Nagoya Math. J., **246** (2022), 372–411
DOI 10.1017/nmj.2021.6

ON THE HEIGHT AND RELATIONAL COMPLEXITY OF A FINITE PERMUTATION GROUP

NICK GILL*, BIANCA LODÀ AND PABLO SPIGA*

Abstract. Let G be a permutation group on a set Ω of size t. We say that $\Lambda \subseteq \Omega$ is an *independent set* if its pointwise stabilizer is not equal to the pointwise stabilizer of any proper subset of Λ. We define the *height* of G to be the maximum size of an independent set, and we denote this quantity $\mathrm{H}(G)$. In this paper, we study $\mathrm{H}(G)$ for the case when G is primitive. Our main result asserts that either $\mathrm{H}(G) < 9\log t$ or else G is in a particular well-studied family (the *primitive large–base groups*). An immediate corollary of this result is a characterization of primitive permutation groups with large *relational complexity*, the latter quantity being a statistic introduced by Cherlin in his study of the model theory of permutation groups. We also study $\mathrm{I}(G)$, the maximum length of an irredundant base of G, in which case we prove that if G is primitive, then either $\mathrm{I}(G) < 7\log t$ or else, again, G is in a particular family (which includes the primitive large–base groups as well as some others).

Contents

1 Introduction **373**
1.1 Definition of statistics . . . 373
1.2 Main results . . . 375
1.3 Structure of the paper . . . 376

2 Some background lemmas **376**

3 Groups with a regular normal subgroup **381**

4 Diagonal action **381**

5 Product actions **384**

6 Almost simple groups: Bounds for $\mathrm{I}(G)$ **385**
6.1 Standard actions of A_n and S_n . . . 386
6.2 The symplectic/orthogonal case . . . 386
6.3 Cases involving graph automorphisms . . . 395

7 Almost simple groups: Bounds on $\mathrm{H}(G)$ **396**
7.1 $G_0 = \mathrm{PSL}_n(q)$ and elements of Ω are m-spaces . . . 396
7.2 Some preliminaries . . . 397
7.3 $G \leq \mathrm{PGL}_n(q)$ . . . 398

Received September 10, 2020. Revised March 25, 2021. Accepted April 25, 2021.
2020 Mathematics subject classification: 20B25, 20B15.
Keywords and phrases: permutation group, height of a permutation group, relational complexity, base size.
*Nick Gill and Pablo Spiga would like to acknowledge the support of the Engineering and Physical Sciences Research Council grant EP/R028702/1 over the course of this research.

© *(2021) The Authors. The publishing rights in this article are licenced to Foundation Nagoya Mathematical Journal under an exclusive license.*

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

7.4 $G \not\leq \mathrm{PGL}_n(q)$ 401
7.5 $G_0 = \mathrm{PSL}_n(q)$ and elements of Ω are pairs of subspaces 402
7.6 Suppose that Γ is a complete matching 403
7.7 Suppose that Γ is not a complete matching 404
7.8 The other classical groups 408

§1. Introduction

In this paper, we study a number of different statistics pertaining to primitive permutation groups. We are interested in understanding which families of primitive permutation groups exhibit large values for these various statistics.

From here on, let G be a finite primitive permutation group on a set Ω of size $t < \infty$. The statistic of most interest to us is the *relational complexity* of a permutation group, $\mathrm{RC}(G)$, a statistic that was first introduced in [CMS96]. We were motivated by a remark in the same paper in which the authors suggest that it should be possible to classify those primitive groups G for which $\mathrm{RC}(G) > \sqrt{t}$. Our main result (more or less) yields this classification; indeed, it applies with the $\sqrt{t}$ replaced by the asymptotically weaker $9\log t + 1$ (here, and everywhere, logarithms are base 2).

In the course of our investigations, we were inspired by the following result of Liebeck [Lie84] concerning the minimum base size, $\mathrm{b}(G)$, of the permutation group G. The definition of this quantity is given below.

THEOREM 1.1. *Let G be a finite primitive group of degree t. Then, one of the following holds:*

1. *G is a subgroup of $\mathrm{Sym}(m) \wr \mathrm{Sym}(r)$ containing $(\mathrm{Alt}(m))^r$, where the action of $\mathrm{Sym}(m)$ is on k-subsets of $\{1,\dots,m\}$ and the wreath product has the product action of degree $t = \binom{m}{k}^r$;*
2. *$\mathrm{b}(G) < 9\log t$.*

Theorem 1.1 leads us to make the following definition.

DEFINITION 1.2. The group G is a *primitive large–base group* if G is a subgroup of $\mathrm{Sym}(m) \wr \mathrm{Sym}(r)$ containing $(\mathrm{Alt}(m))^r$, where the action of $\mathrm{Sym}(m)$ is on k-subsets of $\{1,\dots,m\}$ and the wreath product has the product action of degree $t = \binom{m}{k}^r$.

In this paper, we prove two results which are variants of the main result of Theorem 1.1; in both, the family of primitive large–base groups appear in a similar way to Theorem 1.1, in that they exhibit exceptional behavior with respect to certain statistics. In order to state these results, we must first define the statistics of interest.[1]

1.1 Definition of statistics

For $\Lambda = \{\omega_1,\dots,\omega_k\} \subseteq \Omega$, we write $G_{(\Lambda)}$ or $G_{\omega_1,\omega_2,\dots,\omega_k}$ for the pointwise stabilizer. If $G_{(\Lambda)} = \{1\}$, then we say that Λ is a *base*. The size of a smallest possible base is known as the *base size* of G and is denoted $\mathrm{b}(G)$.

We say that a base is a *minimal base* if no proper subset of it is a base. We denote the maximum size of a minimal base by $\mathrm{B}(G)$.

[1] There has been recent improvement on Liebeck's result. We now know that if G is not a primitive large–base group, then $\mathrm{b}(G) \leq \max\{\lceil \log t \rceil + 1, 7\}$ [MRD].

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Given an ordered sequence of elements of Ω, $[\omega_1,\omega_2,\ldots,\omega_k]$, we can study the associated *stabilizer chain*:

$$G \geq G_{\omega_1} \geq G_{\omega_1,\omega_2} \geq G_{\omega_1,\omega_2,\omega_3} \geq \cdots \geq G_{\omega_1,\omega_2,\ldots,\omega_k}.$$

If all the inclusions given above are strict, then the stabilizer chain is called *irredundant*. If, furthermore, the group $G_{\omega_1,\omega_2,\ldots,\omega_k}$ is trivial, then the sequence $[\omega_1,\omega_2,\ldots,\omega_k]$ is called an *irredundant base*. The size of the longest possible irredundant base is denoted $\mathrm{I}(G)$. Note that an irredundant base is not a base (because it is an ordered sequence, not a set).

Finally, let Λ be any subset of Ω. We say that Λ is an *independent set* if its pointwise stabilizer is not equal to the pointwise stabilizer of any proper subset of Λ. We define the *height* of G to be the maximum size of an independent set, and we denote this quantity $\mathrm{H}(G)$.

There is a basic connection between the four statistics we have defined so far:

$$\mathrm{b}(G) \leq \mathrm{B}(G) \leq \mathrm{H}(G) \leq \mathrm{I}(G) \leq \mathrm{b}(G)\log t. \tag{1.1}$$

The proof of (1.1) goes as follows: the first inequality is obvious. For the second, suppose that $\Lambda = \{\omega_1,\ldots,\omega_k\}$ is a minimal base; then, observe that Λ is also an independent set. For the third, suppose that $\Lambda = \{\omega_1,\ldots,\omega_k\}$ is an independent set; then, observe that the ordered list $[\omega_1,\ldots,\omega_k]$ can be extended to form an irredundant base.

The fourth inequality has been attributed to Blaha [Bla92], who, in turn, describes it as an *observation of Babai* [Bab81]. Suppose that G has a base of size $b = \mathrm{b}(G)$. Then, in particular, $|G| \leq t^b$. On the other hand, any irredundant base has size at most $\log|G|$. We conclude that $\mathrm{I}(G) \leq \log(t^b)$, and the result follows.

We have one more statistic to define. To start, suppose that $r,n \in \mathbb{N}$ with $r \leq n$. If $I,J \in \Omega^n$, then we write $I \tilde{r} J$, and say that I is *r-equivalent* to J with respect to the action of G, if G contains elements that map every subtuple of size r in I to the corresponding subtuple in J, that is,

$$\text{for every } k_1,k_2,\ldots,k_r \in \{1,\ldots,n\}, \text{ there exists } h \in G \text{ with } I^h_{k_i} = J_{k_i}, \text{ for every } i \in \{1,\ldots,r\}.$$

Here, I_k denotes the k^{th} element of tuple I and I^g denotes the image of I under the action of g. Note that n-equivalence simply requires the existence of an element of G mapping I to J.

The group G is said to be of *relational complexity* r if r is the smallest integer such that, for all $n \in \mathbb{N}$ with $n \geq r$ and for all n-tuples $I,J \in \Omega^n$,

$$I \tilde{r} J \implies I \tilde{n} J.$$

In this case, we write $\mathrm{RC}(G) = r$. The fact that relational complexity is well defined takes a little proving. There is an equivalent definition of $\mathrm{RC}(G)$ in terms of relational structures: $\mathrm{RC}(G)$ is the least k for which G can be viewed as an automorphism group acting naturally on a homogeneous relational system whose relations are k-ary [Che16].

The relationship between relational complexity and the statistics defined above is given by the following inequality (which is proved in Section 2):

$$\mathrm{RC}(G) \leq \mathrm{H}(G) + 1. \tag{1.2}$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Note that in what follows, we may sometimes include the set on which we are acting in our statistical notation, if this set is in any doubt. So, for instance, $\mathrm{H}(H)$ and $\mathrm{H}(H,\Delta)$ both mean *the height of the permutation group H on a set Δ*.

1.2 Main results

Our first main result is a Liebeck-type result for height.

THEOREM 1.3. *Let G be a finite primitive group of degree t. If G is not a primitive large–base group, then*

$$\mathrm{H}(G) < 9\log t.$$

Equation (1.2) immediately yields the following corollary.

COROLLARY 1.4. *Let G be a finite primitive group of degree t. If G is not a primitive large–base group, then*

$$\mathrm{RC}(G) < 9\log t + 1.$$

Corollary 1.4 yields the classification proposed in [CMS96] in a very strong form (since our bound is logarithmic in t, whereas the original suggestion was for $\sqrt{t}$). Note, however, that we do not assert that all of the primitive large–base groups genuinely violate the bound in Corollary 1.4. A precise analysis of the relational complexity of the primitive large–base groups was started in [CMS96], but, although a great deal of progress was made, much remains to be done. Note, however, that [Che16, Theorem 2] implies that the family of primitive large–base groups contain an infinite number of groups violating the bound in Corollary 1.4 including, in particular, the alternating groups in their natural action. It is clear that the same is true with respect to Theorem 1.3: both $\mathrm{Alt}(t)$ and $\mathrm{Sym}(t)$ in their natural action violate the bound in Theorem 1.3 for t large enough. In future work, we will show that the same is true of $\mathrm{Alt}(t)$ and $\mathrm{Sym}(t)$ in their action on k-sets.

Equation (1.1) yields a second corollary.

COROLLARY 1.5. *Let G be a finite primitive group of degree t. If G is not a primitive large–base group, then*

$$\mathrm{B}(G) < 9\log t.$$

Of course, we could get yet another corollary by replacing $\mathrm{B}(G)$ by $\mathrm{b}(G)$ here, but this would only reprise Theorem 1.1. What, then, of $\mathrm{I}(G)$? Our main result concerning this statistic is the following.

THEOREM 1.6. *Let G be a finite primitive group of degree t. Then, one of the following holds:*

1. *There exists an almost simple group A, with socle S, such that G is a subgroup of $A \wr \mathrm{Sym}(r)$ containing S^r, the action of A is one of the following:*
   (a) *the action of $\mathrm{Sym}(m)$ on k-subsets of $\{1,\dots,m\}$ (so degree $s = \binom{m}{k}$);*
   (b) *the action of a classical group on a set of subspaces of the natural module, or on a set of pairs of subspaces;*

   and the action of the wreath product has the product action of degree $t = s^r$, where s is the degree of the action of A.
2. $\mathrm{I}(G) < 7\log t$.

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

The listed possible actions of A are a subset of the so-called *standard actions* of almost simple groups. In the second case, let V be a finite-dimensional vector space over $\mathbb{F}_q$; if $S = \mathrm{PSL}(V)$, then such an action is on all subspaces of some fixed dimension k in V; if $S = \mathrm{Cl}(V)$, one of the other classical groups defined on V, then such an action is either on all nondegenerate subspaces of some fixed dimension k in V, or on totally isotropic subspaces of some fixed dimension k and of a given type in V.

Theorem 1.6 is not as strong as we would like—we would like to eliminate the groups listed at item (1b), so that we end up with the same exceptional family as in the other main results. Indeed, we conjecture that something like Theorem 1.3 should hold for $\mathrm{I}(G)$ also.

CONJECTURE 1.7. *There exists a constant $C > 0$ such that if G is a finite primitive group of degree t that is not a primitive large–base group, then*

$$\mathrm{I}(G) < C \log t.$$

1.3 Structure of the paper

Our proofs of Theorems 1.3 and 1.6 make use of the O'Nan–Scott theorem, which divides the class of primitive permutation groups into various families. There are various ways of stating the O'Nan–Scott theorem; in this paper, we make use of the division into eight types which is described in [Pra90].

After proving some useful background lemmas in Section 2, the work in Sections 3–5 when combined with the O'Nan–Scott theorem reduces the problem of proving Theorem 1.6 to a question about almost simple groups. This question is addressed in Section 6, where the proof of Theorem 1.6 is completed.

The work in Section 5, when combined with Theorem 1.6, reduces the problem of proving Theorem 1.3 to a question about almost simple groups. This question is addressed in Section 7, where the proof of Theorem 1.3 is completed.

§2. Some background lemmas

In this section, we collect together a number of lemmas, mostly about height, which will be useful later. We also introduce one final statistic: for a group G, we define $\ell(G)$ to be the length of the longest subgroup chain in G. It is clear that $\ell(G)$ is greater than or equal to $\mathrm{I}(G)$ (and, hence, all the other statistics appearing in Equation (1.1)); it is equally clear that $\ell(G) \leq \log |G|$.

We start with a proof of (1.2).

LEMMA 2.1. $\mathrm{RC}(G) \leq \mathrm{H}(G) + 1$.

Proof. Let $h = \mathrm{H}(G)$ and consider a pair $(I, J) \in \Omega^n$ such that $I \nsim_r J$ with $r = h + 1$. We must show that $I \nsim_n J$.

Observe that we can reorder the tuples without affecting their equivalence. Hence, without loss of generality, we can assume that

$$G_{I_1} > G_{I_1, I_2} > \cdots > G_{I_1, I_2, \ldots, I_\ell},$$

for some $\ell \leq h$, and then this chain stabilizes, that is,

$$G_{I_1, \ldots, I_\ell} = G_{I_1, \ldots, I_{\ell+j}},$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

for all $1 \leq j \leq n-\ell$. From the assumption of $(h+1)$-equivalence, it follows that there exists an element $g \in G$ such that $I_i^g = J_i$, for all $1 \leq i \leq \ell$, and observe that the set of all such elements g forms a coset of $G_{I_1,\ldots,I_\ell}$.

The assumption of $(h+1)$-equivalence implies, moreover, that for all $1 \leq j \leq n-\ell$, there exists $g_j \in G$ such that

$$\begin{cases} I_i^{g_j} = J_i, & \text{for} \quad 1 \leq i \leq \ell; \\ I_{\ell+j}^{g_j} = J_{\ell+j}. \end{cases}$$

The set of all such elements g_j forms a coset of $G_{I_1,\ldots,I_\ell,I_{\ell+j}}$, which is, again, a coset of $G_{I_1,\ldots,I_\ell}$. Indeed, since any coset of $G_{I_1,\ldots,I_\ell}$ is defined by the image of the points $I_1,\ldots,I_\ell$ under an element of the coset, we conclude that elements of the same coset of $G_{I_1,\ldots,I_\ell}$ map $I_{\ell+j}$ to $J_{\ell+j}$, for all $1 \leq j \leq n-\ell$. In particular,$I \mathrel{\widetilde{n}} J$, as required. □

The next two lemmas are little more than observations and require no proof.

LEMMA 2.2. *Let $H \leq G$ and let $\Lambda \subseteq \Omega$ be an independent set with respect to H. Then, Λ is an independent set with respect to G. In particular,* $\mathrm{H}(H) \leq \mathrm{H}(G)$.

LEMMA 2.3. *Let $\Lambda = \{\alpha_1,\ldots,\alpha_n\}$. Then, Λ is independent if and only if*

$$G_{(\Lambda)} \lneqq G_{(\Lambda\setminus\{\alpha_i\})},$$

for all $i = 1,\ldots,n$.

LEMMA 2.4. *Let $\Lambda \subseteq \Omega$. Then, there exists $\Gamma \subseteq \Lambda$ such that Γ is independent and $G_{(\Lambda)} = G_{(\Gamma)}$.*

Proof. If Λ is an independent set, then the claim is trivially true. So, assume it is not. Then, there exists a proper subset $\Lambda_0 \subset \Lambda$ such that $G_{(\Lambda_0)} = G_{(\Lambda)}$. If Λ_0 is independent, then we are done; otherwise, consider Λ_0 and repeat the same argument. □

The next three results give information about the behavior of the statistics $\mathrm{H}(G)$, $\mathrm{I}(G)$, and $\ell(G)$ when the permutation group G is a direct product.

LEMMA 2.5. *Let A and B be permutation groups on Ω_1 and Ω_2, then*

$$\mathrm{H}(A \times B, \Omega_1 \times \Omega_2) \leq \mathrm{H}(A,\Omega_1) + \mathrm{H}(B,\Omega_2).$$

Proof. Let $\pi_A : A \times B \to A$ be the projection onto A and observe that the kernel of π_A is $\{1\} \times B$. Let $\Lambda = \{(\alpha_1,\beta_1),\ldots,(\alpha_k,\beta_k)\}$ be an independent set for $A \times B$. Set $G := A \times B$, and, for every $j \in \{1,\ldots,k\}$, set $G_j := G_{(\alpha_1,\beta_1)} \cap \cdots \cap G_{(\alpha_j,\beta_j)}$.

Now, consider the set $\Lambda_1 = \{\alpha_1,\ldots,\alpha_k\} \subseteq \Omega_1$. Observe that, *a priori*, this set may not be independent. However, by Lemma 2.4, there exists a subset $\{j_1,\ldots,j_s\} \subseteq \{1,\ldots,k\}$ such that the set $\Delta_1 = \{\alpha_{j_1},\ldots,\alpha_{j_s}\}$ is independent and $A_{(\Delta_1)} = A_{(\Lambda_1)}$. After relabeling elements, we can assume $\Delta_1 = \{\alpha_1,\ldots,\alpha_s\}$ and observe that $A_{(\Delta_1)} = \pi_A(G_s)$. In particular, $s \leq \mathrm{H}(A,\Omega_1)$.

Now, let $H := G_s$. It is clear that the set $\{(\alpha_{s+1},\beta_{s+1}),\ldots,(\alpha_k,\beta_k)\}$ is independent with respect to H. Moreover, for all $j \in \{s+1,\ldots,k\}$, we have

$$\pi_A\left(\bigcap_{i=s+1}^{k} H_{(\alpha_i,\beta_i)}\right) = \pi_A\left(\bigcap_{\substack{i=s+1 \\ i\neq j}}^{k} H_{(\alpha_i,\beta_i)}\right),$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

and this implies that

$$\bigcap_{i=s+1}^{k} H_{(\alpha_i,\beta_i)} \cap (\{1\} \times B) \lneqq \left(\bigcap_{\substack{i=s+1 \\ i \neq j}}^{k} H_{(\alpha_i,\beta_i)} \right) \cap (\{1\} \times B).$$

Therefore, by Lemma 2.3, the set $\Delta_2 = \{\beta_{s+1}, \ldots, \beta_k\}$ is an independent set for $B \cap H$, and hence, $k - s \leq \mathrm{H}(B \cap H, \Omega_2)$. As $\mathrm{H}(B \cap H, \Omega_2) \leq \mathrm{H}(B, \Omega_2)$, by Lemma 2.2, we have

$$k \leq \mathrm{H}(A, \Omega_1) + \mathrm{H}(B, \Omega_2),$$

and we conclude that $\mathrm{H}(G, \Omega_1 \times \Omega_2) \leq \mathrm{H}(A, \Omega_1) + \mathrm{H}(B, \Omega_2)$. □

LEMMA 2.6. *Let A and B be nonidentity permutation groups on Ω_1 and Ω_2, then*

$$\mathrm{I}(A \times B, \Omega_1 \times \Omega_2) = \mathrm{I}(A, \Omega_1) + \mathrm{I}(B, \Omega_2) - 1.$$

Proof. Let $[\alpha_1, \ldots, \alpha_k]$ and $[\beta_1, \ldots, \beta_s]$ be irredundant bases for A and B acting on Ω_1 and on Ω_2 and with $k := \mathrm{I}(A, \Omega_1)$ and $s := \mathrm{I}(B, \Omega_2)$. As A and B are not the identity, $s, k \geq 1$. Now, consider the ordered sequence

$$[(\alpha_1, \beta_1), (\alpha_1, \beta_2), \ldots, (\alpha_1, \beta_s), (\alpha_2, \beta_s), (\alpha_3, \beta_s), \ldots, (\alpha_k, \beta_s)].$$

Using the fact that $[\alpha_1, \ldots, \alpha_k]$ and $[\beta_1, \ldots, \beta_s]$ are irredundant bases, it follows that this sequence of elements of $\Omega_1 \times \Omega_2$ is an irredundant base for $A \times B$ acting on $\Omega_1 \times \Omega_2$. Thus,

$$\mathrm{I}(A \times B, \Omega_1 \times \Omega_2) \geq s + k - 1 = \mathrm{I}(A, \Omega_1) + \mathrm{I}(B, \Omega_2) - 1.$$

Now, let $[(\alpha_1, \beta_1), \ldots, (\alpha_t, \beta_t)]$ be an irredundant base for $A \times B$ acting on $\Omega_1 \times \Omega_2$ with $t = \mathrm{I}(A \times B, \Omega_1 \times \Omega_2)$. Let $\mathcal{S}_A := \{i \in \{2, \ldots, t\} \mid A_{\alpha_1,\ldots,\alpha_{i-1}} > A_{\alpha_1,\ldots,\alpha_{i-1},\alpha_i}\}$ and $\mathcal{S}_B := \{i \in \{2, \ldots, t\} \mid B_{\beta_1,\ldots,\beta_{i-1}} > B_{\beta_1,\ldots,\beta_{i-1},\beta_i}\}$. As, for each $i \in \{2, \ldots, t\}$, we have

$$A_{\alpha_1,\ldots,\alpha_{i-1}} \times B_{\beta_1,\ldots,\beta_{i-1}} = \bigcap_{j=1}^{i-1} (A \times B)_{(\alpha_j,\beta_j)} > \bigcap_{j=1}^{i} (A \times B)_{(\alpha_j,\beta_j)} = A_{\alpha_1,\ldots,\alpha_i} \times B_{\beta_1,\ldots,\beta_i},$$

we deduce $\mathcal{S}_A \cup \mathcal{S}_B = \{2, \ldots, t\}$. Moreover, if $\mathcal{S}_A = \{i_1, i_2, \ldots, i_r\}$, then $[\alpha_1, \alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_r}]$ is an irredundant base for A acting on Ω_1, and hence, $\mathrm{I}(A, \Omega_1) \geq |\mathcal{S}_A| + 1$. Similarly, $\mathrm{I}(B, \Omega_2) \geq |\mathcal{S}_B| + 1$. Therefore,

$$\begin{aligned} \mathrm{I}(A, \Omega_1) + \mathrm{I}(B, \Omega_2) &\geq |\mathcal{S}_A| + |\mathcal{S}_B| + 2 \geq |\mathcal{S}_A \cup \mathcal{S}_B| + 2 = |\{2, \ldots, t\}| + 2 = t + 1 \\ &= \mathrm{I}(A \times B, \Omega_1 \times \Omega_2) + 1. \end{aligned}$$

□

LEMMA 2.7. *Let A and B be groups. Then, $\ell(A \times B) \leq \ell(A) + \ell(B)$.*

Proof. Let $G = A \times B$, and let

$$G_\kappa < G_{\kappa-1} < \cdots < G_1 < G_0 = G$$

be a chain of subgroups of length κ. Let $\pi : G \to B$ be the projection onto B. Let

$$\mathcal{I} := \{i \in \{0, \ldots, \kappa - 1\} \mid \pi(G_i) > \pi(G_{i+1})\},$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

and let $s := |\mathcal{I}|$. We may write $\mathcal{I} = \{i_1, \dots, i_s\}$ with $i_1 < i_2 < \cdots < i_s$. Now,

$$\pi(G_{i_s+1}) < \pi(G_{i_s}) < \cdots < \pi(G_{i_2}) < \pi(G_{i_1}) \leq B$$

is a strictly increasing chain of subgroups of B, and hence, $s \leq \ell(B)$.

Let j be an arbitrary element in $\{1, \dots, s-1\}$. Define $i_{s+1} = \kappa$. By definition,

$$\pi(G_{i_{j+1}}) = \cdots = \pi(G_{i_j+3}) = \pi(G_{i_j+2}) = \pi(G_{i_j+1}). \tag{2.1}$$

Since, by assumption, we have the strictly increasing sequence

$$G_{i_{j+1}} < \cdots < G_{i_j+3} < G_{i_j+2} < G_{i_j+1},$$

the first isomorphism theorem and (2.1) give

$$A \cap G_{i_{j+1}} < \cdots < A \cap G_{i_j+3} < A \cap G_{i_j+2} < A \cap G_{i_j+1}.$$

This increasing sequence consists of $i_{j+1} - i_j$ subgroups.

The argument in the previous paragraph can be applied for every $j \in \{1, \dots, s-1\}$, and hence, we obtain s chains of strictly increasing sequences of subgroups of A consisting of $i_1 - i_2, i_2 - i_3, \dots, i_{s-1} - i_s$ terms. By sticking these strictly increasing sequences together, we obtain a longer increasing sequence of subgroups of A of length κ. This longer increasing sequence is not necessarily strictly increasing; however, the only positions where an equality can occur are the positions where we attach two strictly increasing chains, that is, in the positions

$$A \cap G_{i_{j+1}} \leq A \cap G_{i_{j+1}-1}.$$

Since the number of these positions is s and since $s \leq \ell(B)$, we have a strictly increasing chain in A of length $\kappa - \ell(B)$. We conclude that $\kappa - \ell(B) \leq \ell(A)$, and the result follows. □

LEMMA 2.8. *Let $G \leq \mathrm{Sym}(\Omega)$ and let $N \trianglelefteq G$. Then,*

$$\mathrm{H}(G) \leq \mathrm{H}(N) + \ell(G/N)$$

and

$$\mathrm{I}(G) \leq \mathrm{I}(N) + \ell(G/N).$$

Proof. Let $\{\omega_1, \dots, \omega_k\}$ be an independent set for G in its action on Ω.

Let us consider the action of N on Ω. Observe that, *a priori*, the set $\{\omega_1, \dots, \omega_k\}$ is not independent for this action. However, by Lemma 2.4, there exists a subset $J := \{j_1, \dots, j_s\} \subseteq \{1, \dots, k\}$, such that $\{\omega_{j_1}, \dots, \omega_{j_s}\}$ is independent for N on Ω and

$$N \cap \left(\bigcap_{i=1}^{s} G_{\omega_{j_i}} \right) = N \cap \left(\bigcap_{i=1}^{k} G_{\omega_i} \right).$$

After relabeling elements, we can assume $J = \{1, \dots, s\}$. In particular, this yields that

$$s \leq \mathrm{H}(N, \Omega).$$

Observe that, for all $j \in \{s, \dots, k\}$, we have

$$N \cap \left(\bigcap_{i=1}^{s} G_{\omega_i} \right) = N \cap \left(\bigcap_{i=1}^{j} G_{\omega_i} \right). \tag{2.2}$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Let $\pi\colon G \to G/N$ be the natural projection, and let $I = \{s,\dots,k\}$. Suppose that there exist $i_1, i_2 \in I$, with $i_1 < i_2$, such that

$$\pi\left(\bigcap_{j=1}^{i_1} G_{\omega_j}\right) = \pi\left(\bigcap_{j=1}^{i_2} G_{\omega_j}\right).$$

This equality, (2.2), and the first isomorphism theorem imply that

$$\bigcap_{j=1}^{i_1} G_{\omega_j} = \bigcap_{j=1}^{i_2} G_{\omega_j}.$$

Therefore,

$$\bigcap_{j=1}^{k} G_{\omega_j} = \bigcap_{j=1}^{i_2} G_{\omega_j} \cap \bigcap_{j=i_2+1}^{k} G_{\omega_j} = \bigcap_{j=1}^{i_1} G_{\omega_j} \cap \bigcap_{j=i_2+1}^{k} G_{\omega_j}.$$

Since $\{1,\dots,i_1\} \cup \{i_2+1,\dots,k\} \subsetneq \{1,\dots,k\}$, we have a contradiction with the independence of $\{\omega_1,\dots,\omega_k\}$ for the action of G on Ω. We conclude that, for $i_1, i_2 \in I = \{s,\dots,k\}$, with $i_1 < i_2$, we have

$$\pi\left(\bigcap_{j=1}^{i_1} G_{\omega_j}\right) < \pi\left(\bigcap_{j=1}^{i_2} G_{\omega_j}\right).$$

We conclude that $\ell(G/N) \geq k - s = \mathrm{H}(G) - \mathrm{H}(N)$, as required.

We now prove that $\mathrm{I}(G) \leq \mathrm{I}(N) + \ell(G/N)$. Let $[\omega_1,\dots,\omega_s]$ be an irredundant base for G with $s := \mathrm{I}(G)$. Now, set $G_0 := G$ and, for $i \in \{1,\dots,t\}$, $G_i := \cap_{j=1}^{i} G_{\omega_j}$. In particular, we have a strictly decreasing chain

$$G_0 > G_1 > \cdots > G_{s-1} > G_s = 1.$$

Now, let $\mathcal{S}_1 := \{i \in \{0,\dots,s-1\} \mid G_i N > G_{i+1} N\}$ and $\mathcal{S}_2 := \{i \in \{0,\dots,s-1\} \mid G_i \cap N > G_{i+1} \cap N\}$. We claim that $\mathcal{S}_1 \cup \mathcal{S}_2 = \{0,\dots,s-1\}$. We argue by contradiction, and we let $i \in \{0,\dots,s-1\} \setminus (\mathcal{S}_1 \cup \mathcal{S}_2)$. This means that $G_i N = G_{i+1} N$ and $G_i \cap N = G_{i+1} \cap N$. Let $g \in G_i \subseteq G_{i+1} N$. Then, there exist $h \in G_{i+1}$ and $n \in N$ with $g = hn$. Thus, $h^{-1} g = n \in G_i \cap N = G_{i+1} \cap N$. As $h \in G_{i+1}$, this yields $g \in G_{i+1}$. Therefore, $G_i = G_{i+1}$, which is clearly a contradiction.

We have $|\mathcal{S}_1| \leq \ell(G/N)$. Now, let $\mathcal{S}_2 = \{i_1, i_2, \dots, i_r\}$ with $i_1 < i_2 < \cdots < i_r$, and consider the sequence $[\omega_{i_1+1}, \omega_{i_2+1}, \dots, \omega_{i_r+1}]$. We claim that this is an irredundant sequence for N, from which it follows that $\mathrm{I}(N) \geq r = |\mathcal{S}_2|$. Suppose, by contradiction, that

$$N_{\omega_{i_1+1}} \cap \cdots \cap N_{\omega_{i_j+1}} = N_{\omega_{i_1+1}} \cap \cdots \cap N_{\omega_{i_j+1}} \cap N_{\omega_{i_{j+1}+1}}.$$

This gives

$$G_{\omega_{i_1+1}} \cap \cdots \cap G_{\omega_{i_j+1}} \cap N = G_{\omega_{i_1+1}} \cap \cdots \cap G_{\omega_{i_j+1}} \cap G_{\omega_{i_{j+1}+1}} \cap N.$$

We now intersect both sides of this equality with $G_{i_{j+1}} = \bigcap_{u=1}^{i_{j+1}} G_{\omega_u}$, and we obtain

$$G_{i_{j+1}} \cap N = (G_{i_{j+1}} \cap G_{\alpha_{i_{j+1}+1}}) \cap N = G_{i_{j+1}+1} \cap N,$$

contradicting the fact that $i_{j+1} \in \mathcal{S}_2$.

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Summing up,

$$\mathrm{I}(N)+\ell(G/N)\geq|\mathcal{S}_1|+|\mathcal{S}_2|\geq|\mathcal{S}_1\cup\mathcal{S}_2|=|\{0,\dots,s-1\}|=s=\mathrm{I}(G).$$ □

§3. Groups with a regular normal subgroup

In this section, we prove Theorem 1.6 under the supposition that G contains a regular normal subgroup. In fact, here, we use a general argument, which also holds in the case of imprimitive groups containing a regular normal subgroup. The primitive groups containing a regular normal subgroup are the groups of type HA, TW, HS, or HC in the language of [Pra90].

The main result of this section is the following.

PROPOSITION 3.1. *Let G be a permutation group on a finite set Ω of size t. If G contains a regular normal subgroup, then*

$$\mathrm{I}(G)\leq\log t+1.$$

Proof. Suppose that G has a regular normal subgroup N, and fix $\omega_0\in\Omega$. The action of G on Ω is permutation isomorphic to the *affine* action of G on N, where the group N acts on N via translations and the group G_{ω_0} acts by group conjugation. In particular, if $n,v\in N$ and $x\in G_\omega$, then

$$v^{xn}=v^x\cdot n=x^{-1}vx\cdot n.$$

In what follows, we identify Ω with N. We let $[\omega_1,\dots,\omega_k]$ be an irredundant base, and we set $H:=G_{\omega_0}$. We may assume, without loss of generality, that $\omega_1=1\in N$. Now, $H_{\omega_2}=C_H()\omega_2$ fixes ω_2, and hence, it fixes each element of $\langle\omega_2\rangle$. Similarly,

$$H_{\omega_2,\omega_3}=C_H(\omega_2)\cap C_H(\omega_3)=C_H(\langle\omega_2,\omega_3\rangle).$$

Continuing in this way, we obtain a chain of subgroups of H,

$$\begin{aligned}H\geq H_{\omega_2}=C_H(\omega_2)\geq H_{\omega_2,\omega_3}=C_H(\langle\omega_2,\omega_3\rangle)\geq\cdots\\ \cdots\geq H_{\omega_2,\omega_3,\dots,\omega_k}=C_H(\langle\omega_2,\omega_3,\cdots,\omega_k\rangle),\end{aligned}\tag{3.1}$$

and a chain of subgroups of N,

$$\langle1\rangle\leq\langle\omega_2\rangle\leq\cdots\leq\langle\omega_2,\dots,\omega_k\rangle\leq N.\tag{3.2}$$

Since $[\omega_1,\dots,\omega_k]$ is an irredundant base, the inequalities in (3.1) are strict inequalities. This yields that the inequalities in (3.2) must also be strict. In particular,

$$k\leq\log|N|+1.$$ □

§4. Diagonal action

In this section, we prove Theorem 1.6 under the supposition that G is a primitive group of type SD (see [Pra90]). Our main result is the following.

PROPOSITION 4.1. *Let G be a primitive permutation group on a finite set Ω of size t. If G is of type SD, then*

$$\mathrm{I}(G)\leq\log t.$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

We start by reviewing the structure of primitive groups of diagonal type; this will also help us to set some notation.

Let T be a nonabelian simple group, let m be a positive integer with $m \geq 2$, and let $S := T^m$ be the Cartesian product of m copies of T. We denote by $D := \{(t, \dots, t) \mid t \in T\}$ the diagonal subgroup of S, and we set

$$\Omega := [D : S]$$

the set of right cosets of D in S. Each element of Ω has a distinguished coset representative, that is, an element whose first coordinate begins with a 1. In other words,

$$D(t_1, t_2, \dots, t_m) = D(1, t_1^{-1}t_2, \cdots, t_1^{-1}t_m).$$

In particular, Ω is in one-to-one correspondence with the elements of T^{m-1} and hence

$$\log|\Omega| = (m-1)\log|T|. \tag{4.1}$$

Observe, first, that the elements of S act on Ω by coset multiplication, that is, for every $D(t_1, \dots, t_m) \in \Omega$ and $(x_1, \dots, x_m) \in S$, we have

$$D(t_1, \dots, t_m)^{(x_1, \dots, x_m)} = D(t_1x_1, \dots, t_mx_m).$$

Observe, second, that the elements of $\mathrm{Aut}(T)$ act on Ω *componentwise*, that is, for every $D(t_1, \dots, t_m) \in \Omega$ and $\varphi \in \mathrm{Aut}(T)$, we have

$$D(t_1, \dots, t_m)^{\varphi} = D(t_1^{\varphi}, \dots, t_m^{\varphi}).$$

Two comments are in order. First, this does indeed define an action of $\mathrm{Aut}(T)$ on Ω, because D is $\mathrm{Aut}(T)$-invariant. Second, the inner automorphisms of $\mathrm{Aut}(T)$ induce on Ω permutations appearing in S. Let us denote by ι_x the innerautomorphism of T defined by $x \in T$. Indeed, for every $D(t_1, \dots, t_m) \in \Omega$ and $x \in T$, we have

$$\begin{aligned} D(t_1, \dots, t_m)^{\iota_x} &= D(t_1^x, \dots, t_m^x) = D(x^{-1}t_1x, \dots, x^{-1}t_mx) \\ &= D(t_1x, \dots, t_mx) = D(t_1, \dots, t_m)^{(x, \dots, x)}. \end{aligned}$$

Therefore, ι_x and $(x, \dots, x)$ induce the same permutation on Ω.

Observe, third, that $\mathrm{Sym}(m)$ acts on Ω by permuting the coordinates. Again, this action is well defined, because D is $\mathrm{Sym}(m)$-invariant. It is easy to see that $\mathrm{Aut}(T)$ and $\mathrm{Sym}(m)$ centralize each other and they normalize S. We define

$$W := S(\mathrm{Aut}(T) \times \mathrm{Sym}(m)) \cong T^m \cdot (\mathrm{Out}(T) \times \mathrm{Sym}(m)).$$

The group W acts primitively on Ω and any subgroup G of W containing the socle S and projecting primitively on $\mathrm{Sym}(m)$ is said to be a primitive group of diagonal type. With the notation just established, we have the following.

LEMMA 4.2. *Let G be a primitive group of SD type with socle T^m, for some nonabelian simple group T and for some integer $m \geq 2$. Then,*

$$\mathrm{I}(G) \leq \begin{cases} \frac{3m}{2} + \log|\mathrm{Aut}(T)|, & \text{when } m \geq 3, \\ \log|T|, & \text{when } m = 2. \end{cases}$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Proof. By Lemma 2.2, we can assume $G = W$. Set $\omega := D(1, \dots, 1)$. A computation shows that

$$W_\omega = \mathrm{Aut}(T) \times \mathrm{Sym}(m).$$

Suppose first that $m \geq 3$. Lemma 2.7 implies that

$$\ell(W_\omega) \leq \ell(\mathrm{Aut}(T)) + \ell(\mathrm{Sym}(m)).$$

By [CST89], we know that $\ell(\mathrm{Sym}(m)) \leq \frac{3m}{2}$. On the other hand, $\ell(\mathrm{Aut}(T)) \leq \log|\mathrm{Aut}(T)| - 1$, where the -1 accounts for the fact that $|T|$ is divisible by at least two distinct odd primes. We conclude that

$$\mathrm{I}(G) \leq \ell(W_\omega) + 1 \leq \frac{3m}{2} + \log|\mathrm{Aut}(T)|. \tag{4.2}$$

Suppose next that $m = 2$. We identify Ω with T. Set $H := T^2 \cdot \mathrm{Out}(T)$, and observe that H has a regular normal subgroup and that $|W : H| = 2$. Arguing as in the proof of Proposition 3.1, we deduce that

$$\mathrm{I}(H) \leq \omega(|T|) + 1,$$

where $\omega(|T|)$ is the number of prime divisors of $|T|$. Write $|T| = 2^v o$, where $v, o \in \mathbb{N}$ and o is odd. If $o \geq 43$, we have $\log_3(o) \leq \log(o) - 2$, and hence, $\omega(|T|) \leq \log(2^v)\log_3(o) \leq \log|T| - 2$, and hence, $\mathrm{I}(H) \leq \log|T| - 1$. If $o < 43$, then T has a Sylow 2-subgroup of index at most 41 in T, and hence, T admits a faithful primitive permutation representation of degree at most 41. Thus, we have only a finite number of simple groups satisfying this property. A direct analysis yields that T is either $\mathrm{Alt}(5)$ or $\mathrm{PSL}_3(2)$. When $T = \mathrm{PSL}_3(2)$, we have $\omega(|T|) = 5 < \log|T| - 2$, and hence, we obtain again $\mathrm{I}(H) \leq \log|T| - 1$. Therefore, except when $T = \mathrm{Alt}(5)$, as $|G : H| = 2$, we have $\mathrm{I}(G) \leq \mathrm{I}(H) + 1 \leq \log|T|$. When $T = \mathrm{Alt}(5)$, we have checked that $\mathrm{I}(G) = 5 \leq \log|T|$. □

Proof of Proposition 4.1. When $m = 2$, the proof follows immediately from Lemma 4.2. Assume that $m \geq 3$. Lemma 4.2 and (4.1) imply that it is sufficient to prove that

$$\frac{3m}{2} + \log|\mathrm{Aut}(T)| \leq (m-1)\log|T|.$$

We argue by contradiction, and we suppose that this inequality does not hold.

In [Qui04, Lemma 2.2], it is shown that, for a nonabelian simple group T,

$$\frac{|T|}{|\mathrm{Out}(T)|} \geq 30.$$

Now, for a centerless group, we have $|\mathrm{Aut}(T)| = |\mathrm{Out}(T)||T|$; hence, $|\mathrm{Aut}(T)| \leq \frac{1}{30}|T|^2$. Thus,

$$\frac{3m}{2} + \log|\mathrm{Aut}(T)| \leq \frac{3m}{2} + 2\log|T| - \log 30.$$

Therefore,

$$(m-1)\log|T| < \frac{3m}{2} + 2\log|T| - \log 30.$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Rearranging the terms and dividing by $\log|T|$, we obtain

$$m-3<\frac{3m}{2\log|T|}-\frac{\log 30}{\log|T|}.$$

An easy computation (using $|T|\geq 60$) shows that this is never satisfied. □

§5. Product actions

In the breakdown described in [Pra90], there are three classes of groups left to deal with to prove Theorems 1.3 and 1.6. In this section, we deal with groups of type CD or PA. Note that our result for type CD will be definitive, whereas our result for type PA will involve input from groups of type AS—and these will be dealt with in the remainder of the paper.

Note, furthermore, that we will need a result for groups of type PA that is specific to $\mathrm{I}(G)$, and another that is specific to $\mathrm{H}(G)$. In all cases, we let H be a primitive nonregular group on Δ of type AS or SD and let n be a positive integer with $n\geq 2$. We define $W:=H\mathrm{wr}\,\mathrm{Sym}(m)$ endowed with its primitive product action on the Cartesian product $\Omega:=\Delta^m$, that is, for every $(h_1\ldots,h_m)\sigma\in W$ and $(\delta_1,\ldots,\delta_m)\in\Omega$, we have

$$(\delta_1,\ldots,\delta_m)^{(h_1,\ldots,h_m)\sigma}=\left(\delta_1^{h_1},\ldots,\delta_m^{h_m}\right)^{\sigma}=\left(\delta_{1^{\sigma^{-1}}}^{h_{1^{\sigma^{-1}}}},\ldots,\delta_{m^{\sigma^{-1}}}^{h_{m^{\sigma^{-1}}}}\right).$$

Let $t:=|\Omega|=|\Delta|^m$, and let $\pi:W\to\mathrm{Sym}(m)$ be the natural projection. Observe that the kernel of π is the base group of W, that is, H^m.

With the notation just established, the result we need is the following.

Proposition 5.1. *Let G be a primitive permutation group on a finite set, $\Omega=\Delta^m$, of size t.*

1. *If H is of type SD, then $\mathrm{I}(G)<2\log t$.*
2. *If H is of type AS and $\mathrm{I}(H,\Delta)<C\log|\Delta|$, then $\mathrm{I}(G)<(C+\frac{3}{2\log|\Delta|})\log t$.*
3. *If H is of type AS and $\mathrm{H}(H,\Delta)<C\log|\Delta|$, then $\mathrm{H}(G)<(C+\frac{3}{2\log|\Delta|})\log t$.*

Proof. We use, once again, the fact that a subgroup chain in $\mathrm{Sym}(m)$ has length at most $\frac{3}{2}m$, that is, $\ell(\mathrm{Sym}(m))\leq\frac{3}{2}m$ [CST89].

We start with (1): from Lemma 2.8, we have

$$\mathrm{I}(W,\Omega)\leq\mathrm{I}(H^m,\Delta^m)+\ell(W/H^m)=\mathrm{I}(H^m,\Delta^m)+\ell(\mathrm{Sym}(m))\leq\mathrm{I}(H^m,\Delta^m)+\frac{3m}{2}.$$

Now, Lemma 2.6 yields $\mathrm{I}(G,\Omega)\leq m\mathrm{I}(H,\Delta)+3m/2$. Thus,

$$\mathrm{I}(W,\Omega)\leq m\mathrm{I}(H,\Delta)+\frac{3}{2}m. \tag{5.1}$$

Let $G\leq W$ be primitive. Then, we obtain that

$$\mathrm{I}(G,\Omega)\leq m\mathrm{I}(H,\Delta)+\frac{3m}{2}.$$

If H is of type SD, then Proposition 4.1 implies that $\mathrm{I}(H,\Delta)\leq\log|\Delta|$, and hence,

$$\mathrm{I}(G)\leq m\log|\Delta|+\frac{3}{2}m<2m\log|\Delta|=2\log t,$$

where the second inequality follows with a computation using $|\Delta|\geq 60$.

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

For (2) and (3), recall that if H is of type AS, then G is of type PA. Now, for item (2), we know, by supposition, that $\mathrm{I}(H) < C\log t$. This fact combined with (5.1) yields the result. For item (3), the argument is the same, provided all occurrences of $\mathrm{I}(X,Y)$ are replaced with $\mathrm{H}(X,Y)$ (for varying X and Y) and replacing Lemma 2.6 with Lemma 2.5. □

§6. Almost simple groups: Bounds for I(G)

Here, we deal with groups *of type AS* in the language of [Pra90]; in other words, we study almost simple primitive permutation groups. The main result of this section is the following.

PROPOSITION 6.1. *Let G be a primitive almost simple permutation group on a set, Ω, of size t. Either $\mathrm{I}(G) \leq 6\log t$ or else G is one of the groups listed at (1) in Theorem 1.6.*

The proof of Proposition 6.1 splits into several parts. To start with, we use well-known results bounding $\mathrm{b}(G)$ to deal with the so-called *non-standard actions*. The terminology below follows [Bur07].

DEFINITION 6.2. Let G be a classical group with socle G_0, and associated natural module V. A subgroup H of G not containing G_0 is a *subspace subgroup* if for each maximal subgroup M of G_0 containing $H \cap G_0$ one of the following holds:

1. M is the stabilizer in G_0 of a proper nonzero subspace U of V, where U is totally singular, nondegenerate, or, if G_0 is orthogonal and $p = 2$, a nonsingular 1-space (U can be any subspace if $G_0 = \mathrm{PSL}(V)$).
2. $G_0 = \mathrm{Sp}_{2m}(q)$, $p = 2$, and $M = \mathrm{O}^{\pm}_{2m}(q)$.

A *subspace action* of the classical group G is the action of G on the coset space $[G : H]$, where H is a subspace subgroup of G.

Note that the definition above amounts precisely to this: a maximal subgroup of G is a subspace subgroup if it lies in any $\mathcal{C}_1$ class, or is the even-characteristic symplectic case in the $\mathcal{C}_8$ class. This definition requires that we follow [KL90] in labeling the classes $\mathcal{C}_1$– $\mathcal{C}_8$. A small extra collection of maximal subgroups arises when $G_0 = \mathrm{Sp}_4(2^a)$, and G contains a graph automorphism, or if $G_0 = \mathrm{P}\Omega^+_8(q)$ and G contains a triality graph automorphism. We note that [KL90] explicitly exclude these cases.

DEFINITION 6.3. A transitive action of G on a set Ω is said to be *standard* if, up to equivalence of actions, one of the following holds:

1. $G_0 = \mathrm{Alt}(m)$ and Ω is an orbit of subsets or uniform partitions of $\{1,\ldots,m\}$.
2. G is a classical group in a subspace action.

For an almost simple primitive permutation group in a nonstandard action, the base size is bounded by an absolute constant. This was conjectured by Cameron and Kantor ([Cam92], [CK97]) and then settled in the affirmative by Liebeck and Shalev in [LS99, Theorem 1.3]. The constant was then made explicit in subsequent work [BGS11], [BLS09], [BOW10], [Bur07]. The following theorem summarizes these results.

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

THEOREM 6.4. *Let G be a finite almost simple group in a primitive faithful nonstandard action. Then,* $\mathrm{b}(G) \leq 7$*, with equality if and only if G is the Mathieu group M_{24} in its natural action of degree 24.*

Theorem 6.4 and (1.1) immediately yield Proposition 6.1 for nonstandard actions.

LEMMA 6.5. *Let G be a finite almost simple permutation group on a set Ω of size t, in a nonstandard action. Then,*

$$\mathrm{I}(G) \leq 6 \log t.$$

One can compute that $\mathrm{I}(M_{24}) = 8$ (where we consider the natural action of M_{24} on 24 points); hence, we have a constant *6* in the statement of Lemma 6.5 rather than *7*.

6.1 Standard actions of A_n and S_n

Let G be $\mathrm{Sym}(n)$ or $\mathrm{Alt}(n)$. We must prove Theorem 1.6 for the action of G on partitions of n. Let $n = ab$ for some positive integers a, b with $a \geq 2$ and $b \geq 2$. We denote by Ω_a^b the set of all uniform partitions of n, with a parts of size b. Let $t = |\Omega_a^b|$, then

$$t = \frac{(ab)!}{a!(b!)^a}.$$

We consider the action of G on Ω_a^b. We have the following result.

LEMMA 6.6. *Let G be an almost simple group with socle $\mathrm{Alt}(n)$ acting on $\Omega = \Omega_a^b$. Then,*

$$\mathrm{I}(G) < 2 \log t.$$

Proof. By [GMPS15, Lemma 5.6], we know that if $a \geq 2$, $b \geq 2$, and $n \geq 17$, then

$$\frac{(ab)!}{(b!)^a (a!)} \geq 3^{ab/2}.$$

Hence,

$$\log t \geq \frac{\log(3)}{2} ab. \tag{6.1}$$

Once again, we use the fact that $\ell(\mathrm{Sym}(n)) \leq \frac{3}{2}n = \frac{3}{2}ab$. As $3/2 < \log(3)$ for $n \geq 17$, we have

$$\mathrm{I}(G) \leq \ell(G) \leq \frac{3}{2} ab < \log(3) ab \leq 2 \log t,$$

and the result follows. If $n < 17$, we check directly that $\frac{3}{2}ab < 2 \log t$, unless

$$(a, b) \in \{(2,3), (3,2), (4,2), (5,2), (6,2)\}.$$

For these remaining cases, we have computed explicitly the value of $\mathrm{Sym}(n)$ acting on partitions of $\{1, \ldots, n\}$ in b parts of cardinality a, and we have verified that in each case $\mathrm{I}(G) < 2 \log t$. □

6.2 The symplectic/orthogonal case

In this section, we deal with the actions listed at item (2) of Definition 6.2. In particular, G is almost simple with socle $G_0 = \mathrm{Sp}_{2m}(q)$, q a power of 2. Consulting [BHR13] and [KL90], it is clear that if H is a maximal subgroup of G not containing G_0, and $H \cap G_0$ is

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

a subgroup of $M = \mathrm{O}^{\pm}_{2m}(q)$, then $H \cap G_0 = M$. In light of this, the result that we need is the following.

LEMMA 6.7. *Let G be almost simple with socle $G_0 = \mathrm{Sp}_{2m}(q)$, q a power of 2. Let H be a subgroup of G such that $H \cap \mathrm{Sp}_{2m}(q) = \mathrm{O}^{\pm}_{2m}(q)$, let Ω be the set of cosets of H in G, and write $t = |\Omega|$. Then,*

$$\mathrm{I}(G,\Omega) < \frac{11}{3}\log t.$$

The treatment that follows is inspired by [DM96, Section 7.7], where the case of $\mathrm{Sp}_{2m}(2)$ is considered. Let e and m be positive integers, let $q := 2^e$, let $\mathbb{F}_q$ be the finite field with q elements, and let $V := \mathbb{F}_q^{2m}$ be the $2m$-dimensional vector space of row vectors over $\mathbb{F}_q$. To start with, we adjust notation slightly, and assume that G is simple: let $G := \mathrm{Sp}_{2m}(q)$ be the symplectic group defined by the symmetric matrix

$$f := \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix},$$

where 0 and I are the zero and identity $m \times m$-matrices, respectively. In particular, G is the group of invertible matrices preserving the bilinear form $\varphi : V \times V \to \mathbb{F}_q$ defined by

$$\varphi(u,v) := ufv^T,$$

for every $u, v \in V$, that is,

$$G = \left\{ g \in \mathrm{GL}_{2m}(q) \mid gfg^T = f \right\}.$$

Note that the bilinear form φ is alternating, that is, for all $u \in V$, we have

$$\varphi(u,u) = 0. \tag{6.2}$$

Moreover, since $\mathbb{F}_q$ is of characteristic 2, the form φ is symmetric, that is, for all $u, v \in V$, we have

$$\varphi(u,v) = \varphi(v,u). \tag{6.3}$$

Now, we let Ω be the set of quadratic forms $\theta : V \to \mathbb{F}_q$ polarizing to φ. Recall that this means that $\theta : V \to \mathbb{F}_q$ is a function satisfying

- $\theta(u+v) - \theta(u) - \theta(v) = \varphi(u,v)$, for every $u, v \in V$, and
- $\theta(cu) = c^2 u$, for every $c \in \mathbb{F}_q$ and $u \in V$.

Next, consider the matrix

$$e := \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix}$$

and the quadratic form $\theta_0 : V \to \mathbb{F}_q$ defined by

$$\theta_0(u) := ueu^T,$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

for every $u \in V$. For every $u, v \in V$, we have

$$\begin{aligned}\theta_0(u+v)-\theta_0(u)-\theta_0(v) &:= (u+v)e(u+v)^T - ueu^T - vev^T \\ &= ueu^T + vev^T + uev^T + veu^T - ueu^T - vev^T \\ &= uev^T + veu^T = uev^T + ue^Tv^T = u(e+e^T)v^T = ufv^T \\ &= \varphi(u,v).\end{aligned} \tag{6.4}$$

In particular, θ_0 is a quadratic form whose polarisation is the symplectic form φ, and hence, $\theta_0 \in \Omega$.

Let $\theta \in \Omega$ and define $\lambda := \theta - \theta_0$. We have

$$\begin{aligned}\lambda(u+v) &= \theta(u+v) - \theta_0(u+v) = \theta(u) + \theta(v) + \varphi(u,v) - \theta_0(u) - \theta_0(v) - \varphi(u,v) \\ &= \lambda(u) + \lambda(v), \\ \lambda(cu) &= \theta(cu) - \theta_0(cu) = c^2\theta(u) - c^2\theta_0(u) = c^2\lambda(u),\end{aligned}$$

for every $u, v \in V$ and for every $c \in \mathbb{F}_q$. Therefore, since $\mathbb{F}_q$ is of characteristic 2, the function $\lambda : V \to \mathbb{F}_q$ is semilinear, and hence, there exists a unique $b \in V$ such that $\lambda(u) = (u \cdot b^T)^2$, for every $u \in V$ (see Lemma 7.3 for a precise statement). Since f is an invertible matrix, there exists a unique $a \in V$ with $b = af$, and hence,

$$\lambda(u) = (ufa^T)^2 = \varphi(u,a)^2,$$

for every $u \in V$. Summing up, we have shown that an arbitrary element of Ω is of the form

$$u \mapsto \theta_0(u) + \varphi(u,a)^2,$$

where $a \in V$. We denote this element of Ω simply by θ_a. Thus,

$$\theta_a(u) = \theta_0(u) + \varphi(u,a)^2, \quad \text{for every } u \in V. \tag{6.5}$$

In particular, the elements of Ω are parameterized by the vectors of V. Moreover, if $\theta_a = \theta_{a'}$, for some $a, a' \in V$, then $\theta_a(u) = \theta_{a'}(u)$, for every $u \in V$, and this implies $\varphi(u,a) = \varphi(u,a')$, for every $u \in V$. Since φ is nondegenerate, we obtain $a = a'$. Hence, the set Ω is in one-to-one correspondence with V. This, in particular, yields that $|\Omega| = q^{2m}$.

LEMMA 6.8. *The group G acts on the set Ω.*

Proof. First, we show that for every $x \in G$ and for every $\theta \in \Omega$, the mapping

$$\begin{aligned}\theta^x &: V \to \mathbb{F}_q \\ &u \mapsto \theta(ux^{-1})\end{aligned} \tag{6.6}$$

gives rise to an element of Ω. For every $u, v \in V$, we have

$$\begin{aligned}\theta^x(u+v) &= \theta((u+v)x^{-1}) = \theta(ux^{-1} + vx^{-1}) = \theta(ux^{-1}) + \theta(vx^{-1}) + \varphi(ux^{-1}, vx^{-1}) \\ &= \theta(ux^{-1}) + \theta(vx^{-1}) + \varphi(u,v) = \theta^x(u) + \theta^x(v) + \varphi(u,v).\end{aligned}$$

Observe that in the fourth equality, we used the fact that $x \in G$ and hence x preserves the bilinear form φ. Moreover,

$$\theta^x(cu) = \theta((cu)x^{-1}) = \theta(c(ux^{-1})) = c^2\theta(ux^{-1}) = c^2\theta^x(u).$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Therefore, $\theta^x \in \Omega$. Finally, for every $x, y \in G$, $\theta \in \Omega$ and $u \in V$, we have

$$(\theta^x)^y(u) = \theta^x(uy^{-1}) = \theta((uy^{-1})x^{-1}) = \theta(u(xy)^{-1}) = \theta^{xy}(u).$$

Therefore, $(\theta^x)^y = \theta^{xy}$, and hence, G defines a genuine right action on Ω. □

Before continuing our discussion, we gather some information on G. Let $a \in V$, and we define the mapping

$$\begin{aligned} t_a \colon \ & V \to V \\ & u \mapsto u + \varphi(u,a)a. \end{aligned} \tag{6.7}$$

Such a function is called a *transvection*. For every $u, v \in V$ and $c \in \mathbb{F}_q$, we have

$$\begin{aligned} (u+v)t_a &= (u+v) + \varphi(u+v,a)a = (u+\varphi(u,a)a) + (v+\varphi(v,a)a) = (u)t_a + (v)t_a; \\ (cu)t_a &= cu + \varphi(cu,a)a = c(u+\varphi(u,a)a) = c(u)t_a. \end{aligned}$$

Hence, t_a is linear. Moreover, for every $u \in V$, we have

$$\begin{aligned} (u)t_a^2 &= (u+\varphi(u,a)a)t_a = u + \varphi(u,a)a + \varphi(u+\varphi(u,a)a,a)a \\ &= u + \varphi(u,a)a + \varphi(u,a)a + \varphi(u,a)\varphi(a,a)a = u + 2\varphi(u,a)a + \varphi(u,a)\varphi(a,a)a \\ &= u + \varphi(u,a)\varphi(a,a)a \underset{6.2}{=} u, \end{aligned}$$

where in the second-last equality we use the fact that the characteristic of $\mathbb{F}_q$ is 2. This shows that t_a is an involution. Finally, for every $u, v \in V$, we have

$$\begin{aligned} \varphi((u)t_a,(v)t_a) &= \varphi(u+\varphi(u,a)a, v+\varphi(v,a)a) \\ &= \varphi(u,v) + \varphi(v,a)\varphi(u,a) + \varphi(u,a)\varphi(a,v) + \varphi(u,a)\varphi(v,a)\varphi(a,a) \\ &\underset{6.3}{=} \varphi(u,v) + 2\varphi(v,a)\varphi(u,a) + \varphi(u,a)\varphi(v,a)\varphi(a,a) \\ &= \varphi(u,v) + \varphi(u,a)\varphi(v,a)\varphi(a,a) \\ &\underset{6.2}{=} \varphi(u,v). \end{aligned}$$

Therefore, t_a preserves φ and hence lies in the symplectic group G.

We are now interested in computing the image of θ_a under the transvection t_c. First, recall that, in a field of characteristic 2, since $x = -x$ for every $x \in \mathbb{F}_q$, the square root $\sqrt{\cdot}\colon \mathbb{F}_q \to \mathbb{F}_q$ is a well-defined map. Moreover, for every $a, b, x, y \in \mathbb{F}_q$ such that $x = a^2$ and $y = b^2$, we have $(\sqrt{x}+\sqrt{y})^2 = (a+b)^2 = a^2 + b^2 = x + y$, which implies $\sqrt{x}+\sqrt{y} = \sqrt{x+y}$. Moreover, recall that θ_a is a quadratic form polarising to φ and that t_a is an involution, in particular $t_c = t_c^{-1}$. By using these facts, given $v \in V$, we have

$$\begin{aligned} \theta_a^{t_c}(u) &\underset{6.6}{=} \theta_a(ut_c^{-1}) = \theta_a(ut_c) \underset{6.7}{=} \theta_a(u + \varphi(u,c)c) \\ &= \theta_a(u) + \theta_a(\varphi(u,c)c) + \varphi(u,\varphi(u,c)c) \\ &= \theta_a(u) + \varphi(u,c)^2\theta_a(c) + \varphi(u,c)^2 \\ &= \theta_a(u) + (\theta_a(c)+1)\varphi(u,c)^2 = \theta_a(u) + (\sqrt{\theta_a(c)+1}\varphi(u,c))^2 \\ &= \theta_a(u) + \varphi(u,(\sqrt{\theta_a(c)}+1)c)^2 \underset{6.5}{=} \theta_0(u) + \varphi(u,a)^2 + \varphi(u,(\sqrt{\theta_a(c)}+1)c)^2 \\ &= \theta_0(u) + \varphi(u,a+(\sqrt{\theta_a(c)}+1)c)^2 \\ &\underset{6.5}{=} \theta_{a+(\sqrt{\theta_a(c)}+1)c}(u). \end{aligned}$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

From this, we deduce

$$\theta_a^{t_c} = \theta_{a+(\sqrt{\theta_a(c)}+1)c}. \tag{6.8}$$

We now recall some facts about Galois theory. For a reference, see [Lan02, Chapter VI]. The Frobenius mapping $\phi\colon x \mapsto x^2$ from $\mathbb{F}_q$ to itself is a generator of the Galois group of $\mathbb{F}_q$ over $\mathbb{F}_2$. There exists a well-defined $\mathbb{F}_2$-linear *trace* mapping $\mathrm{Tr} : \mathbb{F}_q \to \mathbb{F}_2$. In what follows, we need only two basic facts about Tr: first, Tr is surjective, and second, from Hilbert's 90 theorem, the kernel of Tr consists of the set $\{x^2 + x \mid x \in \mathbb{F}_q\}$ and has cardinality $q/2$.

Define

$$\Omega^+ := \{\theta_a \mid \mathrm{Tr}(\theta_0(a)) = 0\},$$
$$\Omega^- := \{\theta_a \mid \mathrm{Tr}(\theta_0(a)) = 1\}.$$

Observe that the above definition is a generalization of the definition of Ω^+ and Ω^- in [DM96, Corollary 7.7 A]. Indeed, if $q = 2$, then the Galois group is the trivial group, and hence, the trace map is the identity.

Let $N := \langle t_a \mid a \in V\rangle$ be the subgroup of G generated by the transvections. Observe that, for all $a, u \in V$ and $x \in G$, we have

$$(u)x^{-1}t_a x = (ux^{-1} + \varphi(ux^{-1}, a)a)x = u + \varphi(u, ax)ax = (u)t_{ax}.$$

In particular, this shows that $N \trianglelefteq G$.

LEMMA 6.9. *The sets Ω^+ and Ω^- are N-orbits on Ω, with*

$$|\Omega^+| = \frac{q^m(q^m+1)}{2}, \quad |\Omega^-| = \frac{q^m(q^m-1)}{2}.$$

These are also orbits for G.

Proof. We first prove that

$$|\Omega^+| = \frac{q^m(q^m+1)}{2}, \quad |\Omega^-| = \frac{q^m(q^m-1)}{2}.$$

Clearly, it suffices to prove the first equality, because the second follows from the equality

$$|\Omega^-| = |\Omega| - |\Omega^+| = q^{2m} - |\Omega^+|.$$

By definition, $\theta_0(a) = aea^T$. Moreover, by using the canonical basis of V, we have

$$\mathrm{Tr}\left(\theta_0\left(\sum_{i=1}^{2m} a_i e_i\right)\right) = \mathrm{Tr}\left(\sum_{i=1}^{m} a_i a_{i+m}\right) = \sum_{i=1}^{m} \mathrm{Tr}(a_i a_{i+m}).$$

Using this equation, we may compute the cardinality of $|\Omega^+|$ arguing inductively on m. When $m = 1$, we have $\mathrm{Tr}(a_1 a_2) = 0$. When $a_1 = 0$, we have q solutions for a_2; however, for every $a_1 \in \mathbb{F}_q \setminus \{0\}$, we have $q/2$ solutions for a_2. Therefore, when $m = 1$, we have

$$q + (q-1)\frac{q}{2} = \frac{q^2}{2} + \frac{q}{2} = \frac{q(q+1)}{2}$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

solutions to the equation $\mathrm{Tr}(a_1a_2)=0$ and we have $\frac{q(q-1)}{2}$ solutions to the equation $\mathrm{Tr}(a_1a_2)=1$. Arguing inductively, we may assume that

$$\sum_{i=1}^{m-1}\mathrm{Tr}(a_ia_{i+m})$$

is equal to 0 for $q^{m-1}(q^{m-1}+1)/2$ choices of $(a_1,\dots,a_{m-1})$ and is equal to 1 for $q^{m-1}(q^{m-1}-1)/2$ choices of $a_1,\dots,a_{m-1}$. Now,

$$\sum_{i=1}^{m}\mathrm{Tr}(a_ia_{i+m})=\sum_{i=1}^{m-1}\mathrm{Tr}(a_ia_{i+m})+\mathrm{Tr}(a_ma_{2m})$$

has value 0 if and only if $\sum_{i=1}^{m-1}\mathrm{Tr}(a_ia_{i+m})$ and $\mathrm{Tr}(a_ma_{2m})$ have the same value. Therefore, altogether, the number of solutions of $\sum_{i=1}^{m}\mathrm{Tr}(a_ia_{i+m})=0$ is

$$\frac{q^{m-1}(q^{m-1}+1)}{2}\frac{q(q+1)}{2}+\frac{q^{m-1}(q^{m-1}-1)}{2}\frac{q(q-1)}{2}=\frac{q^m(q^m+1)}{2}.$$

Next, we prove that Ω^+ and Ω^- are N-orbits. We start by considering Ω^+. We first prove that Ω^+ is N-invariant. To this end, it suffices to show that if $\theta_a\in\Omega^+$ and $c\in V$, then $\theta_a^{t_c}\in\Omega^+$. In other words, using (6.8), if $\mathrm{Tr}(\theta_0(a))=0$, then $\mathrm{Tr}(\theta_0(a+(\sqrt{\theta_a(c)}+1)c))=0$. So let $\mathrm{Tr}(\theta_0(a))=0$ and recall that Tr is a linear map. We have

$$\begin{aligned}\mathrm{Tr}(\theta_0(a+(\sqrt{\theta_a(c)}+1)c)) &\underset{6.4}{=}\mathrm{Tr}\left(\theta_0(a)+\theta_0((\sqrt{\theta_a(c)}+1)c)+\varphi(a,(\sqrt{\theta_a(c)}+1)c)\right)\\ &=\mathrm{Tr}\left(\theta_0(a)\right)+\mathrm{Tr}\left(\theta_0((\sqrt{\theta_a(c)}+1)c)\right)+\mathrm{Tr}\left(\varphi(a,(\sqrt{\theta_a(c)}+1)c)\right)\\ &=\mathrm{Tr}\left(\theta_0((\sqrt{\theta_a(c)}+1)c)\right)+\mathrm{Tr}\left(\varphi(a,(\sqrt{\theta_a(c)}+1)c)\right)\\ &=\mathrm{Tr}\left((\sqrt{\theta_a(c)}+1)^2\theta_0(c)\right)+\mathrm{Tr}\left(\varphi(a,(\sqrt{\theta_a(c)}+1)c)\right)\\ &=\mathrm{Tr}\left((\theta_a(c)+1)\theta_0(c)\right)+\mathrm{Tr}\left((\sqrt{\theta_a(c)}+1)\varphi(a,c)\right)\\ &=\mathrm{Tr}\left(\theta_a(c)\theta_0(c)\right)+\mathrm{Tr}\left(\theta_0(c)\right)+\mathrm{Tr}(\sqrt{\theta_a(c)}\varphi(a,c))+\mathrm{Tr}(\varphi(a,c)).\end{aligned}\tag{6.9}$$

Using $\theta_a(c)=\theta_0(c)+\varphi(a,c)^2$ and $\mathrm{Tr}(x^2)=\mathrm{Tr}(x)$, for every $x\in\mathbb{F}_q$, we obtain

$$\begin{aligned}\mathrm{Tr}\left(\theta_a(c)\theta_0(c)\right) &=\mathrm{Tr}(\theta_0(c)^2)+\mathrm{Tr}(\varphi(a,c)^2\theta_0(c))\\ &=\mathrm{Tr}(\theta_0(c))+\mathrm{Tr}(\varphi(a,c)^2\theta_0(c)),\end{aligned}\tag{6.10}$$

$$\begin{aligned}\mathrm{Tr}(\sqrt{\theta_a(c)}\varphi(a,c)) &=\mathrm{Tr}(\sqrt{\theta_0(c)}\varphi(a,c))+\mathrm{Tr}(\varphi(a,c)^2)\\ &=\mathrm{Tr}(\sqrt{\theta_0(c)}\varphi(a,c))+\mathrm{Tr}(\varphi(a,c)).\end{aligned}\tag{6.11}$$

Putting (6.10) and (6.11) into (6.9), we obtain

$$\begin{aligned}\mathrm{Tr}(\theta_0(a+(\sqrt{\theta_a(c)}+1)c)) =&\mathrm{Tr}(\theta_0(c))+\mathrm{Tr}(\varphi(a,c)^2\theta_0(c))+\mathrm{Tr}\left(\theta_0(c)\right)\\ &+\mathrm{Tr}(\sqrt{\theta_0(c)}\varphi(a,c))+\mathrm{Tr}(\varphi(a,c))+\mathrm{Tr}(\varphi(a,c))\\ =&0.\end{aligned}$$

As $\Omega^-=\Omega\setminus\Omega^+$, we obtain that Ω^- is also N-invariant.

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Next, we show that Ω^+ is an N-orbit. Actually, we prove something stronger, we show that

$$\Omega^+ = \{\theta_0^{t_c} \mid c \in V\}.$$

For every $a \in V$ with $\mathrm{Tr}(\theta_0(a)) = 0$, we need to show that there exists $c \in V$ such that

$$\theta_a = \theta_0^{t_c}. \tag{6.12}$$

If $\theta_0(a) = 0$, then we may take $c := a$ and (6.8) yields

$$\theta_0^{t_a} = \theta_{0+(\sqrt{\theta_0(a)}+1)a} = \theta_a$$

and we are finished. Suppose $\theta_0(a) \neq 0$. Since $\mathrm{Tr}(\theta_0(a)) = 0$, from Hilbert's 90 theorem, there exists $x \in \mathbb{F}_q$ with $\theta_0(a) = x + x^2$. Since $\theta_0(a) \neq 0$, we have $x \neq 0$. Let $y \in \mathbb{F}_q \setminus \{0\}$ with $y^2 = x$, and set $c := y^{-1}a$. Thus,

$$\begin{aligned}\sqrt{\theta_0(c)} + 1 &= \sqrt{\theta_0(y^{-1}a)} + 1 = \sqrt{y^{-2}\theta_0(a)} + 1 = y^{-1}\sqrt{\theta_0(a)} + 1 \\ &= y^{-1}\sqrt{x + x^2} + 1 = y^{-1}(y + y^2) + 1 = y.\end{aligned}$$

From (6.8), we have

$$\theta_0^{t_c} = \theta_{0+(\sqrt{\theta_0(c)}+1)c} = \theta_{yc} = \theta_{yy^{-1}a} = \theta_a.$$

Next, we show that Ω^- is an N-orbit. Actually, we prove something stronger, we show that

$$\Omega^- = \{\theta_\varepsilon^{t_c} \mid c \in V\}.$$

First, we select a distinguished element of Ω^-. Let $\epsilon \in \mathbb{F}_q$ with $\mathrm{Tr}(\epsilon) = 1$, and set $\varepsilon := \epsilon e_1 + e_{m+1}$, where $(e_i)_{i\in\{1,\dots,2m\}}$ is the standard basis of V. Since $\theta_0(\epsilon e_1) = 0 = \theta_0(e_{m+1})$, we have

$$\theta_0(\varepsilon) = \theta_0(\epsilon e_1) + \theta_0(e_{m+1}) + \varphi(\epsilon e_1, e_{m+1}) = \epsilon\varphi(e_1, e_{m+1}) = \epsilon,$$

and hence, $\mathrm{Tr}(\theta_0(\varepsilon)) = \mathrm{Tr}(\epsilon) = 1$. Therefore, $\theta_\varepsilon \in \Omega^-$.

For every $a \in V$ with $\mathrm{Tr}(\theta_0(a)) = 1$, we need to show that there exists $c \in V$ such that

$$\theta_a = \theta_\varepsilon^{t_c}. \tag{6.13}$$

If $\theta_\varepsilon(a+\varepsilon) = 0$, then we may take $c := a + \varepsilon$ and (6.8) yields

$$\theta_\varepsilon^{t_c} = \theta_{\varepsilon+(\sqrt{\theta_\varepsilon(c)}+1)c} = \theta_{\varepsilon+c} = \theta_a$$

and we are finished. Suppose $\theta_\varepsilon(a+\varepsilon) \neq 0$. We have

$$\begin{aligned}\theta_\varepsilon(a+\varepsilon) &\underset{6.5}{=} \theta_0(a+\varepsilon) + \varphi(a+\varepsilon,\varepsilon)^2 = \theta_0(a+\varepsilon) + \varphi(a,\varepsilon)^2 + \varphi(\varepsilon,\varepsilon)^2 \\ &\underset{6.2}{=} \theta_0(a+\varepsilon) + \varphi(a,\varepsilon)^2 \\ &\underset{6.4}{=} \theta_0(a) + \theta_0(\varepsilon) + \varphi(a,\varepsilon) + \varphi(a,\varepsilon)^2.\end{aligned}$$

Since $\mathrm{Tr}(\theta_0(a)) = 1 = \mathrm{Tr}(\theta_0(\varepsilon))$, by using the previous equality, we deduce that

$$\mathrm{Tr}(\theta_\varepsilon(a+\varepsilon)) = 0.$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

From Hilbert's 90 theorem, there exists $x \in \mathbb{F}_q$ with $\theta_\varepsilon(a+\varepsilon) = x + x^2$. Since $\theta_\varepsilon(a+\varepsilon) \neq 0$, we have $x \neq 0$. Let $y \in \mathbb{F}_q \setminus \{0\}$ with $y^2 = x$ and set $c := y^{-1}(a+\varepsilon)$. Thus,

$$\begin{aligned}\sqrt{\theta_\varepsilon(c)}+1 &= \sqrt{\theta_\varepsilon(y^{-1}(a+\varepsilon))}+1 = \sqrt{y^{-2}\theta_\varepsilon(a+\varepsilon)}+1 = y^{-1}\sqrt{\theta_\varepsilon(a+\varepsilon)}+1\\ &= y^{-1}\sqrt{x+x^2}+1 = y^{-1}(y+y^2)+1 = y.\end{aligned}$$

From (6.8), we have

$$\theta_\varepsilon^{t_c} = \theta_{\varepsilon+(\sqrt{\theta_\varepsilon(c)}+1)c} = \theta_{\varepsilon+yc} = \theta_{\varepsilon+yy^{-1}(a+\varepsilon)} = \theta_a.$$

The fact that Ω^- and Ω^+ are also G-orbits follows from the fact that $N \trianglelefteq G$ and $|\Omega^+| \neq |\Omega^-|$. □

We now compute the maximum length of a stabilizer chain in these two actions. Let $\epsilon \in \{+,-\}$ (we deal simultaneously with both cases). Let $\{\theta_{a_1}, \theta_{a_2}, \ldots, \theta_{a_k}\}$ be a subset of Ω^ϵ such that the corresponding chain of stabilizers is strictly decreasing. Without loss of generality, we may suppose that $a_1 = 0$ when $\epsilon = +$ and $a_1 = \varepsilon$ when $\epsilon = -$. Recall that ε is an element of V for which θ_ε is a distinguished element of Ω^-; the definition of ε was given in the proof of the previous lemma.

Let us define

$$C_{G_{\theta_{a_1}}}(a_1+a_i) = \{x \in G_{\theta_{a_1}} \mid (a_1+a_i)x = a_1+a_i\},$$

that is, the set of matrices in $G_{\theta a_1}$ fixing the vector $a_1 + a_i \in V$.

LEMMA 6.10. *For every $i \in \{2,\ldots,k\}$,*

$$G_{\theta_{a_1}} \cap G_{\theta_{a_i}} = C_{G_{\theta a_1}}(a_1+a_i). \tag{6.14}$$

Proof. First observe that, given $j \in \{1,\ldots,k\}$, $x \in G_{\theta_{a_j}}$ if and only if, for every $u \in V$, we have

$$\theta_0(ux^{-1}) + \varphi(ux^{-1}, a_j)^2 = \theta_0(u) + \varphi(u,a_j)^2. \tag{6.15}$$

Fix $i \in \{2,\ldots,k\}$ and let $x \in G_{\theta_{a_1}} \cap G_{\theta_{a_i}}$. Then, by applying (6.15) with $j := 1$ and with $j := i$ and by using the fact that the characteristic of $\mathbb{F}_q$ is 2, we obtain that, for every $u \in V$,

$$\varphi(ux^{-1}, a_1+a_i)^2 = \varphi(u, a_1+a_i)^2.$$

Since φ is G-invariant, then $\varphi(ux^{-1}, a_1+a_i)^2 = \varphi(u,(a_1+a_i)x)^2$. Therefore, as φ is nondegenerate, $(a_1+a_i)x = a_1+a_i$ and hence $x \in C_{G_{\theta_{a_1}}}(a_1+a_i)$.

Conversely, let $x \in C_{G_{\theta_{a_1}}}(a_1+a_i)$. Then, $x \in G_{\theta_{a_1}}$ and $(a_1+a_i)x = a_1+a_i$. We need to show that $x \in G_{\theta_{a_i}}$. Note that (6.15) applied with $j = 1$ and the fact that we are in characteristic 2 imply that

$$\theta_0(ux^{-1}) = \theta_0(u) + \varphi(u,a_1)^2 + \varphi(ux^{-1}, a_1)^2. \tag{6.16}$$

For every $u \in V$, we have:

$$\begin{aligned}\theta_{a_i}^x(u) &\underset{6.6}{=} \theta_{a_i}(ux^{-1}) \underset{6.5}{=} \theta_0(ux^{-1}) + \varphi(ux^{-1}, a_i)^2\\ &\underset{6.16}{=} \theta_0(u) + \varphi(ux^{-1}, a_1)^2 + \varphi(u,a_1)^2 + \varphi(ux^{-1}, a_i)^2\end{aligned}$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

$$\begin{aligned}
&= \theta_0(u) + \varphi(ux^{-1}, a_1 + a_i)^2 + \varphi(u, a_1)^2 \\
&= \theta_0(u) + \varphi(u, (a_1 + a_i)x)^2 + \varphi(u, a_i)^2 \\
&= \theta_0(u) + \varphi(u, a_1 + a_i)^2 + \varphi(u, a_1)^2 \\
&= \theta_0(u) + \varphi(u, a_1)^2 + \varphi(u, a_i)^2 + \varphi(u, a_1)^2 \\
&= \theta_0(u) + 2\varphi(u, a_1)^2 + \varphi(u, a_i)^2 \\
&= \theta_0(u) + \varphi(u, a_i)^2 \underset{6.5}{=} \theta_{a_i}(u).
\end{aligned}$$

Observe that in the sixth and seventh equalities we use, respectively, the fact that x preserves φ and the fact that x fixes the vector $(a_1 + a_i)$. Hence, $x \in G_{\theta_{a_i}}$ and therefore $x \in G_{\theta_{a_1}} \cap G_{\theta_{a_i}}$. □

For simplicity, for each $i \in \{1, \dots, k\}$, write $G_i := G_{\theta_{a_1}} \cap \cdots \cap G_{\theta_{a_i}}$ and, for each $i \in \{2, \dots, k\}$, write $b_i := a_1 + a_i$. From (6.14), the strictly decreasing sequence

$$G_1 > G_2 > \cdots > G_{k-1} > G_k$$

equals

$$\begin{aligned}
G_1 >& C_{G_1}(b_2) = C_{G_1}(\langle b_2 \rangle_{\mathbb{F}_q}) > C_{G_1}(b_2, b_3) = C_{G_1}(\langle b_2, b_3 \rangle_{\mathbb{F}_q}) > \cdots \\
&> C_{G_1}(b_2, \dots, b_{k-1}) = C_{G_1}(\langle b_2, \dots, b_{k-1} \rangle_{\mathbb{F}_q}) > C_{G_1}(b_2, \dots, b_k) = C_{G_1}(\langle b_2, \dots, b_k \rangle_{\mathbb{F}_q}).
\end{aligned}$$

Here, if $v_1, \dots, v_j \in V$, then we write $\langle v_1, \dots, v_j \rangle_{\mathbb{F}_q}$ to denote the $\mathbb{F}_q$-vector space generated by the $v_1, \dots, v_j$. We obtain that

$$0 < \langle b_2 \rangle_{\mathbb{F}_q} < \langle b_2, b_3 \rangle_{\mathbb{F}_q} < \cdots < \langle b_2, \dots, b_k \rangle_{\mathbb{F}_q} \leq V.$$

Observe that a strict inclusion in the above chain implies that the dimension has to go down by at least one at each step. Therefore, $k \leq 1 + 2m$. Thus, we have proved the following lemma.

LEMMA 6.11. *Let $G := \mathrm{Sp}_{2m}(q)$ act on Ω^ϵ, where $\epsilon \in \{+, -\}$. Then, $\mathrm{I}(G, \Omega^\epsilon) \leq 1 + 2m$.*

We are now able to prove Lemma 6.7 for $m \geq 3$. The proof is virtually identical when $m = 2$; however, to avoid annoying details, we will use the fact that when $m = 2$ the result follows from Lemma 6.12 which we prove in a moment.

Proof of Lemma 6.7. for $m \geq 3$. Let $q = 2^e$. Since $m \geq 3$,

$$\mathrm{Sp}_{2m}(q) \trianglelefteq G \leq \Gamma\mathrm{Sp}_{2m}(q) = \mathrm{Sp}_{2m}(q) \rtimes \mathcal{C}_e.$$

Since a strictly descending chain of subgroups of $\mathcal{C}_e$ has length at most $\log e$, we conclude that

$$\mathrm{I}(\Gamma\mathrm{Sp}_{2m}(q), \Omega^\epsilon) \leq 2m + 1 + \log e.$$

Suppose $2m + 1 + \log e \geq 2 \log t^\epsilon$. Recall the cardinality of Ω^ϵ from Lemma 6.9. Then,

$$2^{2m+1} \cdot e \geq (t^\epsilon)^2 = 2^{2em-2} \left(2^{em} + \epsilon 1\right)^2 \geq 2^{2em-2} \left(2^{em} - 1\right)^2.$$

When $e \geq 2$, with an easy computation, we obtain a contradiction. When $e = 1$, we have $2^{2m+1} \geq 2^{2m-2}(2^m - 1)^2$, which implies $2^3 \geq (2^m - 1)^2$. Again, this is a contradiction, because $m \geq 3$.

Putting these things together, we obtain that $\mathrm{I}(G, \Omega^\epsilon) < 2 \log t^\epsilon$. □

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

6.3 Cases involving graph automorphisms

We remarked after Definition 6.2 that the almost simple groups with socle $\mathrm{Sp}_4(2^a)$ or $\mathrm{P}\Omega_8^+(q)$ containing a graph automorphism or a triality graph automorphism were not covered by the definition. We briefly deal with these groups here.

LEMMA 6.12. *Let G be an almost simple group with socle $G_0 = \mathrm{Sp}_4(2^a)'$, and let G act faithfully and primitively on Ω, a set of size t. Then,*

$$\mathrm{I}(G,\Omega) < \frac{11}{3}\log t.$$

Proof. If $a = 1$, then $G \leq \mathrm{P\Gamma L}_2(9)$. It is not hard to verify that the longest strictly increasing chain of subgroups of $\mathrm{P\Gamma L}_2(9)$ has length 6. Thus, $\mathrm{I}(G) \leq 6$. On the other hand, $t \geq 6$, and so the result follows. Assume, then, that $a > 2$. Note that

$$G_0 \leq G \leq \mathrm{Aut}(\mathrm{Sp}_4(2^a)) = \Gamma\mathrm{Sp}_4(2^a).\langle\gamma\rangle,$$

where γ is a graph automorphism of order $2a$. We obtain that

$$\begin{aligned} |G| &\leq |\mathrm{Aut}(\mathrm{Sp}_4(2^a))| = |\mathrm{Sp}_4(2^a)|\cdot 2\cdot a \\ &= 2^{4a}(2^{2a}-1)(2^{4a}-1)\cdot 2\cdot a \leq 2^{11a}. \end{aligned} \tag{6.17}$$

On the other hand, [KL90, Theorem 5.2.2] implies that

$$t \geq \frac{2^{4a}-1}{2^a-1} = 2^{3a}+2^{2a}+2^a+1 > 2^{3a}. \tag{6.18}$$

Then, (6.17) and (6.18) yield

$$\mathrm{I}(G) \leq \log 2^{11a} = \log(2^{3a})^{\frac{11}{3}} = \frac{11}{3}\log 2^{3a} < \frac{11}{3}\log t.$$

□

LEMMA 6.13. *Let G be an almost simple group with socle $G_0 = \mathrm{P}\Omega_8^+(q)$, where $q = p^f$ for some prime p and positive integer f, and let G act primitively on Ω, a set of size t. Then,*

$$\mathrm{I}(G,\Omega) < \frac{16}{3}\log t.$$

Proof. Note that

$$G_0 \leq G \leq \mathrm{Aut}(\mathrm{P}\Omega_8^+(q)) = \mathrm{P\Gamma O}_8^+(q).\langle\tau\rangle,$$

with τ a triality graph automorphism of order 3. Now, using the fact that $6f < q^4$, and recalling that

$$\begin{aligned} |G_0| &= \frac{1}{d}q^{12}(q^4-1)\prod_{i=1}^{3}(q^{2i}-1), \\ |\mathrm{Out}(G_0)| &= 6df, \end{aligned}$$

where $d = (4, q^4-1)$, we obtain that

$$\begin{aligned} |G| &\leq |\mathrm{Aut}(G_0)| = |G_0|\cdot|\mathrm{Out}(G_0)| \\ &= 6df\cdot\tfrac{1}{d}q^{12}(q^4-1)(q^2-1)(q^4-1)(q^6-1) \\ &< 6fq^{28} < q^{32}. \end{aligned} \tag{6.19}$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

On the other hand, [KL90, Section 5.2] implies that

$$\begin{cases} t \geq \frac{(q^4-1)(q^3+1)}{(q-1)} = (q^3+q^2+q+1)(q^3+1) > q^6, & \text{for } q > 2; \\ t \geq 2^3(2^4-1) > 2^6, & \text{for } q = 2. \end{cases} \tag{6.20}$$

Then, (6.19) and (6.20) yield

$$\mathrm{I}(G) \leq \log q^{32} = \log(q^6)^{\frac{32}{6}} < \frac{16}{3}\log t. \qquad \square$$

Now Proposition 6.1 is a consequence of Lemmas 6.5–6.7, 6.12, and 6.13. Similarly, Theorem 1.6 is a consequence of Propositions 3.1, 4.1, 5.1, and 6.1 (observe that in Proposition 5.1 we are using the fact that $|\Delta| \geq 5$ and hence $3/(2\log|\Delta|) < 1$).

§7. Almost simple groups: Bounds on H(G)

Now that Theorem 1.6 is proved, we turn our attention to the proof of Theorem 1.3. In light of Theorem 1.6 and Proposition 5.1, all that is required is that we deal with the almost simple groups listed at item (1b) of Theorem 1.6—we must show that these conform to the bound given in Theorem 1.3. Thus, the result that we need is the following.

PROPOSITION 7.1. *Let G be an almost simple group with socle, G_0, a simple classical group. Let V be the associated natural module, of dimension n over a field $\mathbb{F}_q$, let m be an integer with $0 < m < n$, let Ω be a set of m-dimensional subspaces of V on which G acts primitively, and let $t = |\Omega|$. Then,*

$$\mathrm{H}(G) < \frac{17}{2}\log t.$$

7.1 $G_0 = \mathrm{PSL}_n(q)$ and elements of Ω are m-spaces

In this case, we suppose that $G_0 = \mathrm{PSL}_n(q)$, and that Ω is the set of all m-dimensional subspaces of V, where V is the natural n-dimensional module for G_0 over $\mathbb{F}_q$. Note that G is transitive on Ω. We prove the following result.

LEMMA 7.2. *Let G be an almost simple group with socle $\mathrm{PSL}_n(q)$ acting on Ω, the set of all m-spaces in V. Then,*

$$\mathrm{H}(G) < \frac{11}{2}\log t.$$

Moreover, if $q = p^f$ with p a prime number and $G \leq \mathrm{P\Gamma L}_n(q)$, then

$$\mathrm{H}(G) < 2\min(m, n-m)n + \log\log_p q. \tag{7.1}$$

Using the fact that duality conjugates m-space stabilizers to $(n-m)$-space stabilizers, in what follows, we may suppose that $m \leq n/2$. Two notes concerning the actions considered in Lemma 7.2. First, it is easy to check that

$$t > q^{m(n-m)}. \tag{7.2}$$

Second, we note that if $m < \frac{n}{2}$, then $G \leq \mathrm{P\Gamma L}_n(q)$. If $m = \frac{n}{2}$, then we must allow for the possibility that G contains a graph automorphism of $\mathrm{PSL}_n(q)$.

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

7.2 Some preliminaries

Let V and W be finite-dimensional vector spaces over $\mathbb{F}_q$. We denote by $\mathrm{End}^s_{\mathbb{F}_q}(V)$ the set of all semilinear transformations $V \to V$. Moreover, we write $\mathrm{Hom}_{\mathbb{F}_q}(V,W)$ for the set of all linear maps $V \to W$ and $\mathrm{Hom}^s_{\mathbb{F}_q}(V,W)$ for the set of all semilinear maps $V \to W$.

If $n = \dim(V)$, then we write $M_n(q)$ to denote the set of linear transformations $V \to V$; where a basis has been chosen, we also allow $M = M_n(q)$ to denote the set of all n-by-n matrices over $\mathbb{F}_q$. If $W_1, \ldots, W_k \leq V$, then we define

$$M_{W_1,\ldots,W_k} = \{g \in M \mid W_i g \leq W_i \text{ for all } i = 1, \ldots, k\}.$$

Using this definition, it is natural to extend the concept of base, irredundant base, and independent set in the context of the algebra $M = M_n(q)$. For instance, given a set $\Lambda := \{W_1, \ldots, W_k\}$ of subspaces of V, we say that Λ is an independent set for $M = M_n(q)$ if, for each $i \in \{1, \ldots, k\}$, we have

$$M_{W_1,\ldots,W_k} < M_{W_1,\ldots,W_{i-1},W_{i+1},\ldots,W_k}.$$

Recall that any element of $\mathrm{Aut}(\mathbb{F}_q)$ induces an automorphism of V: we first fix a basis for V and then act coordinatewise. The following lemma is standard, but we include it for completeness.

LEMMA 7.3. *Let $g \in Hom^s_{\mathbb{F}_q}(V,W)$, $g \neq 0$. Then:*

(a) *there exists a unique $\sigma \in Aut(\mathbb{F}_q)$ such that g is σ-semilinear. We say that σ is the associated automorphism of g;*
(b) *if σ is the associated automorphism of g, then $g\sigma^{-1} \in Hom_{\mathbb{F}_q}(V,W)$;*
(c) *there exists a unique $h \in Hom_{\mathbb{F}_q}(V,W)$ such that $g = h\sigma$.*

Proof.

(a) Assume that there exist $\sigma_1, \sigma_2 \in \mathrm{Aut}(\mathbb{F}_q)$ such that g is σ_1-semilinear and σ_2-semilinear. Hence, for any $v \in V$ and $k \in \mathbb{F}_q$, we have $(kv)g = k^{\sigma_1}(vg) = k^{\sigma_2}(vg)$. In particular, as $g \neq 0$, there exists $v_0 \in V$ such that $(v_0)g \neq 0$. Therefore, $k^{\sigma_1}(v_0 g) = k^{\sigma_2}(v_0 g)$, which implies $k^{\sigma_1} = k^{\sigma_2}$, for any $k \in \mathrm{Aut}\,(\mathbb{F}_q)$, and so $\sigma_1 = \sigma_2$.
(b) Let $v_1, v_2 \in V$ and $k \in \mathbb{F}_q$, then
   – $(v_1 + v_2)g\sigma^{-1} = (v_1)g\sigma^{-1} + (v_2)g\sigma^{-1}$;
   – $(kv_1)g\sigma^{-1} = (k^\sigma(v_1 g))\sigma^{-1} = k((v_1)g\sigma^{-1})$.
   Hence, $g\sigma^{-1}$ is linear.
(c) From (b), it follows that there exists $h \in \mathrm{Hom}_{\mathbb{F}_q}(V,W)$ such that $g\sigma^{-1} = h$. Hence, $g = h\sigma$. Assume that there exist $h_1, h_2 \in \mathrm{Hom}_{\mathbb{F}_q}(V,W)$ such that $g = h_1\sigma = h_2\sigma$. Then, for any $v \in V$, we have $(v)h_1\sigma = (v)h_2\sigma$, which implies that $(v)h_1 = (v)h_2$. So, we conclude that $h_1 = h_2$. □

Now, let $V = W \oplus X$, for some $W, X \leq V$. Then, every $v \in V$ can be written uniquely as $v = w + x$, for some $w \in W$ and $x \in X$. Let $f_W \in \mathrm{End}^s_{\mathbb{F}_q}(W)$ and $f_X \in \mathrm{End}^s_{\mathbb{F}_q}(X)$. We define $f_W \oplus f_X \in \mathrm{End}^s_{\mathbb{F}_q}(V)$ as $v(f_W \oplus f_X) = wf_W + xf_X$.

REMARK 7.4. Let $V = W \oplus X$, for some $W, X \leq V$. Let $g \in \mathrm{End}^s_{\mathbb{F}_q}(V)$. Then, g can be written as $g = g|_W \oplus g|_X$, provided $g(W) \leq W$ and $g(X) \leq X$.

Finally, if $v_1, \ldots, v_k \in V$, then we write $\langle v_1, \ldots, v_k \rangle_{\mathbb{F}_q}$ to denote the $\mathbb{F}_q$-span of $v_1, \ldots, v_k$.

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

7.3 $G \leq \mathbf{PGL}_n(q)$

Let V be n-dimensional over $\mathbb{F}_q$. In what follows, we write $\mathrm{M}_n(q)$ for the set of all linear transformations $V \to V$. Note that it will be convenient to swap between thinking of $\mathrm{GL}_n(q)$ or $\mathrm{PGL}_n(q)$, depending on context—this makes no difference to the calculations in question, as the center of $\mathrm{SL}_n(q)$ is the kernel of the action on Ω.

The following lemma is crucial.

LEMMA 7.5. *Let $\mathrm{SL}_n(q) \leq G \subset \mathrm{GL}_n(q)$ act on Ω. Let $\Lambda = \{W_1, \ldots, W_\ell\} \subseteq \Omega$ be an independent set for the action of G on Ω, of maximal size. Then,*

$$\dim \langle W_1, \ldots, W_\ell \rangle_{\mathbb{F}_q} > n - m.$$

Proof. Let $W = \langle W_1, \ldots, W_\ell \rangle_{\mathbb{F}_q}$ and suppose, first, that $k = \dim W = n - m$. Then, we can assume, by transitivity, that $W = \langle e_1, \ldots, e_{n-m} \rangle_{\mathbb{F}_q}$.

Hence,

$$G_{(\Lambda)} \leq \left\{ \left(\begin{array}{c|c} A & 0 \\ \hline B & C \end{array} \right) \mid A \in \mathrm{M}_{n-m}(q), B \in \mathrm{M}_{m,n-m}(q), C \in \mathrm{M}_m(q) \right\}.$$

Now, we add to Λ the following m-subspace:

$$X = \langle e_{n-m+1}, \ldots, e_n \rangle_{\mathbb{F}_q},$$

and we denote $\bar{\Lambda} = \Lambda \cup \{X\}$. Hence, by the maximality of Λ, we get

$$G_{(\bar{\Lambda})} = G_{(\bar{\Delta})},$$

for some $\bar{\Delta} \subseteq \bar{\Lambda}$.

We distinguish three cases:

Case I : if $\bar{\Delta} = \Lambda$, then $G_{(\bar{\Lambda})} = G_{(\Lambda)}$. Now,

$$G_{(\bar{\Lambda})} \leq \left\{ \left(\begin{array}{c|c} A & 0 \\ \hline 0 & C \end{array} \right) \mid A \in \mathrm{M}_{n-m}(q), C \in \mathrm{M}_m(q) \right\},$$

but

$$G_{(\Lambda)} \geq \left\{ \left(\begin{array}{c|c} I & 0 \\ \hline B & I \end{array} \right) \mid B \in \mathrm{M}_{m,n-m}(q), B \neq 0 \right\},$$

and this is a contradiction.

Case II : if $\bar{\Delta} \subset \Lambda$, then $G_{(\Lambda)} \leq G_{(\bar{\Delta})}$. On the other hand, $G_{(\bar{\Delta})} = G_{(\bar{\Lambda})} \leq G_{(\Lambda)}$, which implies $G_{(\Lambda)} = G_{(\bar{\Delta})}$, and hence a contradiction of the independence of Λ.

Case III : if $\bar{\Delta} \subsetneq \Lambda$, then $X \in \bar{\Delta}$ and we denote $\Delta = \bar{\Delta} \setminus \{X\}$. It is sufficient to prove that $G_{(\Lambda)} = G_{(\Delta)}$, as this leads to a contradiction. Observe that $V = W \oplus X$, and so, for any $g \in \mathrm{End}_{\mathbb{F}_q}(V)$, we can define

$$\begin{aligned} \pi_W : \mathrm{End}_{\mathbb{F}_q}(V) &\to \mathrm{Hom}_{\mathbb{F}_q}(W, V) \\ g &\mapsto g|_W \end{aligned}$$

and

$$\begin{aligned} \pi_X : \mathrm{End}_{\mathbb{F}_q}(V) &\to \mathrm{Hom}_{\mathbb{F}_q}(X, V) \\ g &\mapsto g|_X. \end{aligned}$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Now, if $g \in G_{(\bar{\Lambda})}$ or if $g \in G_{(\bar{\Delta})}$, then $g|_W \in \mathrm{End}_{\mathbb{F}_q}(W)$ and $g|_W \in \mathrm{End}_{\mathbb{F}_q}(X)$, which, by Remark 7.4, imply that$g = g|_W \oplus g|_X$. Therefore, we have that

$$g \in G_{(\bar{\Lambda})} \Leftrightarrow \begin{cases} g|_W \in \pi_W(G_{(\Lambda)}) \\ g|_X(X) = X \end{cases}$$

and

$$g \in G_{(\bar{\Delta})} \Leftrightarrow \begin{cases} g|_W \in \pi_W(G_{(\Delta)}) \\ g|_X(X) = X. \end{cases}$$

Hence, $G_{(\bar{\Lambda})} = G_{(\bar{\Delta})}$ implies

$$\pi_W(G_{(\Lambda)}) = \pi_W(G_{(\Delta)}). \tag{7.3}$$

On the other hand,

$$g \in G_{(\Lambda)} \Leftrightarrow g|_W \in \pi_W(G_{(\Lambda)})$$

and

$$g \in G_{(\Delta)} \Leftrightarrow g|_W \in \pi_W(G_{(\Delta)}).$$

Therefore, by (7.3), we get $G_{(\Lambda)} = G_{(\Delta)}$.

Finally, if $k < n - m$, then we can assume, by transitivity, that $W = \langle e_1, \dots, e_k \rangle_{\mathbb{F}_q}$ and consider $V_0 = \langle e_1, \dots, e_{k+m} \rangle_{\mathbb{F}_q} \leq V$. Let $X = \langle e_{k+1}, \dots, e_{k+m} \rangle_{\mathbb{F}_q}$. Hence, we can apply to V_0 the same argument as before, which leads to a contradiction. □

The next lemma yields Lemma 7.2 provided $G \leq \mathrm{PGL}_n(q)$.

LEMMA 7.6. *Let G be an almost simple group with socle $\mathrm{PSL}_n(q)$, such that $\mathrm{PSL}_n(q) \trianglelefteq G \leq PGL_n(q)$, acting on Ω, the set of all m-spaces in V. Then,*

$$\mathrm{H}(G) < 4 \log t.$$

Moreover,

$$\mathrm{H}(G) < 2 \min(m, n-m) n.$$

Proof. As usual, we may suppose $m \leq n/2$ and $G = \mathrm{PGL}_n(q)$. Moreover, from the opening paragraph of Section 7.3, we may work with the linear group $\mathrm{GL}_n(q)$ and hence we may suppose $G = \mathrm{GL}_n(q)$. Let Λ be an independent set for the action of G on Ω, of maximal size. The previous lemma allows us to order $\Lambda = \{W_1, \dots, W_k, \dots\}$, so that

$$\dim \langle W_1, \dots, W_j \rangle_{\mathbb{F}_q} > \dim \langle W_1, \dots, W_{j-1} \rangle_{\mathbb{F}_q},$$

for $j \leq k$, where $\dim(\Lambda) = \dim(\langle W_1, \dots, W_k \rangle_{\mathbb{F}_q}) > n - m$. Now, Λ is an independent set for $M = M_n(q)$ (see Section 7.2 for the definition of independent set for the algebra of n-by-n matrices over the finite field $\mathbb{F}_q$). Consider the following iterated process for $i \leq k$:

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Step 1. Let $W_1 \in \Omega$. We define $W = W_1$. We recall that we may assume, without loss of generality, that $W = \langle e_1, \ldots, e_m \rangle_{\mathbb{F}_q}$ and that

$$M_W = \left\{ \left(\begin{array}{c|c} A & 0 \\ \hline B & C \end{array} \right) \mid A \in \mathrm{M}_m(q), B \in \mathrm{M}_{n-m,m}(q), C \in \mathrm{M}_{n-m}(q) \right\}.$$

Step*i*. We have $W_1, \ldots, W_{i-1}$ and we define $W = \langle W_1, \ldots, W_{i-1} \rangle_{\mathbb{F}_q}$. Let $\dim W = h$ and assume, without loss of generality, that $W = \langle e_1, \ldots, e_h \rangle_{\mathbb{F}_q}$. Since we have chosen a specific ordering, we have that W_i satisfies

$$\dim \langle W_1, \ldots, W_i \rangle_{\mathbb{F}_q} > \dim W.$$

Consider V/W and $(W_i + W)/W \cong W_i/(W_i \cap W)$. Now let $\dim(W_i + W)/W = j$, observe that $j \in \{1, \ldots, m\}$. Then, we may assume

$$W_i = \langle v_1, \ldots, v_{m-j}, e_{h+1}, \ldots, e_{h+j} \rangle_{\mathbb{F}_q},$$

for some $v_1, \ldots, v_{m-j} \in W$ and $e_{h+1}, \ldots, e_{h+j} \notin W$.

Hence, $\dim(W \cap W_i) = m - j$, and we can write

$$W_i = (W \cap W_i) \oplus \langle e_{h+1}, \ldots, e_{h+j} \rangle_{\mathbb{F}_q}.$$

Let $r \in \{h+1, \ldots, h+j\}$, and let $g \in M_{W_1,\ldots,W_i}$; observe that $e_r^g \in W_i$. Hence, $e_r^g = u_1 + u_2$, for some $u_1 \in W \cap W_i$ and $u_2 \in \langle e_{h+1}, \ldots, e_{h+j} \rangle_{\mathbb{F}_q}$. Therefore, we have

$$\dim \langle \{ e_r^g \mid g \in M_{W_1,\ldots,W_i} \} \rangle_{\mathbb{F}_q} \leq m, \quad \text{for all } r = h+1, \ldots, h+j.$$

This iteration will end after k steps, and hence,

$$\dim M_{W_1,\ldots,W_k} \leq m(n-s) + ns,$$

where $s := n - \dim\langle W_1, \ldots, W_k \rangle$. We conclude that $|\Lambda| \leq k + m(n-s) + ns$, and so we have

$$\mathrm{H}(G) \leq k + m(n-s) + ns.$$

Now, observe that k is maximum when $j = 1$ at each step, in which case $k = n - m - s$. Hence,

$$\mathrm{H}(G) \leq (n - m - s) + ns + m(n-s). \tag{7.4}$$

By Lemma 7.5, we have $s \leq m - 1$. We conclude that the right-hand-side of (7.4) is maximized when $s = m - 1$. Then,

$$\mathrm{H}(G) \leq 2mn - m^2 - m + 1,$$

and, as $m \geq 1$ implies $m^2 + m - 1 > 0$, we conclude that

$$\mathrm{H}(G) < 2mn.$$

Now, observe that

$$\frac{mn}{m(n-m)} \leq 2 \Leftrightarrow m \leq \frac{n}{2}.$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Then, as $\log q \geq 1$, we have

$$\frac{mn}{m(n-m)} \leq 2 \leq 2\log q.$$

This implies that

$$2mn \leq 4\log q^{m(n-m)}.$$

Hence, using (7.2), we obtain

$$\mathrm{H}(G) < 2mn < 4\log t. \qquad \square$$

Note that, in Bianca Lodà's thesis, a precise version of Lemma 7.6 is given for the case $m = 1$. It turns out that, for $\mathrm{PSL}_n(q) \trianglelefteq G \leq \mathrm{PGL}_n(q)$ acting on Ω, the set of all 1-spaces in V, we have

$$\mathrm{H}(G,\Omega) = \begin{cases} n, & \text{if} q = 2; \\ 2n-2, & \text{if} q > 2. \end{cases}$$

7.4 $G \not\leq \mathbf{PGL}_n(q)$

We can finally prove Lemma 7.2.

Proof of Lemma 7.2. Let $q = p^f$. Suppose, first, that

$$\mathrm{PSL}_n(q) \trianglelefteq G \leq \mathrm{P\Gamma L}_n(q) = \mathrm{PGL}_n(q) \rtimes \mathcal{C}_f.$$

Now, Lemma 2.8 implies that

$$\mathrm{H}(G) \leq \mathrm{H}(\mathrm{PGL}_n(q)) + \ell(\mathcal{C}_f),$$

where $\ell(\mathcal{C}_f)$ is the maximum length of a strictly descending chain of subgroups in $\mathcal{C}_f$. Since $\ell(\mathcal{C}_f) \leq \log f = \log\log_p q$, we obtain

$$\mathrm{H}(G) < \mathrm{H}(\mathrm{PGL}_n(q)) + \log\log_p q.$$

From Lemma 7.6, we immediately deduce (7.1) when $m < n/2$ and we also get

$$\mathrm{H}(G) < 4\log t + \log\log_p q.$$

By (7.2), we easily conclude $\log\log_p q < \log t$ and hence

$$\mathrm{H}(G) < 5\log t.$$

We must deal with the possibility that $G \not\leq \mathrm{P\Gamma L}_n(q)$, that is, G contains a graph automorphism. In this case, $n \geq 4$, $m = \frac{n}{2}$, and G contains a subgroup, H, of index 2 that lies in $\mathrm{P\Gamma L}_n(q)$. By the argument above, we know that

$$\mathrm{H}(H) < 5\log t.$$

Now, Lemma 2.8 implies that

$$\mathrm{H}(G) < 5\log t + 1 < \frac{11}{2}\log t. \qquad \square$$

We remark that in Lemma 7.6 we chose a particular ordering of our independent set Λ when we came to study the associated stabilizer chain. We can do this because we are

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

studying $\mathrm{H}(G)$ rather than $\mathrm{I}(G)$—it is precisely this ordering step which has prevented us from proving a strong enough version of Theorem 1.6 to confirm Conjecture 1.7.

7.5 $G_0 = \mathrm{PSL}_n(q)$ and elements of Ω are pairs of subspaces

In this section, we consider the primitive subspace actions of those almost simple groups with socle $\mathrm{PSL}_n(q)$ that contain an automorphism of the Dynkin diagram.

Let $n \geq 3$, and fix m, an integer satisfying $1 \leq m < \frac{n}{2}$. We denote by Ω_m the set of all m-subspaces of V. We consider the action of G on $\Omega^{(i)}$, for $i = 1, 2$, where

$$\Omega^{(1)} = \left\{ \{U, W\} \mid U, W \leq V, \dim U = m, \dim W = n-m \text{ with } m < \frac{n}{2} \text{ and } U \oplus W = V \right\},$$
$$\Omega^{(2)} = \left\{ \{U, W\} \mid U, W \leq V, \dim U = m, \dim W = n-m \text{ with } m < \frac{n}{2} \text{ and } U < W \right\}.$$

The main result of this section is the following.

LEMMA 7.7. *Let G be an almost simple group with socle $\mathrm{PSL}_n(q)$ acting on $\Omega^{(1)}$ or $\Omega^{(2)}$. Then,*

$$\mathrm{H}(G) < \frac{17}{2} \log t.$$

We set $t^{(i)} := |\Omega^{(i)}|$, for $i = 1, 2$. Then, we have

$$t^{(1)} = \frac{q^{m(n-m)} \prod_{i=1}^{n} (q^i - 1)}{\prod_{i=1}^{m} (q^i - 1) \prod_{i=1}^{n-m} (q^i - 1)} \quad \text{and} \quad t^{(2)} = \frac{\prod_{i=1}^{n} (q^i - 1)}{\prod_{i=1}^{m} (q^i - 1)^2 \prod_{i=1}^{n-2m} (q^i - 1)}.$$

We saw already, in (7.2), that $|\Omega_m| > q^{m(n-m)}$. We obtain immediately that

$$t^{(i)} \geq q^{m(n-m)}. \tag{7.5}$$

For the time being, for various technical reasons, we consider $G \leq \mathrm{P\Gamma L}_n(q)$ (although the corresponding actions are not primitive any longer). We simultaneously study the action of G on $\Omega^{(1)}$ and $\Omega^{(2)}$. Let $\Lambda \subset \Omega^{(i)}$ be an independent set for the action of G on $\Omega^{(i)}$, with $k := |\Lambda|$. We define

$$\mathcal{U} = \{U \in \Omega_m \mid \text{there exists } W \in \Omega_{n-m} \text{ with } \{U, W\} \in \Lambda\}; \tag{7.6}$$

$$\mathcal{W} = \{W \in \Omega_{n-m} \mid \text{there exists } U \in \Omega_m \text{ with } \{U, W\} \in \Lambda\}. \tag{7.7}$$

Let $k_U := |\mathcal{U}|$ and $k_W := |\mathcal{W}|$. Clearly, $k_U, k_W \leq |\Lambda| = k$.

Notice that, since $G \leq \mathrm{P\Gamma L}_n(q)$, G acts naturally on Ω_m and Ω_{n-m}; however, observe that, *a priori*, the independence of Λ does not imply the independence of Λ_U and Λ_W. The result that we need is the following.

LEMMA 7.8. $k \leq \mathrm{H}(G, \Omega_m) + \mathrm{H}(G, \Omega_{n-m})$.

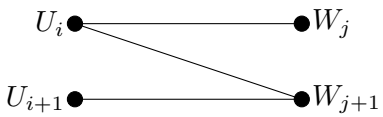
In order to prove Lemma 7.8, we first need to define a graph Γ associated with Λ in the following way: $\Gamma = (V, \Lambda)$, where $V = \mathcal{U} \cup \mathcal{W}$ is the vertex set and the edge set is Λ.

From here, we will first prove a general fact about the length of paths in the graph Γ, and we will then split our proof of Lemma 7.8 into two steps: the first is a special case, which illustrates the general case; this general case will come second.

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

LEMMA 7.9. *The graph* Γ *can only have paths of length at most 2.*

Proof. Assume that in Γ we have a path of length at least 3. Then, relabeling if necessary, we have



for some $1 \leq i < k_U$, $1 \leq j < k_W$. As $G \leq \mathrm{P\Gamma L}_n(q)$, G preserves the two parts Λ_U and Λ_W of the natural bipartition of Γ. Hence,

$$G_{\{U_i,W_j\},\{U_{i+1},W_{j+1}\}} = G_{U_i} \cap G_{W_j} \cap G_{U_{i+1}} \cap G_{W_{j+1}} = G_{\{U_i,W_j\},\{U_i,W_{j+1}\},\{U_{i+1},W_{j+1}\}}.$$

As $\{U_i,W_j\}$, $\{U_i,W_{j+1}\}$, $\{U_{i+1},W_{j+1}\}$ are elements in Λ, this contradicts the independence of Λ. □

7.6 Suppose that Γ is a complete matching

In other words, we suppose Γ has no paths of length 2. This case will illustrate the general argument very well.

In this case, $\Lambda = \{\{U_1,W_1\},\dots,\{U_k,W_k\}\}$. Let us consider the action of G on Ω_m. By Lemma 2.4, there exists $\mathcal{I} \subset \{1,\dots,k\}$, say $\mathcal{I} = \{i_1,\dots,i_s\}$, such that $\{U_{i_1},\dots,U_{i_s}\}$ is an independent set and

$$\bigcap_{j=1}^{k} G_{U_i} = \bigcap_{j=1}^{s} G_{U_{i_j}}.$$

After relabeling elements, we can assume $\mathcal{I} = \{1,\dots,s\}$. This yields

$$s \leq \mathrm{H}(G,\Omega_m).$$

Let

$$H := \bigcap_{i=1}^{s} (G_{U_i} \cap G_{W_i})$$

and consider its action on $\Omega^{(i)}$, for $i \in \{1,2\}$. Recall that, by Lemma 2.2, the set

$$\Delta = \{\{U_{s+1},W_{s+1}\},\dots,\{U_k,W_k\}\}$$

is independent with respect to H. Now, consider the action of H on Ω_{n-m}.

LEMMA 7.10. *The set* $\{W_{s+1},\dots,W_k\}$ *is independent with respect to the action of* H *on* Ω_{n-m}.

Proof. Assume, for contradiction, that the set $\{W_{s+1},\dots,W_k\}$ is not independent with respect to H. Then, by Lemma 2.3, there would exist $j \in \{s+1,\dots k\}$ such that

$$\bigcap_{i=s+1}^{k} H_{W_i} = \bigcap_{\substack{i=s+1 \\ i\neq j}}^{k} H_{W_i}.$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Observe that

$$\bigcap_{i=1}^{s} G_{U_i} = \bigcap_{i=1}^{k} G_{U_i} = \bigcap_{\substack{i=1 \\ i\neq j}}^{k} G_{U_i}.$$

Therefore, we have

$$\begin{aligned} G_{(\Lambda)} &= H \cap \bigcap_{i=s+1}^{k} G_{W_i} = \bigcap_{i=s+1}^{k} H_{W_i} = \bigcap_{\substack{i=s+1 \\ i\neq j}}^{k} H_{W_i} \\ &= H \cap \left(\bigcap_{\substack{i=s+1 \\ i\neq j}}^{k} G_{W_i} \right) = \left(\bigcap_{i=1}^{s} G_{U_i} \cap \bigcap_{i=1}^{s} G_{W_i} \right) \cap \left(\bigcap_{\substack{i=s+1 \\ i\neq j}}^{k} G_{W_i} \right) \\ &= \left(\bigcap_{\substack{i=1 \\ i\neq j}}^{k} G_{U_i} \right) \cap \left(\bigcap_{\substack{i=1 \\ i\neq j}}^{k} G_{W_i} \right) = \bigcap_{\substack{i=1 \\ i\neq j}}^{k} (G_{U_i} \cap G_{W_i}) \\ &= G_{(\Lambda \setminus \{\{U_j, W_j\}\})}. \end{aligned}$$

This contradicts the independence of Λ. □

The previous lemma implies that

$$k - s \leq \mathrm{H}(H, \Omega_{n-m}).$$

Now, by Lemma 2.2, we have $\mathrm{H}(H, \Omega_{n-m}) \leq \mathrm{H}(G, \Omega_{n-m})$. Putting these things together yields that

$$k \leq \mathrm{H}(G, \Omega_m) + \mathrm{H}(G, \Omega_{n-m}),$$

and Lemma 7.8 is proved in this special case.

7.7 Suppose that Γ is not a complete matching

The general argument is very similar but requires some more notation. After reordering the vertices of Γ and using a suitable labeling, the fact that Γ has no paths of length 3 implies that Γ is isomorphic to the graph in Figure 1, where $\ell_1, \ldots, \ell_b, s_1, \ldots, s_c \geq 2$.

Observe that, by definition, we have k_U vertices on the left, k_W vertices on the right, and k edges, because these are the elements of Λ. Let us consider the action of G on Ω_m, that is, we focus on the action of G on the vertices on the left-hand side of the graph. By Lemma 2.4, there exists a subset $\mathcal{U}_1$ of $\mathcal{U}$, such that $\mathcal{U}_1$ is an independent set and

$$\bigcap_{U \in \mathcal{U}} G_U = \bigcap_{U \in \mathcal{U}_1} G_U. \tag{7.8}$$

Let $s := |\mathcal{U}_1|$. We reorder the vertices on the left-hand side of the graph, so that the elements of $\mathcal{U}_1$ occur in the top s positions. Simultaneously, we take to the top exactly one edge, chosen arbitrarily, having one end in U, for each $U \in \mathcal{U}_1$. Note that here we allow crossings between edges. Observe that Lemma 7.9 implies that this operation is well defined. We have

$$s \leq \mathrm{H}(G, \Omega_m).$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

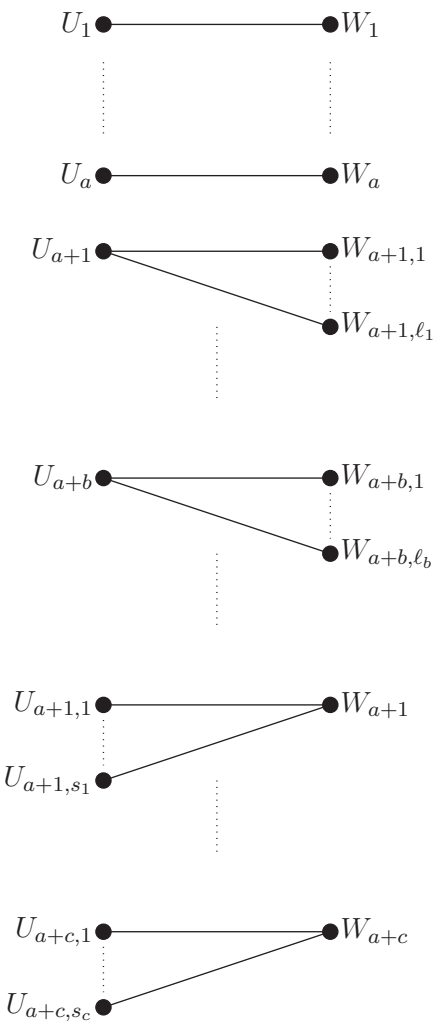


Figure 1.
The graph Γ in the general case.

Let $\Lambda_{\mathcal{U}_1} \subset \Lambda$ be the set of those chosen edges and note that $|\Lambda_{\mathcal{U}_1}| = s$. We define

$$\mathcal{W}_1 = \{W \in \mathcal{W} \mid W \text{ is an end point for an edge in } \Lambda_{\mathcal{U}_1}\},$$
$$\mathcal{W}_2 = \mathcal{W} \setminus \mathcal{W}_1.$$

Observe that $|\mathcal{W}_1| \leq s$.

LEMMA 7.11. *Let $H := G_{(\Lambda_{\mathcal{U}_1})}$. Let $\lambda \in \Lambda \setminus \Lambda_{\mathcal{U}_1}$, and let U_λ and W_λ be, respectively, its left-hand end point and right-hand end point. Then, we have that*

1. *the function*

$$f \colon \Lambda \setminus \Lambda_{\mathcal{U}_1} \to \mathcal{W}_2$$
$$\lambda \mapsto W_\lambda$$

is a bijection. In particular, $|\mathcal{W}_2| = |\Lambda \setminus \Lambda_{\mathcal{U}_1}| = k - s$;
2. *$G_{(\Lambda)} = H_{(\mathcal{W}_2)}$.*

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Proof. To prove claim (1), we first prove that f is well defined. We claim that $W_\lambda \in \mathcal{W}_2$. Indeed, if $W_\lambda \in \mathcal{W}_1$, then

$$\begin{aligned} G_{(\Lambda_{\mathcal{U}_1})} &= \bigcap_{\substack{U\in\mathcal{U}_1\\ W\in\mathcal{W}_1}} (G_U \cap G_W) \\ &= \left(\bigcap_{U\in\mathcal{U}_1\cup\{U_\lambda\}} G_U\right) \cap \left(\bigcap_{W\in\mathcal{W}_1} G_W\right) \\ &= G_{(\Lambda_{\mathcal{U}_1}\cup\{\lambda\})}, \end{aligned}$$

which contradicts the independence of Λ. This fact implies that f is well defined. Moreover, it is onto by definition. Finally, we prove that f is one-to-one. Assume that there exist two distinct edges $\lambda, \mu \in \Lambda \setminus \Lambda_{\mathcal{U}_1}$ such that $W_\lambda = W_\mu$. Then, by using (7.8), we have

$$\begin{aligned} G_{(\Lambda_{\mathcal{U}_1}\cup\{\lambda,\mu\})} &= \bigcap_{\substack{U\in\mathcal{U}_1\\ W\in\mathcal{W}_1}} (G_U \cap G_W) \cap (G_{U_\lambda} \cap G_{W_\lambda}) \cap \left(G_{U_\mu} \cap G_{W_\mu}\right) \\ &= \left(\bigcap_{U\in\mathcal{U}_1\cup\{U_\lambda\}\cup\{U_\mu\}} G_U\right) \cap \left(\bigcap_{W\in\mathcal{W}_1\cup\{W_\lambda\}} G_W\right) \\ &= \left(\bigcap_{U\in\mathcal{U}_1\cup\{U_\lambda\}} G_U\right) \cap \left(\bigcap_{W\in\mathcal{W}_1\cup\{W_\lambda\}} G_W\right) \\ &= G_{(\Lambda_{\mathcal{U}_1}\cup\{\lambda\})}, \end{aligned}$$

which is a contradiction. Hence, f is a bijection, in particular $|\mathcal{W}_2| = |\Lambda \setminus \Lambda_{\mathcal{U}_1}| = k - s$.

For part (2), by using the previous result and (7.8), we have

$$\begin{aligned} G_{(\Lambda)} = G_{(\Lambda_{\mathcal{U}_1})} \cap G_{(\Lambda\setminus\Lambda_{\mathcal{U}_1})} &= H \cap G_{(\Lambda\setminus\Lambda_{\mathcal{U}_1})} \\ &= H \cap \left(\bigcap_{\lambda\in\Lambda\setminus\Lambda_{\mathcal{U}_1}} G_{U_\lambda} \cap G_{W_\lambda}\right) \\ &= \left(H \cap \bigcap_{\lambda\in\Lambda\setminus\Lambda_{\mathcal{U}_1}} G_{U_\lambda}\right) \cap \left(H \cap \bigcap_{\lambda\in\Lambda\setminus\Lambda_{\mathcal{U}_1}} G_{W_\lambda}\right) \\ &= H \cap \left(\bigcap_{W\in\mathcal{W}_2} G_W\right) = H_{(\mathcal{W}_2)}. \end{aligned}$$

∎

Now, consider the action of H on Ω_{n-m}, namely we focus on the vertices on the right-hand side of the graph. As a consequence of the previous lemma, we obtain the following corollary.

COROLLARY 7.12. *The set $\mathcal{W}_2$ is independent for the action of H on Ω_{n-m}.*

Proof. Assume not, then there would exist $\lambda \in \Lambda \setminus \Lambda_{\mathcal{U}_1}$ such that

$$H_{(\mathcal{W}_2)} = \bigcap_{W\in\mathcal{W}_2} H_W = \bigcap_{\substack{W\in\mathcal{W}_2\\ W\neq W_\lambda}} H_W.$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

If $U_\lambda \in \mathcal{U}_1$, then there exists $\mu \in \Lambda_{\mathcal{U}_1}$ such that $U_\lambda = U_\mu$. If $U_\lambda \notin \mathcal{U}_1$, then $\bigcap_{U\in\mathcal{U}} G_U = \bigcap_{U\in\mathcal{U},U\neq U_\lambda} G_U$.

In either case, this, along with (7.8) and part (2) of the previous lemma, implies that

$$\begin{aligned} G_{(\Lambda\setminus\{\lambda\})} &= \left(\bigcap_{U\in\mathcal{U}} G_U\right) \cap \left(\bigcap_{\substack{W\in\mathcal{W}\\ W\neq W_\lambda}} G_W\right) \\ &= \left(\bigcap_{U\in\mathcal{U}_1} G_U \cap \bigcap_{W\in\mathcal{W}_1} G_W\right) \cap \left(\bigcap_{\substack{W\in\mathcal{W}_2\\ W\neq W_\lambda}} G_W\right) \\ &= H \cap \bigcap_{\substack{W\in\mathcal{W}_2\\ W\neq W_\lambda}} G_W \\ &= \bigcap_{\substack{W\in\mathcal{W}_2\\ W\neq W_\lambda}} H_W = H_{(\mathcal{W}_2)} = G_{(\Lambda)}. \end{aligned}$$

This contradicts the independence of Λ. □

The previous lemma implies that

$$k - s \leq \mathrm{H}(H, \Omega_{n-m}).$$

Now, by Lemma 2.2, we have $\mathrm{H}(H,\Omega_{n-m}) \leq \mathrm{H}(G,\Omega_{n-m})$. Putting these things together yields that

$$k \leq \mathrm{H}(G,\Omega_m) + \mathrm{H}(G,\Omega_{n-m}),$$

and Lemma 7.8 is proved in the general case. We are ready to prove Lemma 7.7.

Proof of Lemma 7.7. Suppose, first, that $G \leq \mathrm{P\Gamma L}_n(q)$. Recall that the actions of G on Ω_m and Ω_{n-m} are permutation isomorphic. Thus, Lemma 7.8 and (7.1) imply that

$$\mathrm{H}(G,\Omega^{(i)}) < 4mn + 2\log\log_p q.$$

Here, we are tacitly assuming that $m < n/2$, because, for instance, (7.1) does use the hypothesis that $m < n/2$. However, this is not a restriction, because the duality conjugates the m-space stabilisers in $(n-m)$-space stabilisers.

Now, let $G \not\leq \mathrm{P\Gamma L}_n(q)$, that is, G contains the inverse transpose automorphism. Let $H = G \cap \mathrm{P\Gamma L}_n(q)$. Then, H is a normal subgroup of G of index 2. Then, by Lemma 2.8, we have

$$\mathrm{H}(G,\Omega^{(i)}) \leq \mathrm{H}(H,\Omega^{(i)}) + 1,$$

and therefore,

$$\mathrm{H}(G,\Omega^{(i)}) < 4mn + 2\log\log_p q + 1. \tag{7.9}$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

In view of (7.5), to prove the result, it is sufficient to check that

$$4mn+2\log\log_p q+1<\frac{17}{2}\log(q^{m(n-m)}).$$

The result follows directly. □

7.8 The other classical groups

In this subsection, we suppose that $G_0=\mathrm{Cl}(V)$ is one of the other classical groups defined on a vector space V of dimension n over $\mathbb{F}_q$. In this case, m is an integer such that $0<m<n$ and Ω is a G-orbit on the set of totally isotropic/totally singular/nondegenerate subspaces of dimension m in V. As usual, we set $t=|\Omega|$.

The result that we need is the following.

LEMMA 7.13. *Let G be an almost simple group with socle $Cl_n(q)$, and consider G as a permutation group in a subspace action. Then,*

$$\mathrm{H}(G)<\frac{17}{2}\log t.$$

In [BG16, Table 4.1.2], we find a list of the degrees of all such actions and, using the notation just established, one can easily verify the following fact.

LEMMA 7.14. *Either $G_0\cong \mathrm{P}\Omega_n^+(q)$ and $m=\frac{n}{2}$ or $t>q^{\frac{1}{2}m(n-m)}$.*

Proof. The proof of this lemma is rather tedious: it is simply a case-by-case analysis of the cases described in Table 4.1.2 of [BG16].

We make a couple of brief remarks about the proof and leave the details to the reader. Observe, first, that for a number of the cases (III, X with q odd, XI with q odd, XIII, XV, and XVI), t is divisible by $q^{\frac{1}{2}m(n-m)}$, and so the result follows immediately.

For the rest, one must demonstrate the bound directly. To illustrate, we consider Case VI which is where the exception occurs; in this case, $G_0\cong \mathrm{P}\Omega_n^+(q)$ and $m\leq\frac{n}{2}$. Then, we have

$$\begin{aligned}
t&=\frac{(q^{n/2}-1)(q^{(n-2m)/2}+1)\prod_{i=(n-2m)/2+1}^{(n-2m)/2+(m-1)}(q^{2i}-1)}{\prod_{i=1}^{m}(q^i-1)}\\
&=\frac{(q^{n/2}-1)(q^{(n-2m)/2}+1)}{q^m-1}\cdot\prod_{i=1}^{m-1}\left(\frac{q^{n-2m+2i}-1}{q^i-1}\right)\\
&>\frac{(q^{n/2}-1)(q^{(n-2m)/2}+1)}{q^m-1}\cdot\prod_{i=1}^{m-1}q^{n-2m+i}\\
&>q^{n/2-m+n/2-m}\cdot q^{(n-2m)(m-1)+\frac{1}{2}(m-1)m}\\
&=q^{nm-\frac{3}{2}m^2-m}.
\end{aligned}$$

Now, it is easy to check that the bound is satisfied if and only if $m<\frac{n}{2}$, as required. □

We are ready to prove Lemma 7.13.

Proof of Lemma 7.13. Observe, first, that $G\leq \mathrm{P\Gamma L}_n(q)$ and that the action we are studying is an action on Γ, the set of m-spaces. Then, Lemma 2.2 implies that

$$\mathrm{H}(G,\Omega)\leq \mathrm{H}(\mathrm{P\Gamma L}_n(q),\Gamma).$$

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

In Lemma 7.2, we give an upper bound for $\mathrm{H}(\mathrm{P\Gamma L}_n(q),\Gamma)$:

$$\mathrm{H}(G) < 2\min(m,n-m)n + \log\log_p q.$$

Thus, in light of Lemma 7.14, provided we do not have $(G_0,m) = (\mathrm{P\Omega}_n^+(q), \frac{n}{2})$, it is enough to prove that

$$2\min(m,n-m)n + \log\log_p q < \frac{17}{2}\log(q^{\frac{1}{2}m(n-m)}).$$

This is easily verified. Assume that $(G_0,m) = (\mathrm{P\Omega}_n^+(q), \frac{n}{2})$. In this case, $t \geq q^{\frac{n^2}{8}-\frac{n}{4}}$. Lemma 6.13 implies that we can assume that $n \geq 10$, and thus,

$$|G| \leq 2fq^{n(n-2)/4}(q^{n/2}-1)\prod_{i=1}^{n/2-1}(q^{2i}-1) < 2fq^{\frac{n^2}{2}-\frac{n}{2}} \leq q^{\frac{n^2}{2}-\frac{n}{2}+1}.$$

Now, observe that

$$\mathrm{H}(G) \leq \log|G| < \log(q^{\frac{n^2}{2}-\frac{n}{2}+1}) < 5\log(q^{\frac{n^2}{8}-\frac{n}{4}}) < 5\log t,$$

and we are done. □

Proposition 7.1 is a consequence of Lemmas 7.2, 7.7, and 7.13.

Proof of Theorem 1.3. Observe that $17/2 + 3/(2x) < 9$, as long as $x > 3$ (so $|\Delta| > 8$). With this remark, the proof follows from Theorem 1.6, and Propositions 5.1 and 7.1, except when G is a primitive group of PA type with $G \leq H \wr \mathrm{Sym}(\ell)$ and the domain of G is $\Omega = \Delta^\ell$ and $|\Delta| \leq 8$. If $H = \mathrm{Alt}(\Delta)$ or $H = \mathrm{Sym}(\Delta)$, then G is a large–base group, and hence, we may omit these cases for the rest of the proof. As H is a primitive group of degree $5, 6, 7$, or 8 and as H is neither $\mathrm{Alt}(\Delta)$ nor $\mathrm{Sym}(\Delta)$, we deduce that one of the following holds:

- $|\Delta| = 6$ and $H = \mathrm{PSL}_2(5)$ or $H = \mathrm{PGL}_2(5)$ in its natural action on the projective line,
- $|\Delta| = 7$ and $H = \mathrm{PSL}_3(2)$ in its natural action on the projective plane,
- $|\Delta| = 8$ and $H = \mathrm{PSL}_2(7)$ or $H = \mathrm{PGL}_2(7)$ in its natural action on the projective line.

In each of these groups, we have $\mathrm{H}(H) \leq 3$, and hence, from (5.1), we have $\mathrm{H}(G) \leq 3\ell + 3\ell/2 = 9\ell/2$. As $9\ell/2 < 9\log(t)$, the proof is completed. □

Acknowledgments. The results in this paper are based on the Ph.D. thesis of Bianca Lodà. The authors would like to thank Colva Roney-Dougal and Dugald Macpherson, who, in the course of examining the thesis, made many useful remarks which have helped to improve the current paper.

References

[Bab81] L. Babai, *On the order of uniprimitive permutation groups*, Ann. Math. (2) **113** (1981), 553–568.
[Bla92] K. D. Blaha, *Minimum bases for permutation groups: The greedy approximation*, J. Algorithms **13** (1992), 297–306.
[BHR13] J. N. Bray, D. F. Holt, and C. M. Roney-Dougal, *The Maximal Subgroups of the Low-Dimensional Finite Classical Groups*, Cambridge University Press, Cambridge, MA, 2013.
[Bur07] T. C. Burness, *On base sizes for actions of finite classical groups*, J. Lond. Math. Soc. (2) **75** (2007), 545–562.
[BG16] T. C. Burness and M. Giudici, *Classical Groups, Derangements and Primes*, Aust. Math. Soc. Lect. Ser. **25**, Cambridge University Press, Cambridge, MA, 2016.

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

[BGS11] T. C. Burness, R. M. Guralnick, and J. Saxl, *On base sizes for symmetric groups*, Bull. Lond. Math. Soc. **43** (2011), 386–391.
[BLS09] T. C. Burness, M. W. Liebeck, and A. Shalev, *Base sizes for simple groups and a conjecture of Cameron*, Proc. Lond. Math. Soc. (3) **98** (2009), 116–162.
[BOW10] T. C. Burness, E. A. O'Brien, and R. A. Wilson, *Base sizes for sporadic simple groups*, Israel J. Math. **177** (2010), 307–333.
[Cam92] P. J. Cameron, "Some open problems on permutation groups" in *Groups, Combinatorics and Geometry. Proceedings of the L.M.S. Durham Symposium, Held July 5–15, 1990 in Durham, UK*, Cambridge University Press, Cambridge, MA, 1992, 340–350.
[CK97] P. J. Cameron and W. M. Kantor, "Random permutations: Some group-theoretic aspects" in *Combinatorics, Geometry and Probability. A Tribute to Paul Erdős. Proceedings of the Conference Dedicated to Paul Erdős on the Occasion of His 80th Birthday, Cambridge, UK, 26 March 1993*, Cambridge University Press, Cambridge, MA, 1997, 139–144.
[CST89] P. J. Cameron, R. Solomon, and A. Turull, *Chains of subgroups in symmetric groups*, J. Algebra **127** (1989), 340–352.
[Che16] G. Cherlin, *On the relational complexity of a finite permutation group*, J. Algebraic Combin. **43** (2016), 339–374.
[CMS96] G. L. Cherlin, G. A. Martin, and D. H. Saracino, *Arities of permutation groups: Wreath products andk-sets*, J. Combin. Theory Ser. A **74** (1996), 249–286.
[DM96] J. D. Dixon and B. Mortimer, *Permutation Groups*, Grad. Texts Math. **163**, Springer, New York, 1996.
[GMPS15] S. Guest, J. Morris, C. E. Praeger, and P. Spiga, *On the maximum orders of elements of finite almost simple groups and primitive permutation groups*, Trans. Amer. Math. Soc. **367** (2015), no. 11, 7665–7694.
[KL90] P. Kleidman and M. Liebeck, *The Subgroup Structure of the Finite Classical Groups*, Lond. Math. Soc. Lect. Note Ser. **129**, Cambridge University Press, Cambridge, MA, 1990.
[Lan02] S. Lang, *Algebra*, 3rd revised ed., Grad. Texts Math. **211**, Springer, New York, 2002.
[Lie84] M. W. Liebeck, *On minimal degrees and base sizes of primitive permutation groups*, Arch. Math. **43** (1984), 11–15.
[LS99] M. W. Liebeck and A. Shalev, *Simple groups, permutation groups, and probability*, J. Amer. Math. Soc. **12** (1999), 497–520.
[MRD] M. Moscatiello and C. Roney-Dougal, *Base sizes of primitive permutation groups*, preprint available at https://arxiv.org/abs/2106.01219.
[Pra90] C. E. Praeger, *The inclusion problem for finite primitive permutation groups*, Proc. Lond. Math. Soc. (3) **60** (1990), 68–88.
[Qui04] M. Quick, *Probabilistic generation of wreath products of non-abelian finite simple groups*, Comm. Algebra **32** (2004), 4753–4768.

Nick Gill
Department of Mathematics,
University of South Wales,
Treforest CF37 1DL,
United Kingdom
nick.gill@southwales.ac.uk

Bianca Lodà
Department of Mathematics,
University of South Wales,
Treforest CF37 1DL,
United Kingdom
bianca.loda@southwales.ac.uk

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press

Pablo Spiga
Dipartimento di Matematica e Applicazioni,
University of Milano-Bicocca,
Via Cozzi 55, 20125 Milano,
Italy
pablo.spiga@unimib.it

https://doi.org/10.1017/nmj.2021.6 Published online by Cambridge University Press