 154 (2) (1977) 103–114.
[29] A. Wagner, An observation on the degrees of projective representations of the symmetric and alternating group over an arbitrary field, Arch. Math. (Basel) 29 (6) (1977) 583–589.